

Standard Operating Procedure (SOP) 11: Data QA/QC, Formatting and Archiving, Version 1.0

Dall's Sheep Monitoring Protocol

Arctic (ARCN) and Central Alaska (CAKN) Inventory and Monitoring Networks

Kumi L. Rattenbury, Joshua H. Schmidt and Scott D. Miller

National Park Service
4175 Geist Road
Fairbanks, Alaska 99709

March 2017

U.S. Department of the Interior
National Park Service
Natural Resource Stewardship and Science
Fort Collins, Colorado

Please cite this Standard Operating Procedure as:

Rattenbury, K. L., Schmidt, J. H., and S. D. Miller. 2017. Standard Operating Procedure (SOP) 11: Data QA/QC, Formatting and Archiving, Version 1.0. Dall's Sheep Monitoring Protocol, Arctic (ARCN) and Central Alaska (CAKN) Inventory and Monitoring Networks. National Park Service, Fairbanks, Alaska. Available online at <https://irma.nps.gov/DataStore/Reference/Profile/2238931>.

The main Narrative for the Protocol for Monitoring Dall's Sheep, Arctic (ARCN) and Central Alaska (CAKN) Inventory and Monitoring Networks is available from the NPS Inventory and Monitoring Arctic Network website (<http://science.nature.nps.gov/im/units/arcn/>), the NPS I&M Central Alaska Network website (<http://science.nature.nps.gov/im/units/cakn/>) and the Natural Resource Publications Management website (<http://www.nature.nps.gov/publications/nrpm/>).

SOP 11 – Data QA/QC, Formatting and Archiving

Standard Operating Procedure #11
Version 1.0 (03/01/2017)

Revision History Log:

Previous Version #	Revision Date	Author	Changes Made	Reason for Change	New Version #

Add rows as needed for each change or set of changes associated with each version.

SOP 11 – Contents

SOP 11 - Figures.....	5
SOP 11 – Tables.....	6
SOP 11.1 – Introduction	7
SOP 11.2 – Post Field Campaign Data Handling	7
SOP 11.3 – Initial Quality Assurance/Control Procedures	9
SOP 11.4 – Geodatabase Editing.....	10
SOP 11.5 – Recalculating Group Distances	15
SOP 11.6 – Creating Buffer Files	16
SOP 11.7 – Importing Field Data to the Master Sheep Monitoring Database.....	28
SOP 11.7.1 Overview	28
SOP 11.7.2 Creating a new Survey Record in the Sheep Database	29
SOP 11.7.2.1 Creating a new Survey Record in SQL Server Management Studio	29
SOP 11.7.2.2 Creating a new Survey Record: Insert Query Syntax.....	30
SOP 11.7.3 Field Data Deliverables to Database Table Mappings	30
SOP 11.7.3.1 Prerequisites.....	30

SOP 11.7.4 ARCN Sheep Data Management Tools ArcToolbox	30
SOP 11.7.5 Importing Field Data From NPS.gdb	31
SOP 11.7.5.1 Data Import Part 1: Generate SQL INSERT Query Scripts Using NPS.gdb to SQL Server ArcToolbox Tool	32
SOP 11.7.5.2 Data Import Part 2: Running SQL INSERT Query Scripts	35
SOP 11.7.6 Importing Buffers	36
SOP 11.7.7 Importing Pilot/Observer Waypoints and Tracklogs.....	37
SOP 11.7.7.1 Importing Waypoints Using the WaypointsToSQL.py ArcTool.....	37
SOP 11.7.7.2 Importing Waypoints Using a Spreadsheet	39
SOP 11.7.7.3 Importing Pilot Tracklogs.....	40
SOP 11.7.8 Animal Group Images	41
SOP 11.8 – SOP Executing Large SQL Files	42
SOP 11.9 – Creating Arcmap Query Layers.....	42
SOP 11.10 – Final Quality Assurance/Control Procedures	45
SOP 11.11 – Data Preparation for Analysis.....	46
SOP 11.11.1 Group Size and Distance for Program Distance.....	46
SOP 11.11.2 Group Size and Sex-Age Classes for Program R/WinBUGS/OpenBUGS.....	48
SOP 11.11.3 Group Distance for Program R/WinBUGS/OpenBUGS.....	49
SOP 11.11.4 Transect Data for Program R/WinBUGS/OpenBUGS	49
SOP 11.11.5 Transect Center Point Coordinates for Program R/WinBUGS	50
SOP 11.12 – Seasonal Close Out and Dataset Certification.....	51
SOP 11.13 – Archiving Field Deliverables in IRMA	52
SOP 11.13.1 Creating a Field Deliverables Dataset Reference in IRMA	52
SOP 11.13.2 Link the Field Deliverables Reference as Product of the Master Sheep Monitoring Project Reference	54

SOP 11.14 – Literature Cited.....	56
SOP 11.15 – Appendix A: Transect Editing Illustrations.....	57
SOP 11.16 – Appendix B: Alternate Method to Generate Buffers.....	61
SOP 11.17 – Appendix C: Dataset Certification Form.....	66

SOP 11 - Figures

	Page
Figure SOP11-1. Survey directory structure for the 2011 Itkillik survey showing the working and read-only raw data directories.....	8
Figure SOP11-2a. Screenshot examples of incorrect ridgeline buffers produced with the Buffer tool in ArcMap 10.3.....	16
Figure SOP11-2b. Screenshot examples of corrected ridgeline buffers produced after splitting DB Tracklog and re-running the Buffer tool in ArcMap 10.3.....	17
Figure SOP11-3. ARCN Sheep Data Management Tools ArcToolbox.	31
Figure SOP11-4. Example directory and file structure for the SQL scripts exported by the NPSdotGDBtoSQL.py tool.	31
Figure SOP11-5. Waypoints to SQL ArcTool.....	38
Figure SOP11-6. WaypointsToSQL ArcTool output SQL scripts.	39
Figure SOP11-7. WaypointsToSQL ArcTool example in batch mode.	39
Figure SOP11-8. Pilot tracklog to SQL tool.....	41
Figure SOP11-9. Example of a sheep survey dataset IRMA reference.....	54
Figure SOP11-10a. Copy for Appendix B – Screenshot examples of incorrect ridgeline buffers produced with the Buffer tool in ArcMap 10.3.	61
Figure SOP11-10b. Copy for Appendix B – Screenshot examples of corrected ridgeline buffers produced after splitting DB Tracklog and re-running the Buffer tool in ArcMap 10.3.	61
Figure SOP11-11. Screenshot examples of incorrect ridgeline buffers and location where tracklogs should be split to produce proper buffers.	63
Figure SOP11-12. Screenshots showing where additional buffer editing may be necessary.	64

Figure SOP11-13. Final buffers after checking left/right designation and overlapping transects; splitting tracklogs at ridgelines; running the Buffer, Integrate, Merge and Dissolve tools; Erasing the overlap areas; and Clipping the Buffers to the Survey Boundary..... 65

SOP 11 – Tables

	Page
Table 1. Example directory structure under the main sheep monitoring network directory for raw field data files (read-only) and the same files available for processing.	8
Table 2. Mappings of field data components to database tables.....	30
Table 3. Analytical data formats, tools and database views.....	46

SOP 11.1 – Introduction

SOP 11 covers data processing, quality assurance and quality control (QA/QC), preparation of the data for analysis, and data archiving in the Master Sheep Monitoring Database and IRMA. For analysis, see SOP 12, Schmidt et al. (2012) and Schmidt and Rattenbury (2013).

A general overview of data processing steps appears below with specific procedures described in more detail in the sections that follow.

1. Compile and organize all field data following the surveys (SOP 11.2).
2. Initiate the quality assurance/control (QA/QC) process by verifying the geodatabase(s) with hardcopy data (datasheets and GPS-Photo Linked photographs) and the pilots' tracklogs and waypoints (SOP 11.3)
3. Edit the geodatabase as needed in ArcMap (SOP 11.4).
4. Recalculate the perpendicular distance of each sheep group to the transect lines (SOP 11.5).
5. Create transect buffers (SOP 11.6).
6. Import the data layers from NPS.gdb into the sheep monitoring database (SOP 11.7, SOP 11.8, SOP 11.9).
7. Perform quality control and assurance procedures (SOP 11.0).
8. Prepare data for analysis in program R and WinBUGS (SOP 11.11). If anomalies show up during analysis, another round of data verification, validation and analysis will be initiated.
9. Perform seasonal close out and dataset certification (SOP 11.12, SOP 11.13).

SOP 11.2 – Post Field Campaign Data Handling

Sheep survey data should come from the field well organized according to SOP 9. The first office task is to archive the files on the Network shared drive.

1. Navigate to the sheep monitoring root directory on the Network shared drive (contact the Network data manager for access information).
2. Field survey data is organized in the sheep monitoring network folder first by year, then by survey. If a directory for the current year does not exist, create it under the \data directory. Example: \data\2016.
3. Create a directory in the survey year directory for the current survey if it does not exist already. Give it a logical name using the PI's naming scheme (Park or survey area and year) or following existing directories. Example: \data\2016\WRST2016.
4. In the new survey directory create two subdirectories according to the following naming scheme, one for read-only raw field data to be archived and the other for working copies. The former preserves the data exactly as it came from the field and the other makes them available for processing. You should have two directories similar to the example below:
 - a. \data\2016\WRST2016\FieldData-ReadOnly
 - b. \data\2016\WRST2016\FieldData-Working
5. Compress the files in the \FieldData-ReadOnly directory and give it Read-only attributes.

Table 1 describes in more detail the structure and function of the paired field data directories:

Table 1. Example directory structure under the main sheep monitoring network directory for raw field data files (read-only) and the same files available for processing.

Directory	Purpose
\\Data\\2016\\WRST-2016\\FieldData-ReadOnly	Directory for archiving raw field data. Tagged read-only and never modified. Add all field data files to a compressed (zip) archive and tag it and the directory read-only. If there are multiple datasets from multiple flights create logical subdirectories in the zip archive.
\\Data\\2016\\WRST-2016\\FieldData-Working	Same files as above but these working files may be modified, and processed for importing into the main sheep database. Maintain a change log in a metadata file. If there are multiple datasets from multiple flights create logical subdirectories or a consistent file naming scheme.

Figure SOP11-1 shows an example of the Itkillik 2011 survey directories. The ‘data’ directory is organized by year with subdirectories for each survey. Each survey directory has a working and read-only directory. The read-only directory is a zip archive of survey files as they came from the field.

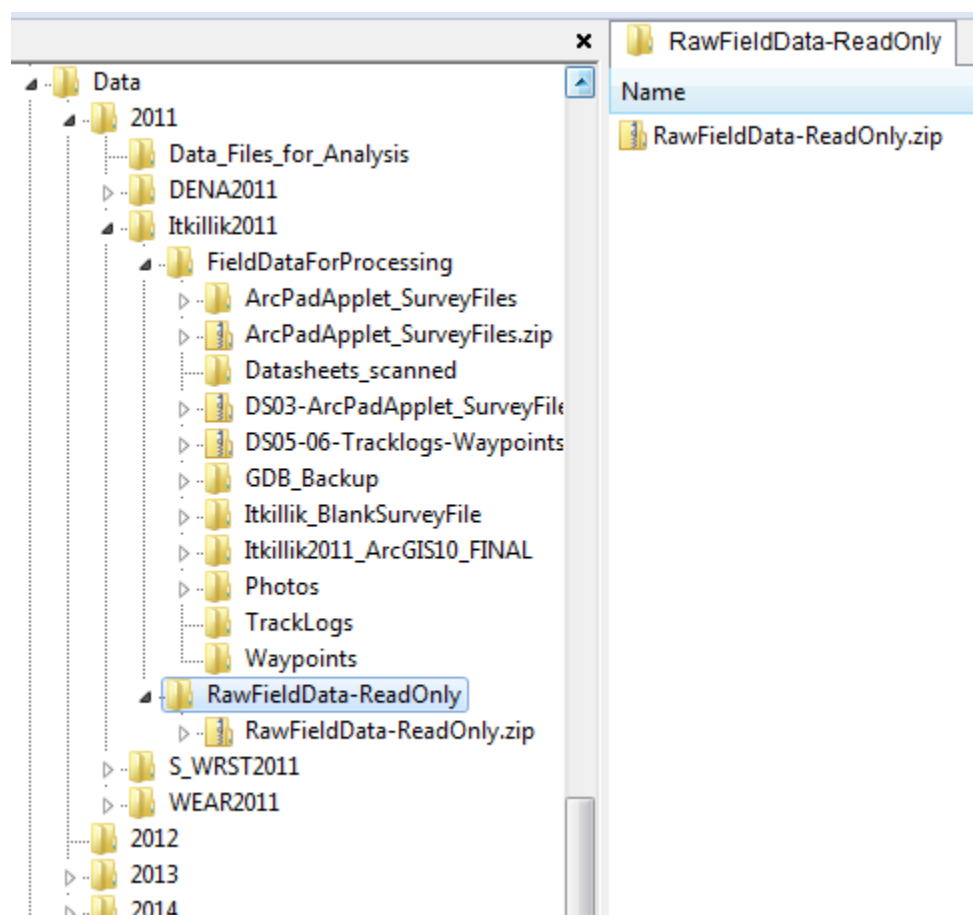


Figure SOP11-1. Survey directory structure for the 2011 Itkillik survey showing the working and read-only raw data directories. The read-only directory contains the survey files as they came from the field.

Fully processed deliverables and associated metadata from the directories above will be integrated into the main sheep monitoring database and then archived in the IRMA data store during seasonal close-out (SOP 11.12 and SOP 11.13).

SOP 11.3 – Initial Quality Assurance/Control Procedures

Data verification and data validation are two processes used to ensure data quality assurance and control (QA/QC). Data verification checks that data entered into a secondary (e.g., electronic) format match the source data, whereas data validation checks that data make sense. Although data entry and verification can be conducted by personnel with a general familiarity of the data (e.g., survey observers), the validation process requires in-depth knowledge about the project and data collected (i.e., Project Manager, AKA Principal Investigator). Data editing to resolve errors can occur simultaneously with the initial verification process. The final round(s) of data verification and validation should be performed by the project manager.

The project manager is responsible for overseeing the initial verification of sheep monitoring data (both physical and electronic sources). The purpose of verifying data is to reduce collection errors thereby enhancing accuracy of data and subsequent analysis and reporting. Physical data (e.g., paper field data forms and notes) are verified for accuracy, legibility and completeness immediately after collection and are then compared with the geodatabase, first by the observer while in the field (SOP 9), and then by the project manager once the survey is complete. The project manager is responsible for resolving any errors as soon as possible according to the procedures in this SOP.

The project manager(s) will assemble all data following the surveys. He/she will check the datasheets for legibility and completeness or will delegate this to a technician who was not involved with the data collection. If photographs have not yet been GPS-Photo Linked (i.e., observers were not able to complete this during the survey), all photographs will be processed according to SOP 10. A minimum of two people not involved with data collection, specifically project manager(s) and/or observers with multiple years of experience with aerial Dall's sheep surveys will verify and correct group size and sex and age classification data by comparing the geodatabase and datasheets with the photographs. Group size may be increased based on photograph verification, but will not be decreased if fewer sheep are found in a photograph than the observer-recorded data because it is often difficult to photograph large and/or scattered groups of sheep.

The project manager(s) will then open the ArcMap project associated with the survey and check the data transect by transect by comparing the geodatabase to the datasheets, pilot tracklogs and waypoints and the GPS-Photo Linked photographs. If there were multiple aircraft involved in the survey and each observer was downloading and checking his/her data separately during the survey, then there will be as many ArcMap projects and associated geodatabases as there are survey base locations (e.g., one for Bettles and one for Galbraith in a GAAR survey), or even one project and geodatabase per observer or aircraft. Each survey base location (or observer/aircraft) will have its own Transect_Generation_ArcMap_Extension folder housing an ArcMap project and geodatabase (see SOP 2.1.1). The project manager should check each of these in separate sessions with the respective sets of datasheets, tracklogs, waypoints, and photographs. Any discrepancies between the geodatabase, the datasheets, the pilots' tracklogs and waypoints and the photographs will be discussed with the observer (if needed), and corrected in an ArcMap editing session (SOP 11.4).

SOP 11.4 – Geodatabase Editing

All transect flight line and sheep location data within the geodatabase are processed and edited as needed before being imported into the main sheep monitoring database. Data will be analyzed using methods similar to those described by Walsh et al. (2010). Editing the geodatabase may be necessary (1) where there are discrepancies between the geodatabase and the pilots' tracklogs and waypoints or with the datasheet information or photographs, (2) where the external GPS lost satellite coverage or connection with the field laptop while flying and (3) where on- and off-transect segments (from the DB Tracklog shapefile) need to be edited because observers are not always able to switch segment type (on/off) in the heat of the moment when classifying a group of sheep.

Although the observer may have time to edit the geodatabase for these items while in the field, more often than not, edits are made during the initial QA/QC effort in the office by one of the project managers or a technician who was not involved with data collection. The observer should note on the datasheets where geodatabase edits are necessary if he/she does not have time to complete those edits in the field. The observer should be consulted if there are any questions about discrepancies between the geodatabase and other data forms; therefore it is important to consult with the observer in a timely manner following the survey when his/her memory is freshest.

If there were multiple aircraft involved in the survey and each observer was downloading and checking his/her data separately during the survey, then there may be as many ArcMap projects and associated geodatabases. Start with one ArcMap project, complete all edits to the associated geodatabase and repeat this with each of the ArcMap projects and associated geodatabases. Then combine the geodatabases in one ArcMap project (step 13), before Recalculating Group Distances (SOP 11.5) and Creating Buffer Files (SOP 11.6).

- 1) Open the ArcMap project (i.e., NPS_Transect_Tool_DOTNET.mxd within the Transect_Generation_ArcMap_Extension folder [see SOP 2.1.1]).
- 2) Add the pilot's tracklog and waypoints and the observer's tracklog to the ArcMap project if they are not already there. If the GPS lost signal during the survey flight(s) and data are missing, you can use the pilot's tracklog and waypoints to edit the NPS.gdb, and specifically the feature classes for DB Tracklog and DB Animals (see steps 6-8 below).
- 3) Add the GPS-Photo Linked shapefile to the ArcMap project if they are not already there (see SOP 10.4). You can use this shapefile to compare the locations where photos were taken (represented by points in the photo link shapefile) with the DB Animals feature class to identify which photos go with which sheep group. The photos can then be accessed and compared with the data for that sheep group to check count and composition accuracy (see step 9 below).
- 4) Start an Editing session in ArcMap, and make sure 'DB Animals' and 'DB Tracklog' are set as the targets to be edited.
- 5) It is recommended to check each transect one at a time in a systematic fashion for each ArcMap project/geodatabase.

- 6) Examine each transect to make sure that the on-transect segments of the DB Tracklog line up with the pilot's and the observer's tracklogs.
 - a. You can identify the start and end points of flown transects and the sheep locations using the pilot's waypoints and the observer's notes recorded on the datasheet. You can also label the pilot's (point shapefile) tracklog with the LTIME field and use those labels to note the flight line direction for DB Tracklog.
 - b. The pilot's and observer's Garmin 296 and 76csx tracklogs are generally always more accurate and consistent than the external GPS-connection with the field computer and therefore more accurate than the DB Tracklog.
 - c. Where there are big differences between the On-Transect DB Tracklog segments and the pilot's/observer's tracklogs, click the Edit Tool (black arrow) from the Editor toolbar and select the problematic segment on the map. The segment will be highlighted in light blue.



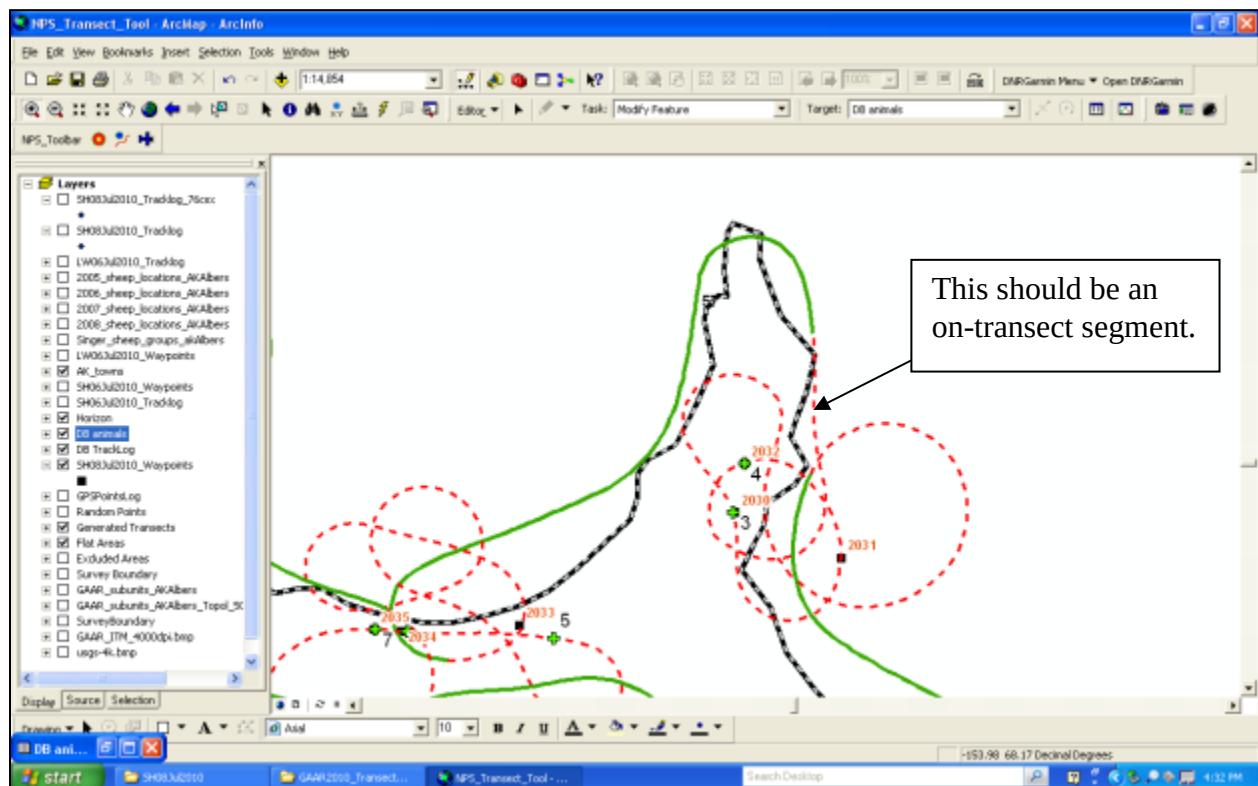
- d. Click the Edit Vertices tool from the Editor Toolbar. The segment will change from light blue to a series of green vertices that can be moved or deleted to adjust the segment to more accurately reflect the tracklog. To insert vertices, right click on a straight line portion of the DB Tracklog segment and select Insert Vertex from the drop-down menu. You only need to edit the on-transect segments.



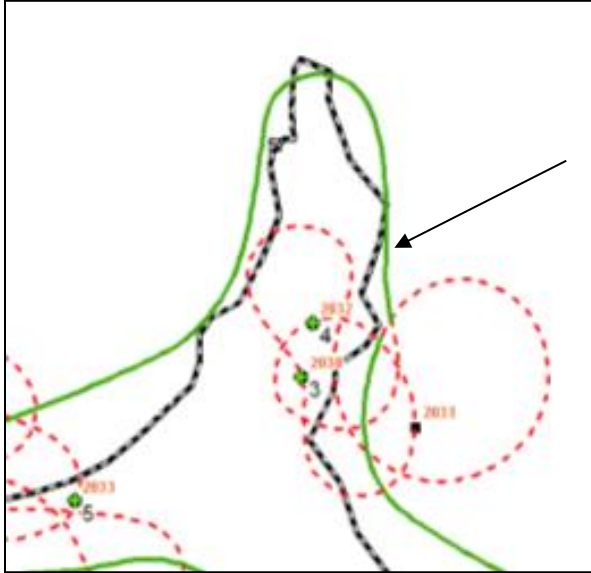
- e. If the total distance for all On-Transect segments of a particular transect exceeds the target transect length by more than 1 km, cut the tail so that it matches the target transect length. For example, if after editing you find that your total On-Transect line is 26 km when the target length is 20 km, then clip and delete the 6-km tail. You will have to eliminate any animals that correspond with the clipped tail from DB Animals.
- 7) If on-transect segments are missing (e.g., no GPS signal transmission to the applet) or if a portion of an on-transect segment needs to be re-digitized, you will use the Create Feature tool in ArcMap.
 - a. The Create Feature window may be open on the right side of your ArcMap project window.
 - b. Double-click on DB Tracklog within that window.
 - c. Use your cursor to trace the portion of the pilot's tracklog that corresponds to the missing portion of the on-transect. You only need to digitize the on-transect segments.

- d. Fill in the attributes table for those created segments or transects. Having to do this is not ideal, but it is a way to salvage the data if the GPS loses contact with your computer during an ArcPad session.
- 8) On-transect segments also must be edited so that the perpendicular distance of a sheep group to its transect and the transect buffer areas are calculated accurately in SOP 11.5 and SOP 11.6, respectively. Please see the example that follows and Appendix A, which shows additional before and after illustrations of edits that must be made to the on-transect lines.

BEFORE edits:



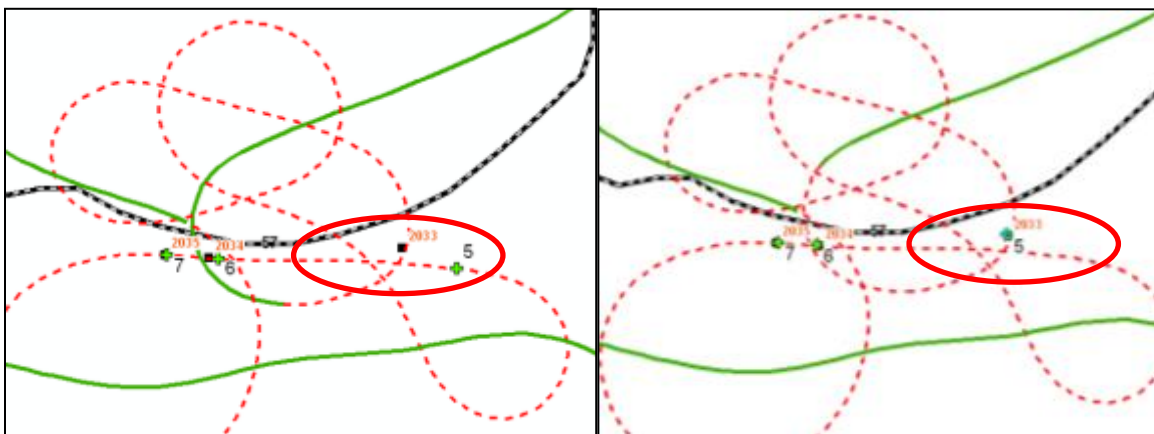
AFTER edits:



- 9) Examine each sheep group in the DB Animals feature class.
 - a. Compare the numbers for each sex-age class and the total count on the DB Animals Attribute Table with the datasheet numbers.
 - b. Compare the photographs (where available) for each group and, where appropriate, correct total count and/or sex-age classification after consulting with at least one other knowledgeable observer or project manager. A minimum of two people not involved with data collection, specifically project manager(s) and/or observers with multiple years of experience with aerial Dall's sheep surveys will verify and correct group size and sex and age classification data by comparing the geodatabase and datasheets with the photographs. Group size may be increased based on photograph verification, but will not be decreased if fewer sheep are found in a photograph than the observer-recorded data because it is often difficult to photograph large and/or scattered groups of sheep.
 - c. Check the sheep group locations with the corresponding pilot's waypoint. The pilot's waypoint number should be on the datasheet. Most pilots take better waypoints for sheep group locations than observers can on the field computers. If necessary, move the DB Animals group to the pilot's waypoint. Please see the example below and specifically sheep group 5.

BEFORE edits:

AFTER edits:



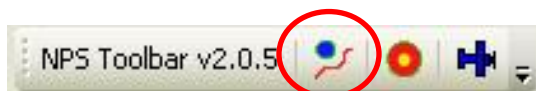
- d. You may have to edit or add sheep group information as well by moving or creating points in DB Animals and/or by making edits in the Attribute Table for DB Animals. Be sure that the datasheet corresponds with the Attribute Table. Mark a simple line through any incorrect data on the datasheet and note the correction beside it.
- 10) Open the Generated Transects Attributes Table. Delete all transects that were not flown by that pilot-observer team whose ArcMap project you are editing ('N' in the Flown field).
 - 11) Save all edits, close the Editing session and close the ArcMap project.
 - 12) Repeat Steps 1-11 for each ArcMap project/geodatabase associated with the survey.
 NOTE again: On the last ArcMap project, be sure to save those generated transects which were not flown by any pilot-observer teams on the survey.
 - 13) You will now combine the geodatabase information into one final geodatabase for the survey.
 - a. Copy one of the observer's Transect_Generation_ArcMap_Extension folder to a new location and rename it with the survey name and the word FINAL (e.g., DENA2012_Arc10_FINAL). Open the ArcMap project in this FINAL folder.
 - b. Use the Append (Data Management) tool in ArcMap to bring the DB Animals data from each observer's geodatabase into the FINAL DB Animals feature class. Repeat this for DB Tracklog and Generated Transects. Be sure that you have a full set of all transects (flown, not flown) in the Generated Transects feature class.
 - c. Save the ArcMap project.
 - d. Make a copy of the FINAL Transect_Generation_ArcMap_Extension folder for backup before you move on.

- 14) When possible, all changes to data will be documented in the edit log (within the comments field of the appropriate geodatabase component) and accompanied by an explanation that includes pre- and post-edit data descriptions. However, time constraints during the field survey period may preclude the Project Manager from entering edits in the electronic log. In these cases, edits should be made on hard copy data forms and later transferred to the electronic log. Allowable changes include those that improve or update the data formats while maintaining data integrity.

SOP 11.5 – Recalculating Group Distances

Distance sampling requires accurate measurements of the perpendicular distance of the object of interest (in this case groups of Dall's sheep) to the transect line (in this case, the corresponding on-transect segment from which the sheep group was first detected). The field DIST2TRANS in the DB Animals Attribute Table houses these data which are calculated when the observer selects Mark Animal while in flight, and the distance is adjusted if the observer updates the group position during the same off-transect session (see SOP 8.12). However, if the sheep group is moved during an ArcMap session (SOP 9.3.3, SOP 11.4), the distance data is not updated. Therefore, you should recalculate and check the group distances for all groups after the initial QA/QC and data editing sessions in the office.

- 1) Open the FINAL ArcMap project where you combined all geodatabase information from all observers and aircraft for the survey.
- 2) Run the Straight Line Tool by clicking that button on the NPS Toolbar v2.0.5. If the tool does not run to completion, you will check and adjust the values in the DIST2TRANS field anyway in the following steps.



- 3) Start an Editing session in ArcMap, and make sure 'DB Animals' is set as the target to be edited.
- 4) Open the Attributes Table for DB Animals.
- 5) Go group by group and, using the ArcMap measuring tool set to sum distances in meters, measure the perpendicular distance from the group to the respective DB Tracklog On-transect segment. Check those measurements with the values in the DIST2TRANS field in the table. Edit as necessary. If there is a question as far as what on-transect segment to use for the distance measurement (e.g., the flight line and on-transect segment wraps around a narrow ridge and the sheep group would have been seen as easily and despite disturbance had the line been flown in the opposite direction), discuss this with one of the project managers.
- 6) Save your edits and close the editing session.
- 7) Open the DB Animals.csv or text file in Excel. Remove any sheep groups that are > 685 m, or your specified effective strip width outer limit taking into account the 22 m blind area next to the on-transect line. If there are any questions about sheep groups just on the

line or that ended up below the line – the project managers should confer to discuss eliminating or retaining these for analysis. Groups within the 0-22 m blind area can be set to 0 m for distance to transect, however the project managers should determine what portion of those if any should be eliminated. Save this as an Excel file.

- 8) The Animal groups are now ready to be imported into the main sheep monitoring database.

SOP 11.6 – Creating Buffer Files

The effective strip width, which is determined using the distance data (Schmidt et al. 2012), is applied as a buffer to the uphill (search) side of the on-transect lines to represent the area sampled per transect.

NOTE - In 2015 there was a new problem with buffering that affected both the custom NPS Transect Tool and the generic ArcMap Buffer tool requiring additional edits to DB Tracklog segments as part of the buffer generation process. This may be a change in the upgrade from ArcMap 10.1/10.2 to 10.2.2/10.3. Where transects run around a ridgeline, the buffer is only produced for the area between the outgoing and return transects, and is not produced for the area outside of those lines which would be the full buffer for the opposite side of the ridgeline (see Figures SOP 11-2a and 2b below).

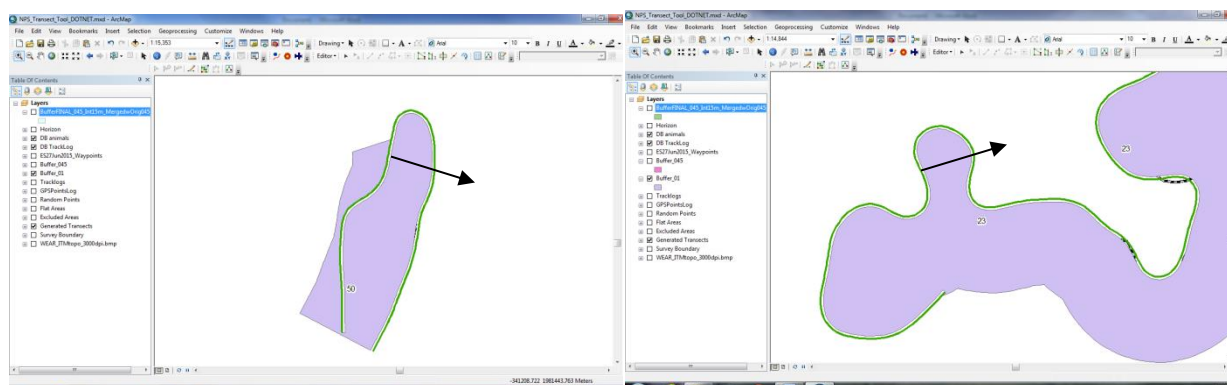


Figure SOP11-2a. Screenshot examples of incorrect ridgeline buffers produced with the Buffer tool in ArcMap 10.3. The buffer from the left side of these ridgelines should overlap the right-side flight line (green DB Tracklog line) to cover the full effective strip width, indicated by the black arrows. The buffer tool in later versions of ArcMap erases this overlap during processing, resulting in underestimated buffer areas. Compare with Figure SOP11-2b. See also Appendix B of SOP 11 for a correction.

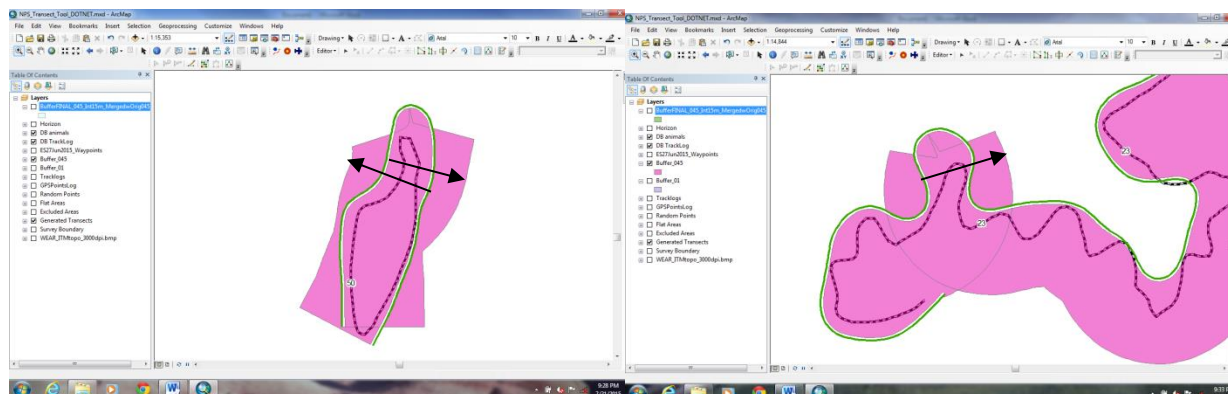
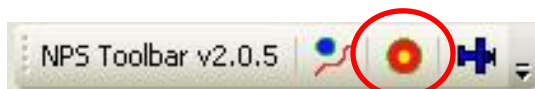


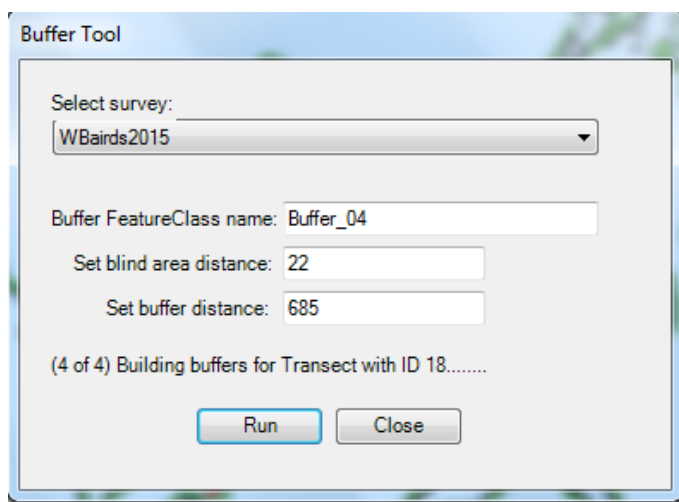
Figure SOP11-2b. Screenshot examples of corrected ridgeline buffers produced after splitting DB Tracklog and re-running the Buffer tool in ArcMap 10.3. The black arrows indicate where the corrected buffers cover the full effective strip width compared with Figure SOP11-2a.

If your first buffer tool attempt (steps 1 and 2 below) produces buffers like Figure SOP11-2a (incorrect, underestimated buffer areas), proceed to Appendix B and make buffers according to those directions. Then skip to SOP 11.7. If your first buffer tool attempt looks like Figure SOP11-2b, proceed with the directions below.

- 1) Open the FINAL ArcMap project where you combined all geodatabase information from all observers and aircraft for the survey.
- 2) Run the Buffer tool for the first time so that you can check that all buffers are on the uphill side of each transect. Click the Buffer (target-shaped) button on the NPS Toolbar v2.0.5.

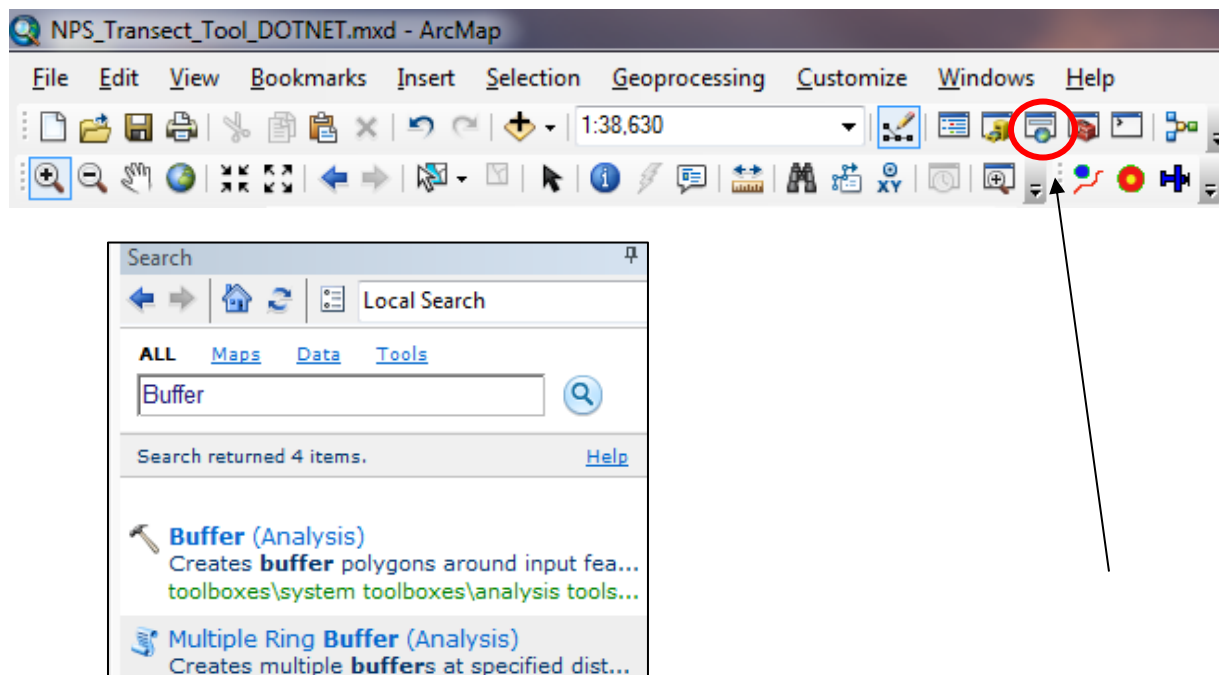


In the Buffer Tool window that opens, pick your Survey ID from the drop-down list, name your Buffer feature class, set the distances at 22 and 685 and hit Run.



You may get an error message that the Buffer table doesn't exist. Close the Buffer Tool window, Save and Close your ArcMap project. Reopen your ArcMap project and select the Buffer tool button again. This time your Buffer feature class should be there for you to select. Once it's running you'll see it processing and making buffers in the Buffer Tool window. This may take 5-30 minutes to run. If the Buffer tool doesn't work, see steps 5-6 and generate one buffer set each for the left and right on-transect segments from DB Tracklog.

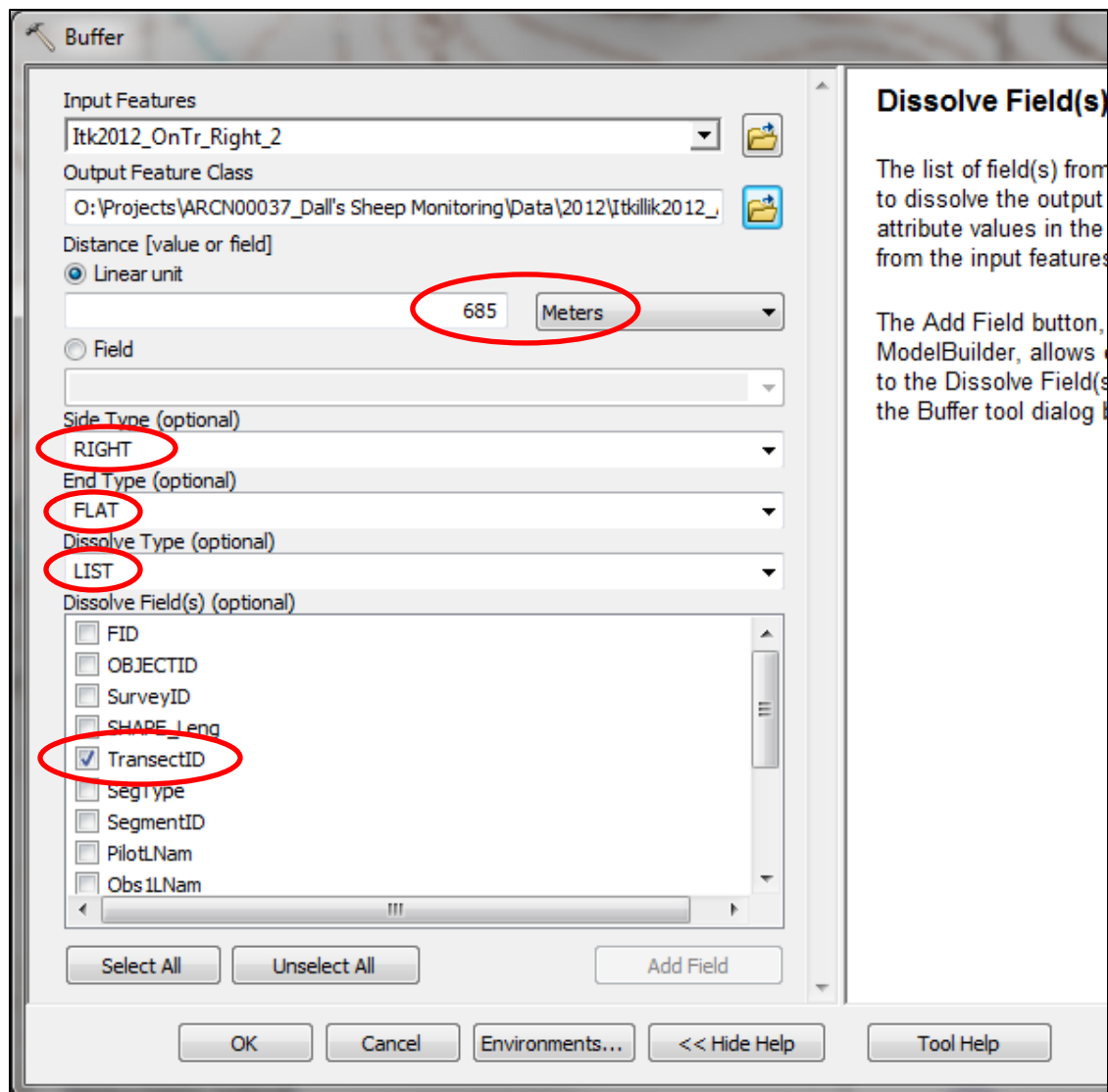
- 3) Review the first buffer set(s) with a topo layer background to make sure all buffers are on the uphill side of their respective transects. For those that are on the downhill side, you will need to change the left or right designation of the associated on-transect segments in the DB Tracklog feature class. Start an Editing session and make sure DB Tracklog is set as the target dataset for editing. Open the Attributes table for DB Tracklog and make the appropriate changes Left or Right in both the PilotDir and Obs1Dir fields. The designation should always indicate the uphill side for the direction that the on-transect segment was flown. Save your edits and close the Editing session.
- 4) Rerun the Buffer Tool per step 2. If the Buffer Tool works skip to step 10 for final verification of the buffers. If the Buffer Tool does NOT work, see steps 5-9.
- 5) If the Buffer Tool isn't working you can "manually" generate buffers. Select the left, on-transect segments in DB Tracklog and save them to a new line shapefile (e.g., Itkillk2012_Left_On-transect.shp). Repeat this with the right, on-transect segments.
- 6) If you are using ArcMap 10, access the Buffer (Analysis) tool via the Search function – which is a button at the top of your menu bars and opens the Search window.



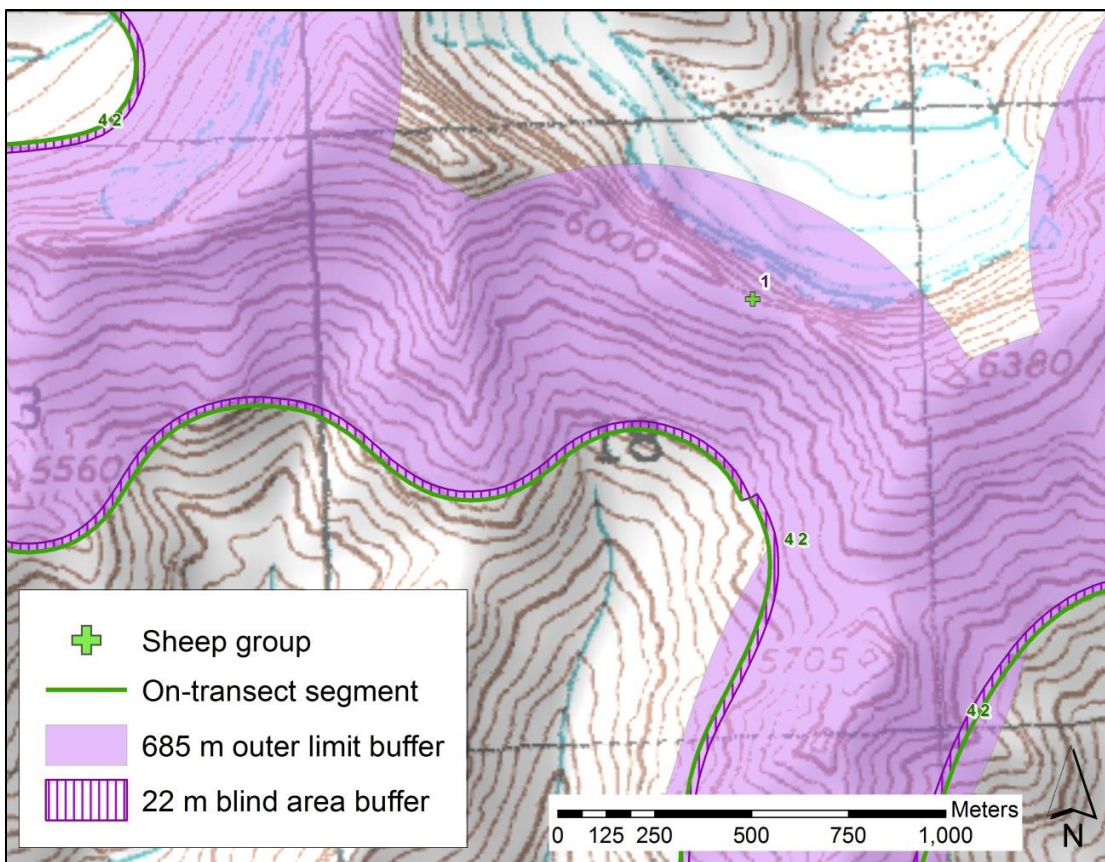
Run the Buffer (Analysis) tool to produce four buffer shapefiles from the left and right shapefiles you created in step 5:

- Left outer limit (maximum width) buffer from the left on-transect lines
(e.g., Itkillik2012_Buffer_685m_Left.shp)
- Left blind area buffer from the left on-transect lines
(e.g., Itkillik2012_Buffer_22m_Left.shp)
- Right outer limit (maximum width) buffer from the right on-transect lines
(e.g., Itkillik2012_Buffer_685m_Right.shp)
- Right blind area buffer from the right on-transect lines
(e.g., Itkillik2012_Buffer_22m_Right.shp)

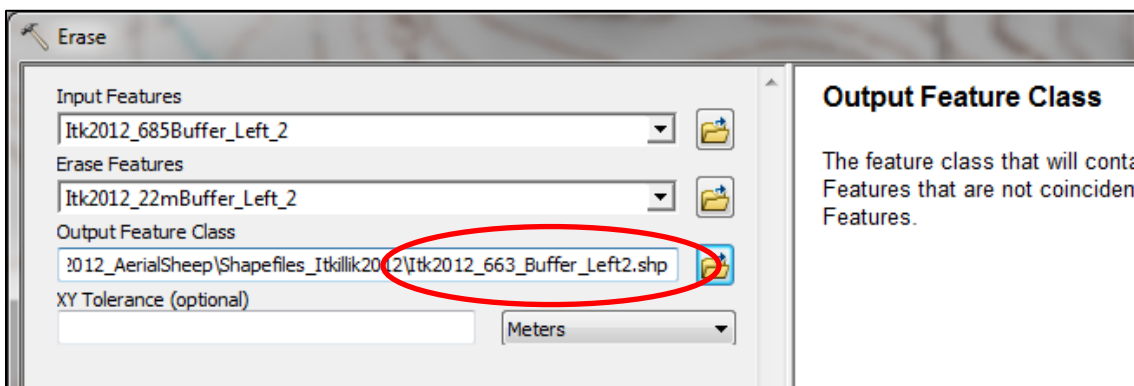
The buffer widths are determined by the effective strip width, which is based on the detection function or distribution of the distances from the line of the observed sheep groups, removing the farthest 5% of groups (outer limit) as well as the 'blind area' closest to the line (i.e., 22 m, Walsh et al. 2010). For Schmidt et al. (2012), the effective strip width was 663 m between an outer limit of 685 m from the on-transect line and a 22 m blind area buffer adjacent to the on-transect line. If you are using the same effective strip width, you will produce 685 m and 22 m buffer shapefiles for both the left and the right on-transect shapefiles (total of four). Please see the example below for producing a right-side, 685 m, outer limit buffer shapefile, and note the settings circled in red.

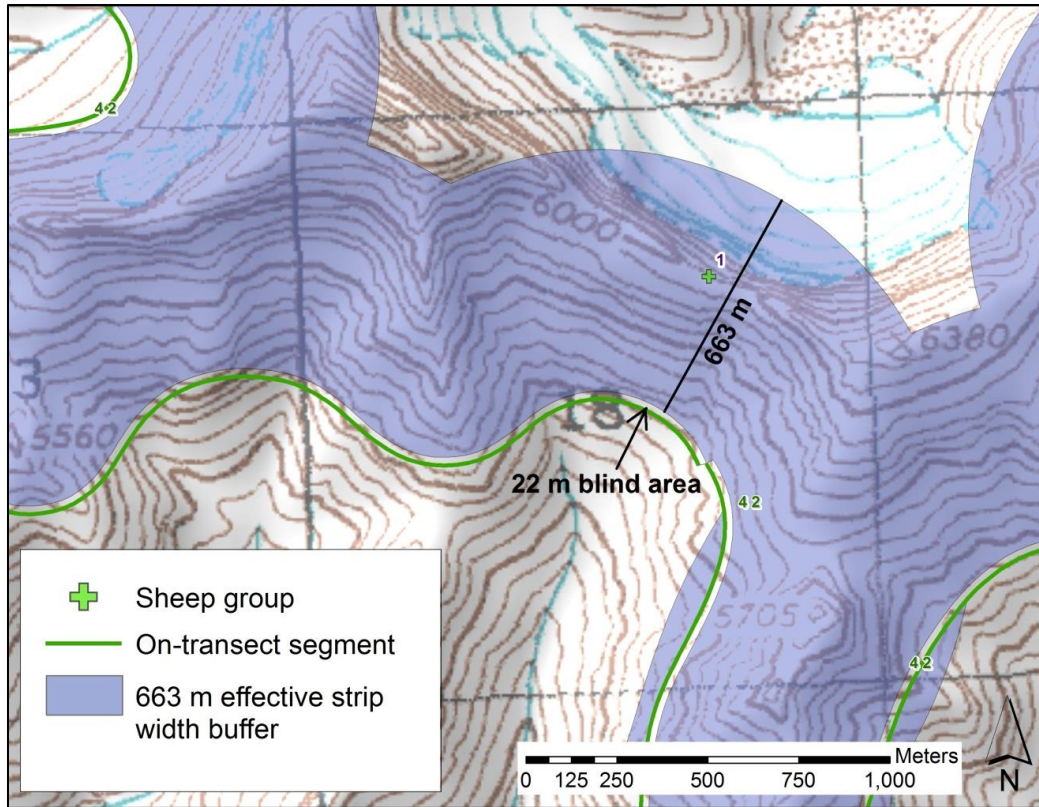


The resulting shapefiles should look like this:

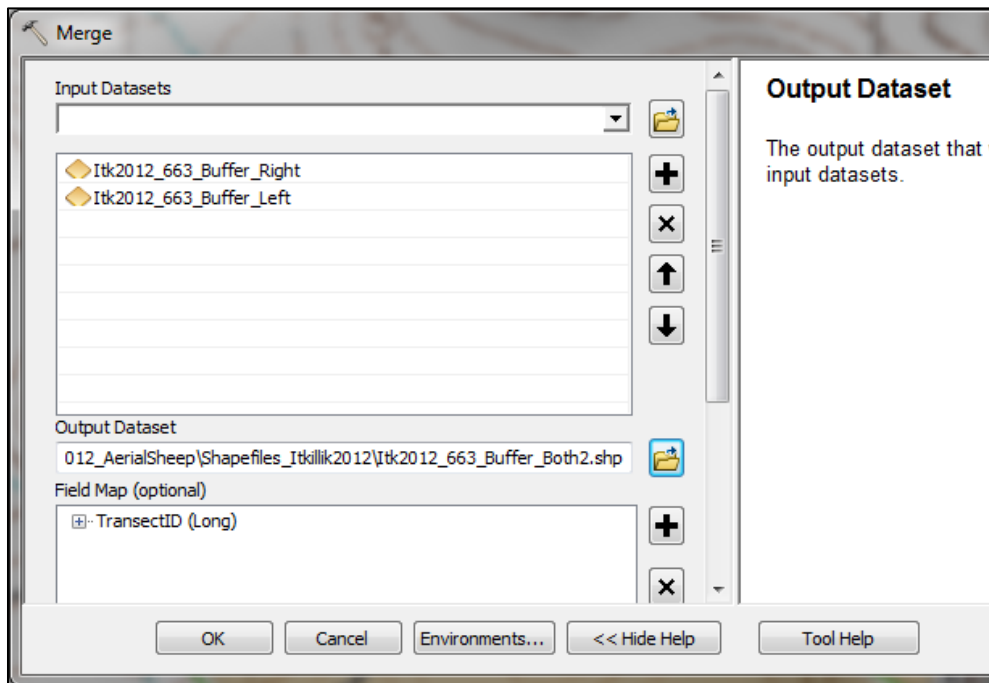


- 7) Use the Erase (Analysis) tool (ArcMap 10) to remove the blind area from the left-side, outer limit buffer shapefile. Repeat this with the right-side, outer limit buffer shapefile to produce two effective strip width buffer shapefiles (one left and one right). The example below shows this being done with the left side buffers. You will create two new shapefiles in this step and this buffer covers the effective strip width (e.g., Itkillik2012_Buffer_663m_Left.shp and Itkillik2012_Buffer_663m_Right.shp).

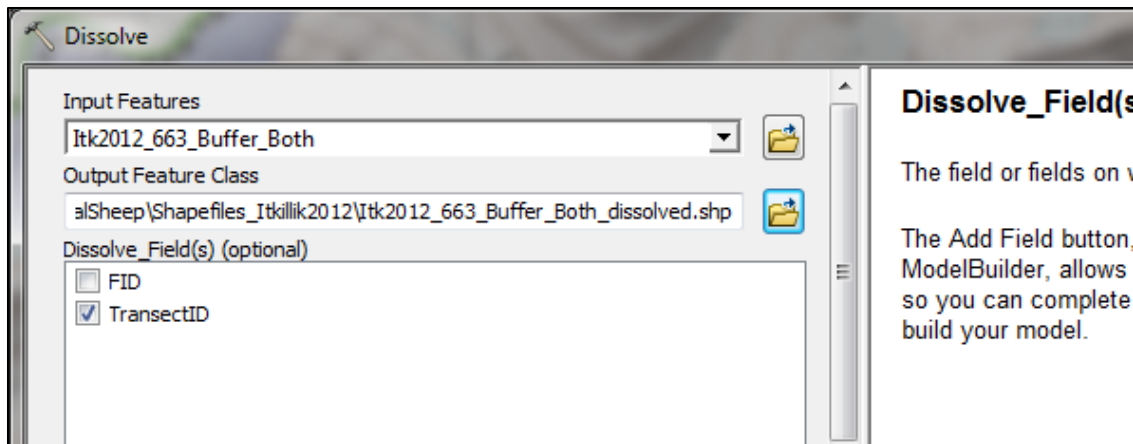




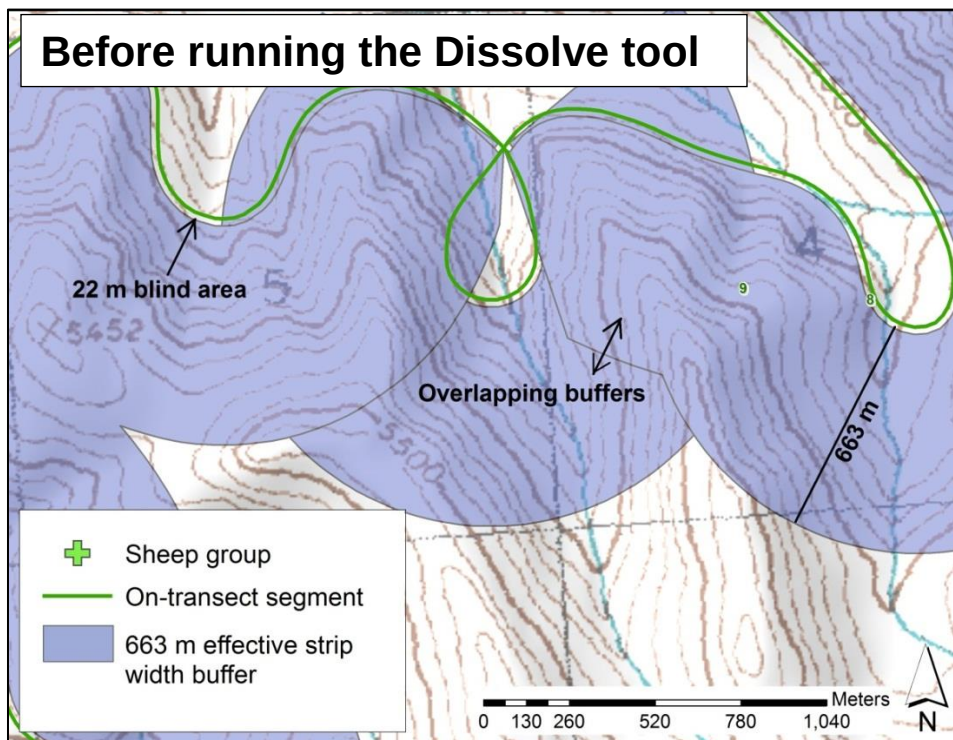
- 8) Use the Merge (Data Management) tool (ArcMap 10) to merge the left and right effective strip width buffer shapefiles to produce one effective strip width buffer shapefile (e.g., Itkillik2012_Buffer_663m_Both.shp).

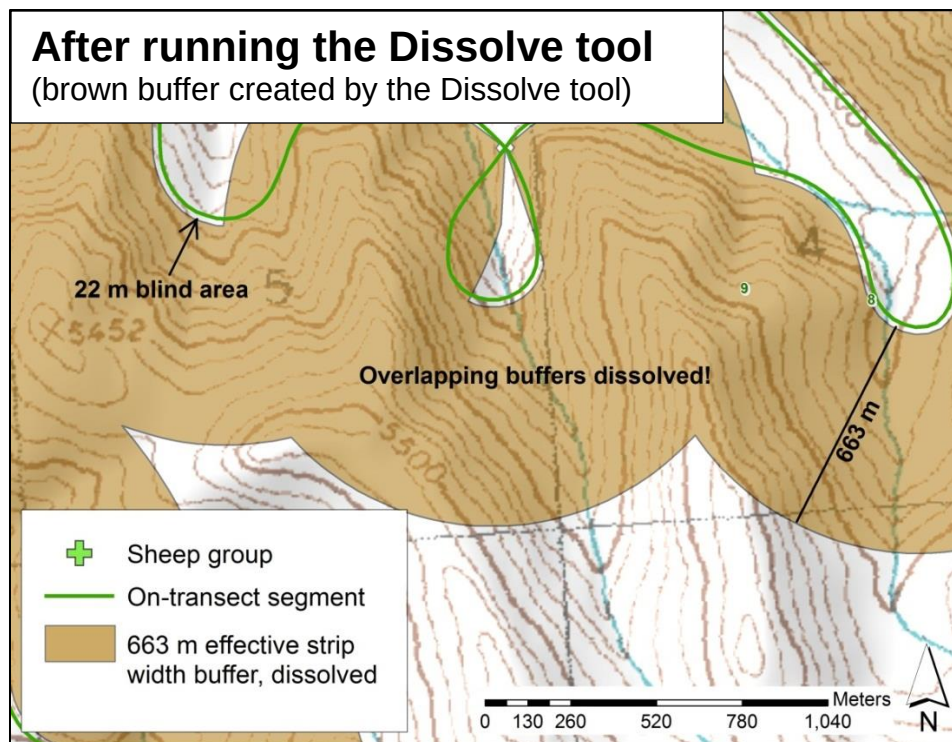


- 9) Run the Dissolve (Data Management) tool (ArcMap 10) on the effective strip width buffer shapefile to dissolve lines of adjacent or overlapping buffer segments. This will produce yet another shapefile (e.g., Itkillik2012_Buffer_663m_Both_Dissolved.shp). If you don't run the Dissolve tool, the overlapping buffer segments will inflate the effective strip width and search area and throw off the sheep density and abundance estimates.

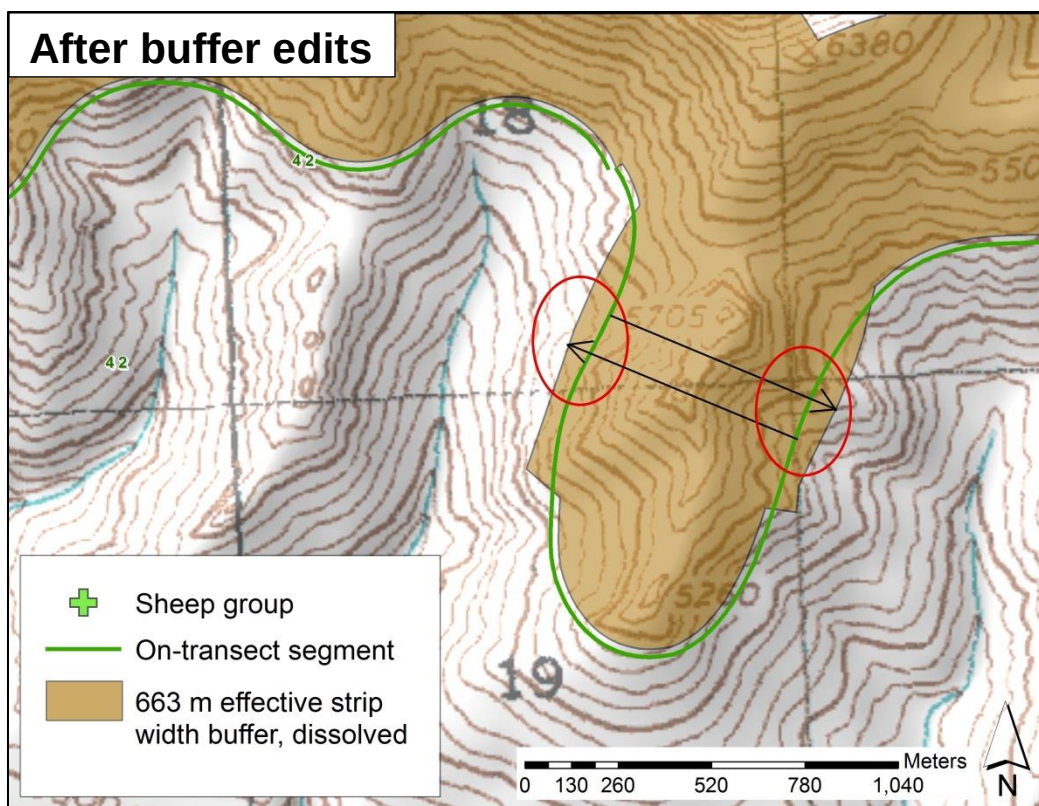
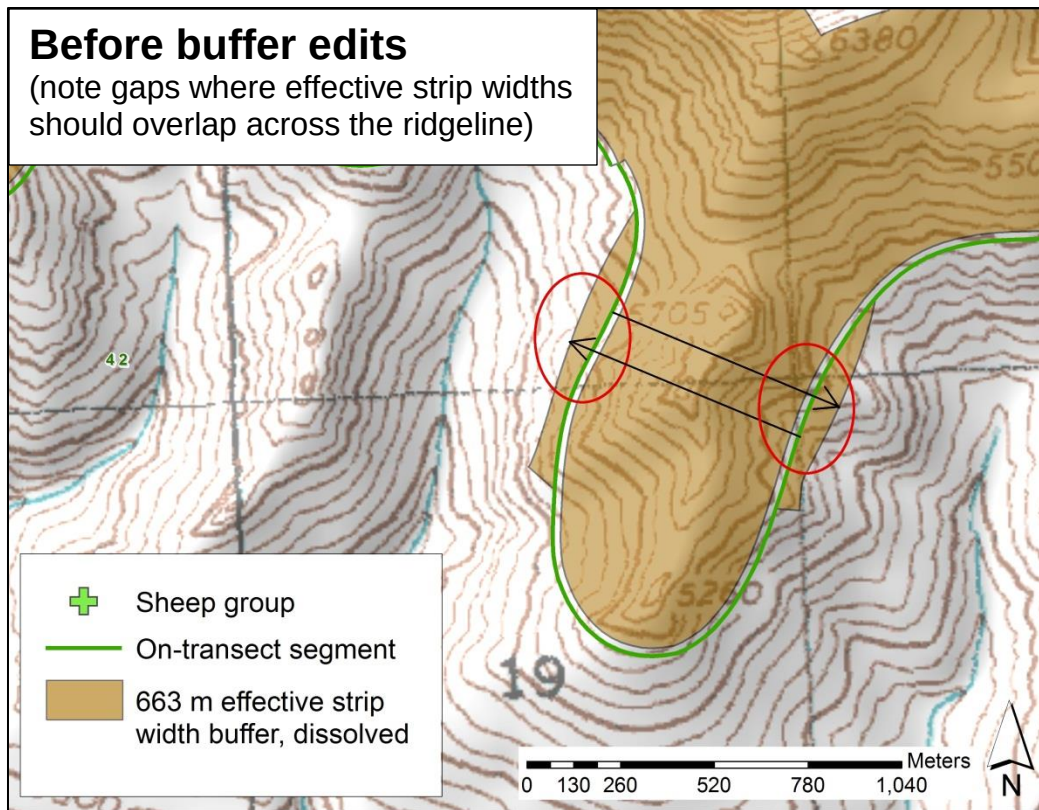


The following figures show the buffer shapefiles before and after running the Dissolve tool.

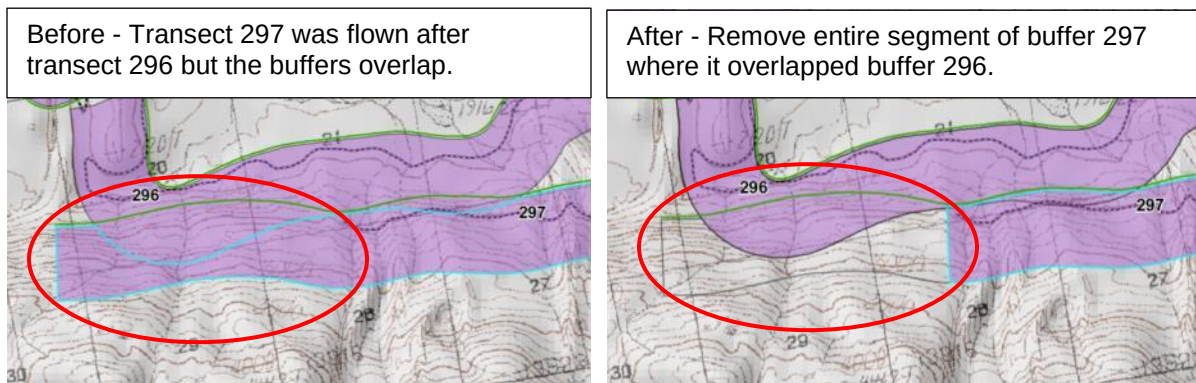




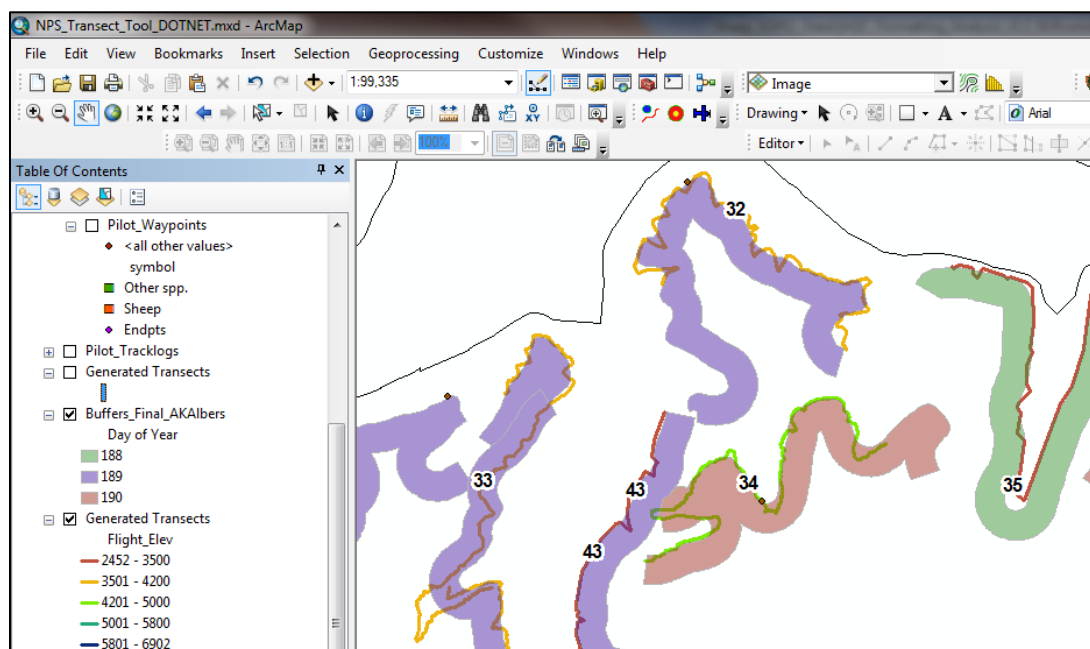
- 10) Check every buffer in the shapefile you just produced in step 9, and edit any overlaps or gaps. Use the ArcMap Editor tools to make these changes. In the following screenshot example, the Erase tool used earlier had cut some portions of the blind area that could be seen from another portion of the flown transect on the other side of the ridgeline. Note the red circles, black arrows and gaps where the effective strip widths on both sides of the ridgeline overlap. Go into Editing mode in ArcMap, select the buffer shapefile for editing and delete, move, etc. vertices for each buffer to eliminate any overlaps or gaps. Close the Editing session.



- 11) Look for overlapping buffers from different transects. If buffers were flown the same day, are < 900' different in elevation, that segment of the buffer flown second needs to be removed completely where the overlap exists. All remaining buffer segments must retain the effective strip width (e.g., 663 m).

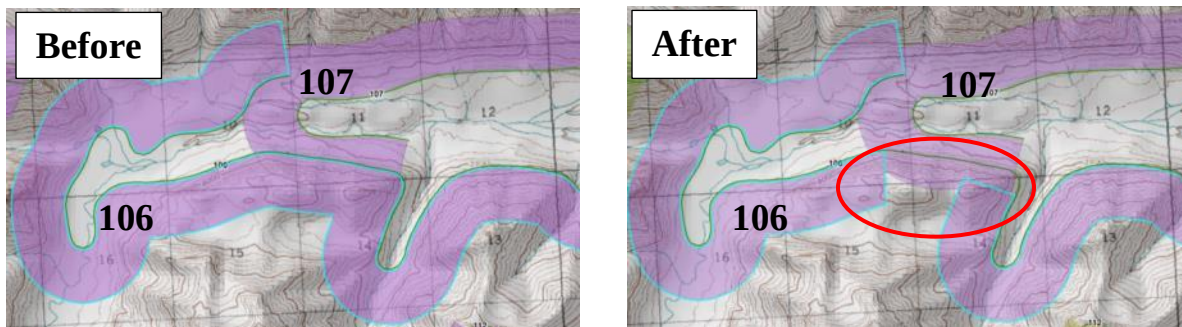


- a. You will need to first determine which buffers have an overlap problem. We recommend the following to make this process easier. Join your Buffer file to the Generated Transects feature class and change symbology to category so buffers are color coded by date (one of the Generated Transects fields). Make these semi-transparent. Then change symbology for Generated Transect to Quantities and color code by elevation (e.g., 5 categories set at each 1000' break). See screenshot below.

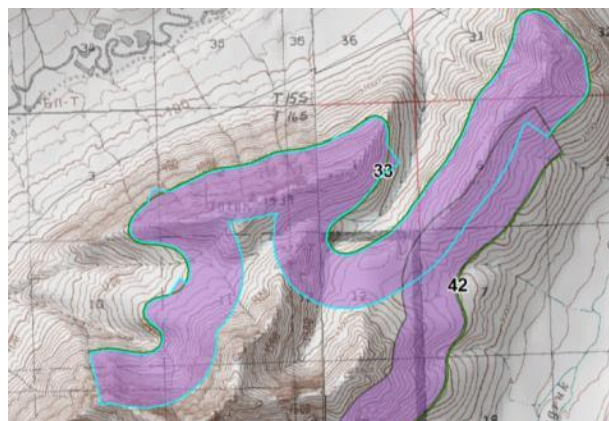


- b. Start an editing session to edit your buffers.
- c. If buffers are on the same side of the hill, overlap by > 30% visually, were flown the same day and are < 900' different in elevation, then remove the overlapping

portion of the buffer that was flown second. In the example below, Transect 106 (outlined in light blue) was flown after 107 so the overlap was removed completely (circled in red).

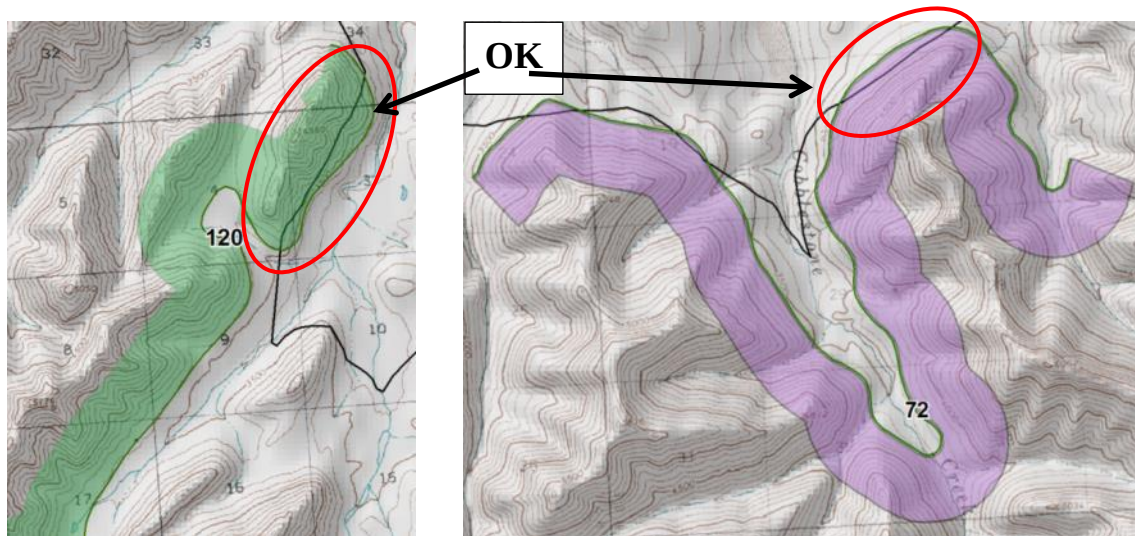


- d. If buffers are on opposite sides of a ridgeline and either transect is < 500 ft from the top of ridgeline, then remove the overlapping portion of the buffer that was flown second. In the example below, both transects are > 500 ft below the ridgeline so we keep both buffers. If one was closer to the ridgeline, the overlapping portion of the buffer flown second would be removed.

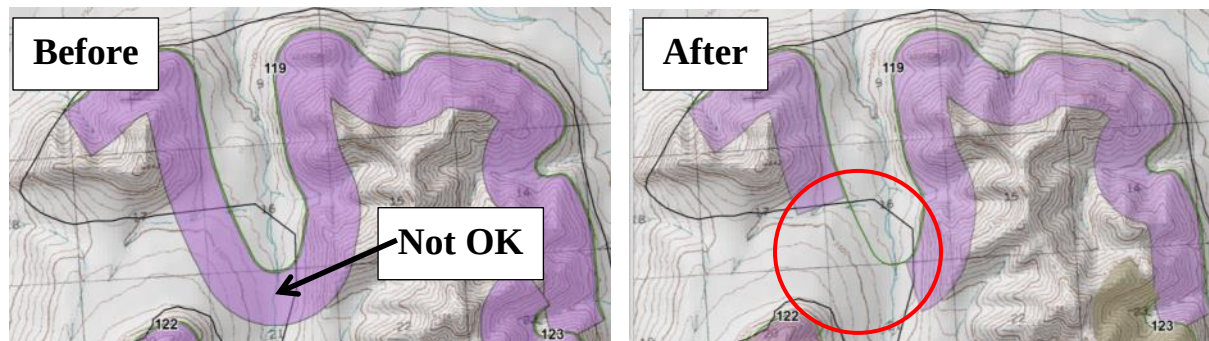


12) If the buffer extends outside of the survey boundary, there are two scenarios to consider.

- a. For the most part you won't have to remove buffers where the flight line just strayed outside for a bit and was more or less parallel to the survey boundary (first pair of screenshots below).



- b. However if the pilot flew outside the boundary, essentially perpendicular to the boundary line, remove any buffer outside of the boundary (see screenshot pair below showing before and after editing screenshots).



- 13) Calculate buffer area (m^2) for the clipped buffer shapefile. In ArcMap 10, you can use the Calculate Areas (Spatial Statistics) tool, which will produce another shapefile. In ArcMap 10.3, just add a new field to the attributes table called Area_m2, right click on the field and select Calculate Geography to calculate and input area for each buffer.
- 14) Export the final buffer shapefile as a Feature Class to your final survey geodatabase (where the final DB Tracklog and DB Animals are housed) and call it Buffers_Final.

SOP 11.7 – Importing Field Data to the Master Sheep Monitoring Database

SOP 11.7.1 Overview

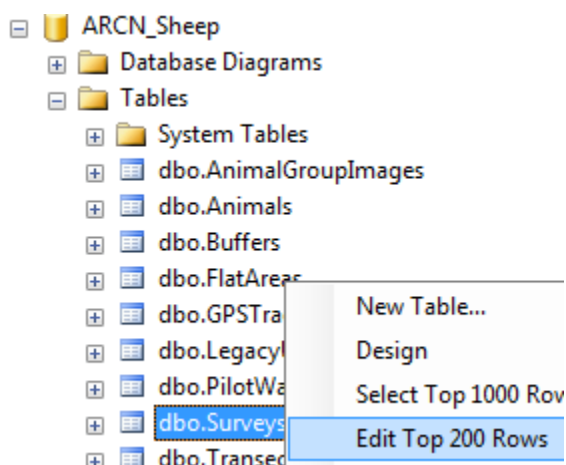
Once field data collection operations cease and initial quality assurance procedures have been done the information is transferred to the ARCN_Sheep SQL Server database. The major task in field data import is converting the field data from various spatial data formats into SQL insert query scripts that can then be executed against the sheep monitoring SQL Server database.

SOP 11.7.2 Creating a new Survey Record in the Sheep Database

The first step in importing any data from a field data collection campaign is to create a new Survey record. All the tables and related records in the sheep monitoring database are subordinate to the Surveys table. A Survey record must exist before any data can be related to it. A survey can be considered to be a container into which all the field data collected during a suite of survey flights can be organized. The principal investigator or Network data manager can create a Survey record.

SOP 11.7.2.1 Creating a new Survey Record in SQL Server Management Studio

1. Open SQL Server Management Studio
2. Log in to the ARCN SQL Server. Contact the Network data manager for login information.
3. Navigate to the ARCN_Sheep database.
4. Right click the Surveys table.



5. Select 'Edit top 200 rows'. A data grid showing the Surveys data appears.
6. Enter a name for the survey, the version of the monitoring protocol followed during the survey and a start date. More information may be filled in as needed.

	Survey	ProtocolVersion	StartDate
	2011 Denali Sheep Surveys	0.00	2011-07-20
	2011 Itkillik Sheep Surveys	0.00	2011-07-03
	2011 Wrangell St. Elias Sheep Surveys	0.00	2011-07-07
	2011 Western Arctic Parklands Sheep Survey	0.00	2011-07-07
	2014 Gates of the Arctic Sheep Survey	0.00	2014-07-10
	2013 Itkillik Sheep Surveys	0.00	2013-07-12
	2012 Itkillik Sheep Survey	0.00	2012-07-06
►*	NULL	NULL	NULL

7. Execute the underlying query and refresh to verify.

SOP 11.7.2.2 Creating a new Survey Record: Insert Query Syntax

Users familiar with relational databases may create a new survey using the tool of their choice by modifying and executing the following example query:

```
INSERT INTO Surveys (Survey,ProtocolVersion,StartDate) VALUES ('Survey name',1.0,'7/10/2016');
```

SOP 11.7.3 Field Data Deliverables to Database Table Mappings

Once the new Survey record is created you may begin to import field data into the master database. The database will ingest information from multiple sources as described in Table 2.

Table 2 shows the raw data sources and the database tables into which they should be entered (Table 2):

Table 2. Mappings of field data components to database tables

Database Table	Field Data Source
Animals	NPS.gdb/Animals feature class (Also Known As [AKA] DB Animals)
Buffers	NPS.gdb/Buffers feature class or shapefile created in SOP 11.6.
GPSTracks	NPS.gdb/GPSPointsLog feature class (not required)
PilotWaypoints	Pilot's GPS waypoints.
PilotTracklogs	Pilot's GPS tracklogs.
Surveys	A new survey record is generated by the database administrator or PI for each survey campaign.
Transect_or_Unit_Information	NPS.gdb/TrnOrig feature class (AKA Generated Transects)
TransectPoints	NPS.gdb/TrnPoints feature class (AKA Random Points)
TransectTracklog	NPS.gdb/Tracklog feature class (AKA DB Tracklog)

SOP 11.7.3.1 Prerequisites

There are a number of prerequisites for importing field data into the sheep monitoring database:

- A Survey record in the sheep monitoring database to which the data being imported will be related (see above).
- NPS Active Directory account
- SQL Server Management Studio or other software for database interactions.
- ArcGIS
- ARCN_Sheep SQL Server database connection information
- Datawriter privileges on the ARCN_Sheep database.
- Knowledge and experience using, troubleshooting and debugging ArcGIS, Python, SQL and Transact-SQL.
- Access to the ARCN Sheep GitHub code repository (<https://github.com/NPS-ARCN-CAKN/Sheep>)

Contact your Network data manager for help with any of the above.

SOP 11.7.4 ARCN Sheep Data Management Tools ArcToolbox

Most field data are imported to the sheep monitoring database using ArcGIS Python scripts. These scripts have been organized in an ArcToolbox toolbox called *ARCN Sheep Data Management Tools* (Figure SOP11-3, below). Each of the three tools converts field data into SQL insert query scripts that can then be executed against the sheep monitoring master database. NPS.gdb to SQL converts the field data collected by the ArcPad field data collection application

while Pilot tracklog to SQL and Pilot waypoints to SQL are the two tools used to convert pilot tracklogs and waypoints into SQL insert query scripts.

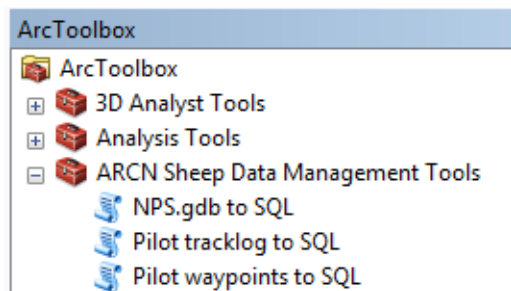


Figure SOP11-3. ARCEN Sheep Data Management Tools ArcToolbox.

The entire toolset is available from the sheep monitoring GitHub repository at <https://github.com/NPS-ARCEN-CAKN/Sheep>.

SOP 11.7.5 Importing Field Data From NPS.gdb

The Python script used to transfer data from NPS.gdb to the SQL Server database is called *NPSdotGDBtoSqlServer.py*. The script has a GUI tool in the ARCEN Sheep Monitoring Data Management Tools ArcToolbox called 'NPS.gdb to SQL' (Figure SOP11-4, below). The script and tool converts the field data records in each layer in NPS.gdb into SQL script (text) files (*.sql) containing insert queries. These scripts are then executed against the master sheep monitoring database to insert the records.

NPSdotGDBtoSqlServer.py Output File Naming Scheme

The SQL scripts are deposited in the same directory as the geodatabase and are named the same as each data layer with '.sql' appended to the layer name. As an example, the SQL script for the *Animals* layer would be generated as '*Animals.sql*'.

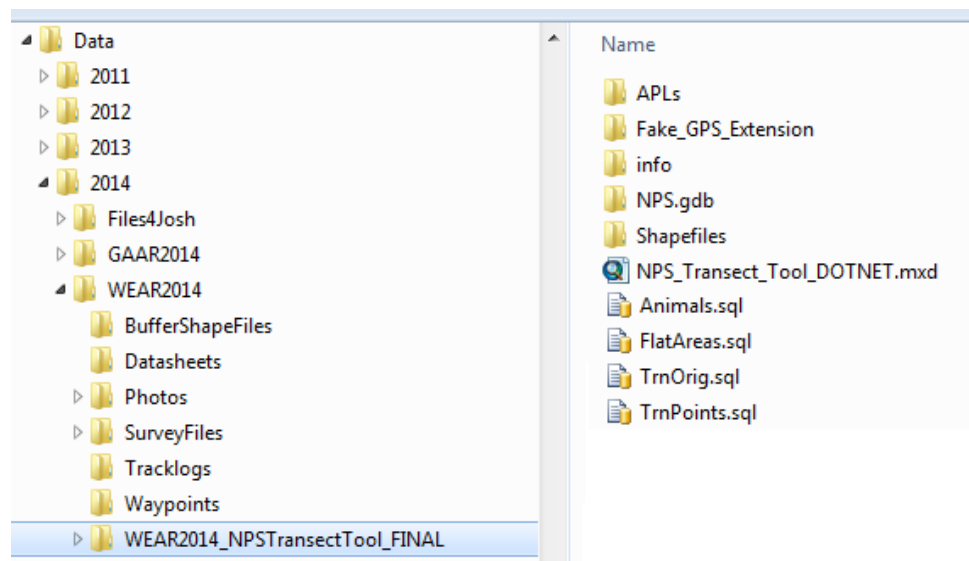


Figure SOP11-4. Example directory and file structure for the SQL scripts exported by the *NPSdotGDBtoSQL.py* tool. Each SQL script is named according to its source data layer in NPS.gdb. SQL scripts are generated in the same directory as NPS.gdb.

Please note that the NPSdotGDBtoSqlServer.py script and tool do not interact with the main sheep database in any way. There is no danger to the master sheep monitoring database in test running it as many times as needed to see, test and debug the outputs.

Once all the data from all the feature classes have been written to SQL script files you may run them using SQL Server Management Studio or Power Shell to insert the records into the database.

SOP 11.7.5.1 Data Import Part 1: Generate SQL INSERT Query Scripts Using NPS.gdb to SQL Server ArcToolbox Tool

Procedure:

1. The records to be imported from NPS.gdb into the sheep monitoring database must be related to a single Survey record. The ArcToolbox tool will require you to provide the SurveyID of this record (see screenshot below of the Surveys table in SQL Server Management Studio and the NPSdotGDBtoSQL.py tool overlaid). The red boxes show the information mappings between the Survey database record and the NPS.gdb to SQL ArcTool.

The screenshot shows a SQL Server Management Studio window with a query titled "SQLQuery1.sql - INPYUGAM...r (57))". The query is:

```
1 USE ARC_N_Sheep
2 SELECT * FROM Surveys ORDER BY Survey
```

Below the query is a table with the following data:

	SurveyID	Survey	ProtocolVersion
1	04C117B8-ECE2-46E0-9552-FE08B51C9612	2011 Denali Sheep Surveys	0.00
2	1AC66891-5D1E-4749-B962-40AB1BCA577F	2011 Itkillik Sheep Surveys	0.00
3	31A2F280-6794-4ECB-9F40-2D04352B17EA	2011 Western Arctic Parklands Sheep Survey	0.00
4	1B2FC2A0-FE84-42F2-BBE1-89FE940465BA	2011 Wrangell St. Elias Sheep Surveys	0.00
5	CD45C62E-89ED-4EEB-918E-64EF3EA490CD	2012 Itkillik Sheep Survey	0.00
6	B76669A3-4512-47FF-8F65-42FFF1C78DB0	2013 Itkillik Sheep Surveys	0.00
7	6F83E7F3-72F3-40A4-96A6-8F15A05E6327	2014 Gates of the Arctic Sheep Survey	0.00
8	50F08FE9-1A64-48F7-BF16-B2575B93B7D5	2014 Western Arctic Parklands Sheep Survey	0.00

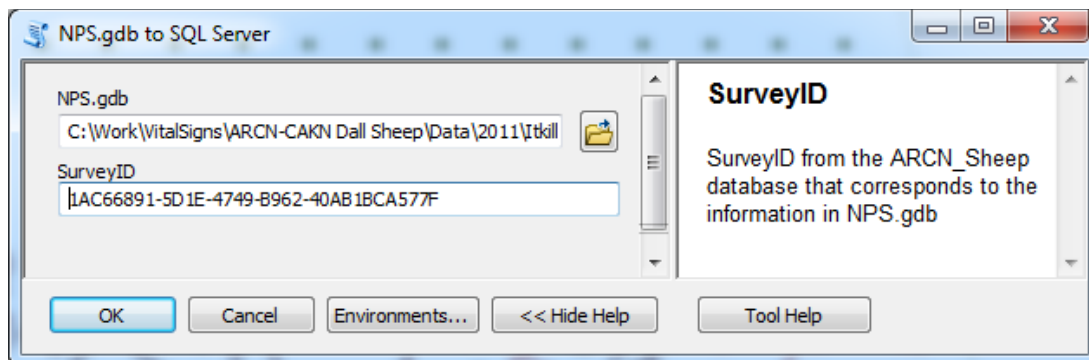
Overlaid on the bottom of the screenshot is the "NPS.gdb to SQL" dialog box. It contains the following fields:

- NPS.gdb: 14\WEAR2014_NPSTransectTool_FINAL\NPS.gdb
- SurveyID: 50F08FE9-1A64-48F7-BF16-B2575B93B7D5

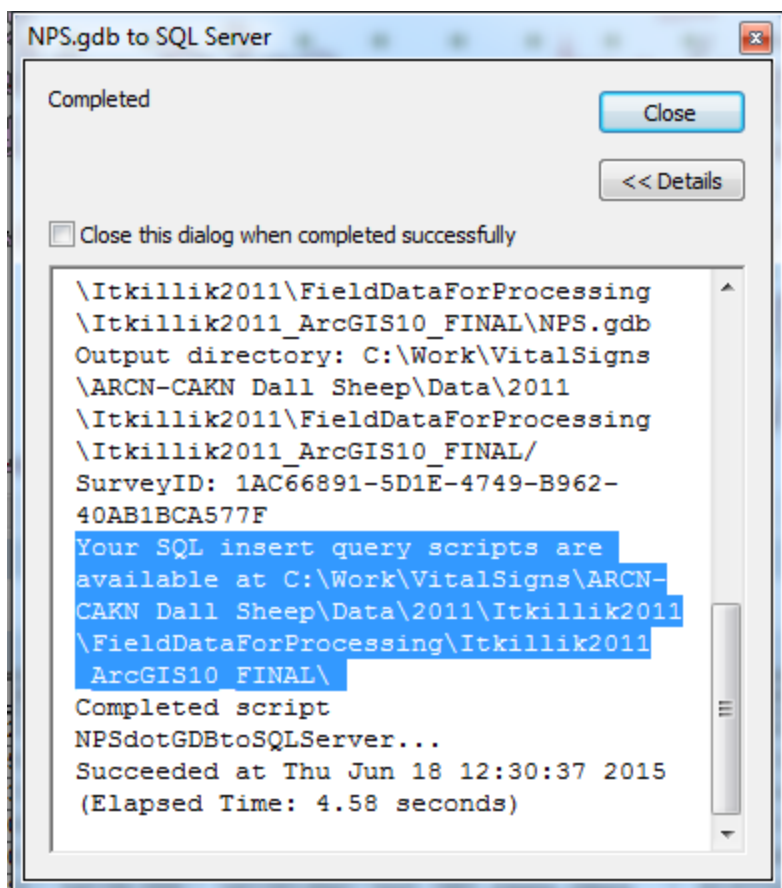
Red boxes highlight the SurveyID field in the dialog box and the corresponding SurveyID in the table above. Red arrows point from the SurveyID in the table to the SurveyID field in the dialog box.

The data entered into ArcToolbox tools must match a Survey database record.

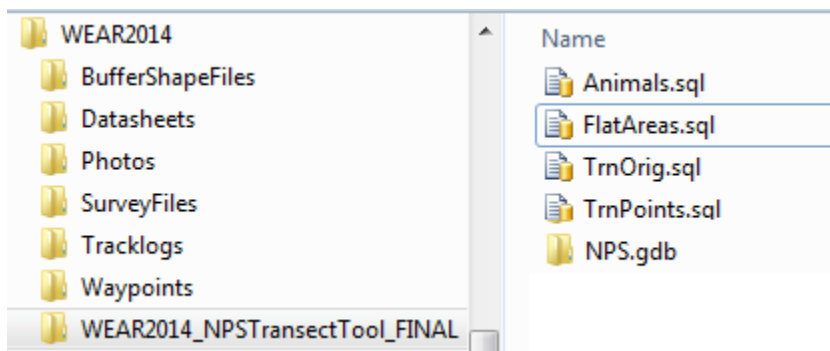
2. Download the sheep monitoring tools from the ARCN Sheep GitHub repository at <https://github.com/NPS-ARCN-CAKN/Sheep>.
3. Start ArcMap.
4. Open ArcToolbox.
5. Add the downloaded ARCN Sheep Tools toolbox to ArcToolbox. Consult the ArcGIS help if needed.
6. Ensure the NPS.gdb to SQL Server tool points to the NPSdotGDBtoSQLServer.py script on your computer by right clicking it and selecting Properties → Source.
7. Open the NPS.gdb to SQL Server tool and enter into the tool the path to NPS.gdb and the SurveyID from above.



8. Run the tool and examine the output for problems.



9. Inspect the output .sql files. They are sent to the same directory as NPS.gdb:



10. Examine the scripts in SQL Server Management Studio or other suitable interface and ensure they are as expected. Ensure you received a script from every required layer in NPS.gdb. *Note: Some layers such as tracklog points (GPSPointsLog from the gdb) are very time consuming to process. Such layers may be commented out in the underlying NPSdocGDBtoSQLServer.py script to improve processing performance. Open the Python script in an IDE or text editor and uncomment code according to your needs.*

SOP 11.7.5.2 Data Import Part 2: Running SQL INSERT Query Scripts

At this point you should have a series of SQL scripts ready to be executed to insert data from NPS.gdb into the sheep monitoring database. The next procedure will lead you through the process. You will require at least *datawriter* privileges on the sheep database.

Procedure:

1. Open SQL Server Management Studio (SSMS) or a similar tool.
2. Connect to the ARCN_Sheep database.
3. Start by opening the TrnOrig.sql script. Other scripts depend on these transects existing in the database so you must insert these records first. Here is what a typical script will look like in Windows Explorer and in SSMS:

```
1  -- Insert queries to transfer data from ARCN Sheep monitoring field geodat
2  -- File generated 6/18/2015 12:30:34 PM by SDMiller
3  -- If this file is too big to run in Sql Server Management Studio then run
4  USE ARCN_Sheep
5  BEGIN TRANSACTION -- Do not forget to COMMIT or ROLLBACK the changes after
6
7  -- insert the generated transects from TrnOrig -----
8  DECLARE @SurveyID nvarchar(50) -- SurveyID of the record in the Surveys ta
9  SET @SurveyID = '1AC66891-5D1E-4749-B962-40AB1BCA577F'
10 INSERT INTO [ARCN_Sheep].[dbo].[Transect_or_Unit_Information] (SurveyID,Ele
11 INSERT INTO [ARCN_Sheep].[dbo].[Transect_or_Unit_Information] (SurveyID,Ele
12 INSERT INTO [ARCN_Sheep].[dbo].[Transect_or_Unit_Information] (SurveyID,Ele
```

Note that the script starts by requesting use of the ARCN_Sheep database and opens a transaction. This transaction ensures that all the insert queries succeed or fail together.

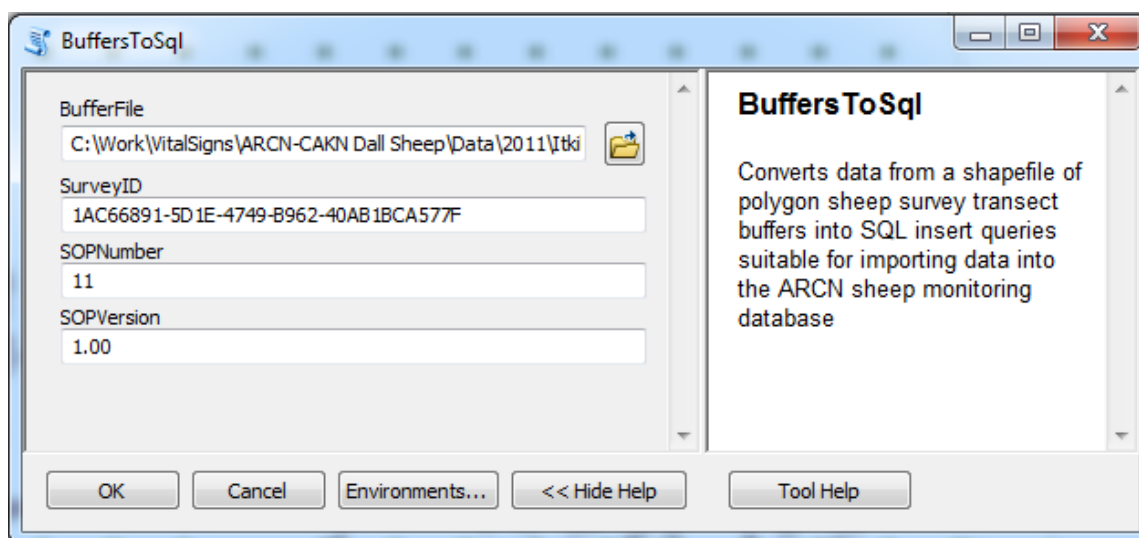
4. Ensure the @SurveyID parameter is the correct SurveyID to receive the data from NPS.gdb.
5. Execute the script.
6. Examine the output. If SSMS indicates all the queries succeeded then issue a COMMIT command. If any failed then issue ROLLBACK and fix the problems. *IMPORTANT: You must COMMIT or ROLLBACK once the transaction has been opened. Doing neither will leave the database in a hanging state.*
7. Repeat for each of the layers in NPS.gdb to be imported into the sheep database.
8. Check your work:
 - a. Run the DBQualityReport stored procedure In SQL Server Management Studio by executing the command `EXEC Dbqualityreport` (send the results to the screen or to a text file to examine the results). The results will give you an idea of how your import went, indicate any data that was omitted and any problem areas to fix.
 - b. Create or open ArcMap Query Layers to examine the data visually.
 - c. Run ad-hoc queries against the database to ensure the data imported correctly.

SOP 11.7.6 Importing Buffers

Import Buffers using the *ARCN Sheep Monitoring Data Management Tools* tool called *Buffers to SQL*.

Procedure:

1. Start ArcMap.
2. Add the *ARCN Sheep Monitoring Data Management Tools* ArcToolbox toolbox.
3. Open the *Buffers to SQL* tool.
4. Enter the path to the buffers shapefile, the SurveyID to which the buffers belong (ask the PI or data manager), the number and version of the SOP used to collect the buffers data (e.g., Buffers_Final feature class in the gdb).



5. Click OK to run the tool.
6. Examine the output. If there were problems fix them in the BuffersToSQLServer.py script and run the tool again until it works.
7. If the tool ran successfully it will have created a .sql text file with the same name as the buffers shapefile (or Buffers_Final feature class in the gdb) plus '.sql' in the shapefile's directory. The tool output will indicate where the file was created.
8. Open the .sql file generated by the tool in SQL Server Management Studio.
9. Verify the SQL, the SurveyID and that the queries are wrapped in a transaction.
10. Execute the script (you will need minimum *datawriter* permissions). If there were no problems execute COMMIT. If there were problems with any of the queries then execute ROLLBACK, fix the problems and try again until the queries all execute successfully as a batch.
11. Check your work:
 - a. Execute the DBQualityReport stored procedure and read the report
 - b. Using GIS, compare the buffers in the original Buffers_Final feature class from the NPS.gdb with the buffers in the database

SOP 11.7.7 Importing Pilot/Observer Waypoints and Tracklogs

Waypoints and tracklogs are important data components that are captured during a survey by the pilot with handheld GPS units, and these are used in the QA/QC process of reviewing the ArcPad collected data that have been imported into the NPS.gdb. The pilot's waypoints are a larger set than those collected in ArcPad and include sheep seen off-transect, sheep missed on-transect (which may be incorporated into analysis in the future), start and stop points for transects and other wildlife observations. The observer collects a tracklog on a handheld GPS unit as a second backup for the flight line (usually a Garmin GPS76csx which may be specified in the file name of the downloaded observer tracklog). Typically the observer does not collect waypoints on his/her handheld GPS as long as the field computer and ArcPad are working. If the field computer/ArcPad system crashes in the plane, then the observer will start using his/her GPS to also collect waypoints as a backup to the pilot's which becomes the primary for all locations including sheep on-transect (see SOP 8.19). If there was a problem with the pilot's tracklogs (or waypoints) then the observer's should be used in QA/QC of the NPS.gdb files. The section below covers the processing of pilot waypoints. The processing of waypoints and tracklogs captured by the ArcPad application, if required by the PI, are covered elsewhere (SOP 9).

This procedure assumes that the field data collection files have been produced, processed and organized on the Network file share according to SOP 9. The pilot/observer waypoint files will consequently be available in the /Waypoints directory and the pilot and observer tracklogs in the /Tracklogs directory.

SOP 11.7.7.1 Importing Waypoints Using the WaypointsToSQL.py ArcTool

Procedure

Use the Waypoints to SQL tool in the ARC� Sheep Monitoring Data Management ArcToolbox to convert the pilot's waypoints into a script file of SQL insert queries.

Important note on duplicate waypoints: Unless a pilot/observer clears their GPS on a daily basis waypoints will accumulate. Each day's GPS export file will consequently be a partial duplicate of the one exported the day before. The database has a unique constraint to disallow such duplicates but considerable time and effort can be saved by examining all the waypoint files for a particular pilot during a survey and determining if the last waypoints export file contains all the waypoints collected during a survey. If such is the case you may simply import the data from the latest waypoints file.

Procedure

1. Start ArcMap
2. Add the ARC� Sheep Monitoring Data Management ArcToolbox to the Toolboxes palette.
3. Start the Waypoints to SQL tool (Figure SOP11-5, below).
4. Supply the required information:
 - a. Waypoints shapefile you would like to convert to insert queries.
 - b. SurveyID of the Survey in the master database to which the waypoints should be associated.

- c. Pilot's name.
 - d. Aircraft tail number.
 - e. Waypoints source. Defaults to PILOT GPS. May be OBSERVER GPS if the pilot's waypoint file is not usable.
 - f. Standard Operating Procedure number describing how the waypoints were collected
 - g. Version of the SOP above.
5. Click OK.
 6. Examine the source waypoints shapefile directory for an SQL script file named the same as the input shapefile with a '.sql' suffix (Figure SOP11-6, below).
 7. Execute the SQL scripts in SSMS inside a transaction.
 8. If there were no errors reported by SSMS COMMIT the insert queries, otherwise ROLLBACK and fix any problems.
 9. Query the PilotWaypoints table and/or examine the waypoints in a GIS to ensure the import process happened correctly.

Waypoints to SQL

Pilot waypoints shapefile
C:\Work\VitalSigns\ARCN-CAKN Dall Sheep\Data\2013\Itkillik2013_AerialSheepSurvey\Waypoints\SH12Jul2013_Waypoints.shp

SurveyID
B76669A3-4512-47FF-8F65-42FFF1C78DB0

Pilot Name
Sandy Hamilton

Tail Number
N3967

Waypoints Source
PILOT GPS

SOP Number
8

✔ SOP Version
d

OK Cancel Environments... << Hide Help

Figure SOP11-5. Waypoints to SQL ArcTool.

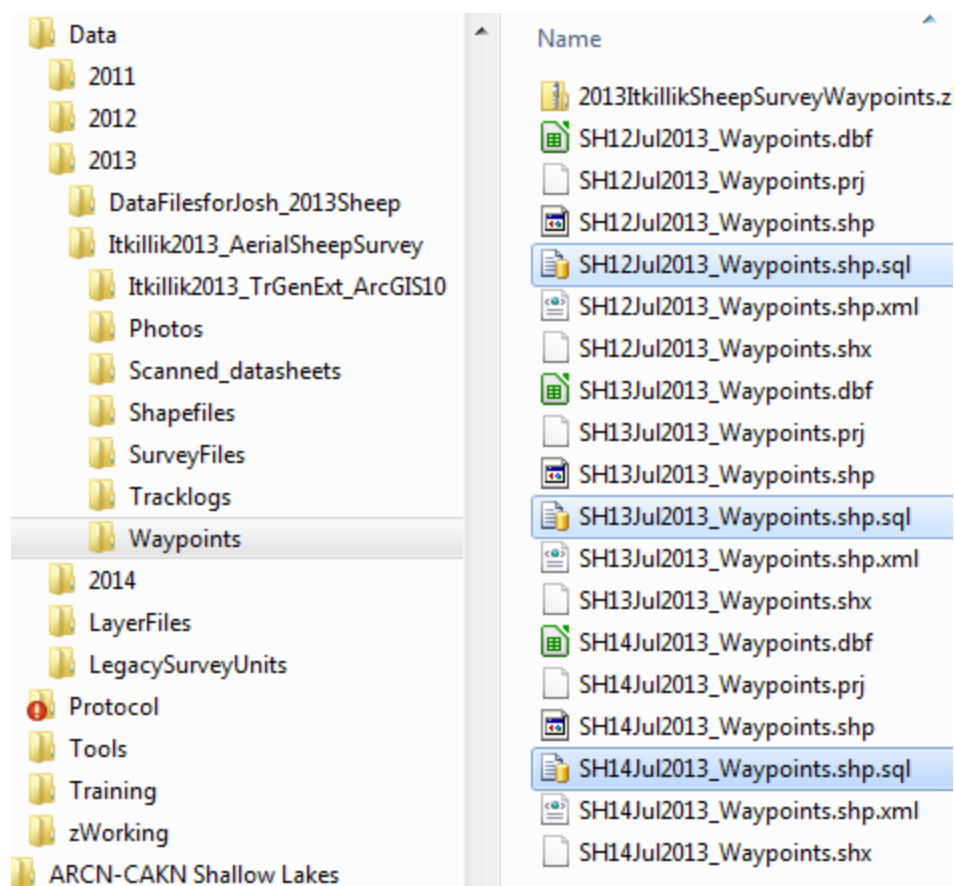


Figure SOP11-6. WaypointsToSQL ArcTool output SQL scripts. The .sql files are named the same as the input shapefiles with a '.sql' suffix.

The procedure above described running the WaypointsToSQL ArcTool. The tool may also be run in batch mode by right clicking the tool and selecting batch mode. In batch mode the user must supply the needed information in a grid with one input file per row (Fig. SOP11-7, below).

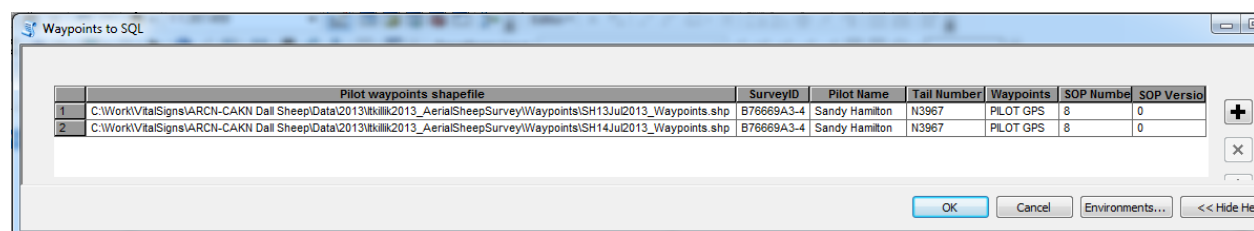


Figure SOP11-7. WaypointsToSQL ArcTool example in batch mode.

SOP 11.7.7.2 Importing Waypoints Using a Spreadsheet

The WaypointsToSQL tool is the best way to convert pilot waypoints into SQL insert scripts but another approach exists for users familiar with spreadsheets and formulas. The data in a shapefile such as the pilot's waypoints shapefile is stored in its .dbf file. Most spreadsheets such as Microsoft Excel or LibreOffice Calc are able to open and manipulate a .dbf file (it is recommended to only do this on a copy of the dbf file to avoid corrupting the source shapefile). A user could then create a formula in an empty column that will generate an insert query for each

row using the formula in Equation 1 as an example. These queries can then be executed against the sheep monitoring database to insert the records. Care must be taken to ensure foreign key constraints are satisfied depending on which database table the records are being inserted into. See your spreadsheet's documentation for help on formulas.

Equation 1. Example of a spreadsheet cell formula to generate an SQL insert query suitable for executing against the master sheep monitoring database.

```
= "INSERT INTO  
PilotWaypoints (SurveyID,WaypointName,PilotName,CaptureDate,Altitude,GPSModel,  
SourceFilename,PointFeature,Comments,TailNo) VALUES ('1AC66891-5D1E-4749-B962-  
40AB1BCA577F', '" & B2 & "', 'L. Williams', '" & AG2 & "', " & N2 & ", '" & AE2 &  
"', 'LW03-06Jul2011_ItkillikSheep_AllWaypoints_WGS84.dbf',  
geography::STGeomFromText('POINT(" & D2 & " " & C2 & ")',4326), '" & G2 & "',  
, '" & E2 & "');" 
```

In the formula above the waypoint name is found in cell B2 and the CaptureDate comes from AG2 and so on. Other columns are mapped accordingly. Some columns are hard-coded such as the SurveyID ('1AC66891-5D1E-4749-B962-40AB1BCA577F') and 'L. Williams'. Once the formula is ready, dragging it down the INSERT query column will replicate the query for each record substituting the correct values. Copy the formulas into SSMS and execute within a transaction to insert the records. See your spreadsheet and SSMS help for more information.

Please note the following regarding the formula above:

- The column mappings may be different for your spreadsheet.
- The SurveyID and SourceFilename column values are hard-coded into the query. You will have to look up the SurveyID in the database to ensure the imported waypoints are associated with the correct survey. The source file name should be the name of the waypoints file.
- The WKT string submitted to the STGeomFromText() function is a Lon/Lat (not Lat/Lon) and there is a space between the values, not a comma.

SOP 11.7.7.3 Importing Pilot Tracklogs

Pilot tracklogs are imported using the 'Pilot tracklog to SQL' (PilotTracklogToSQL.py) ArcToolbox script tool.

Procedure

1. Start ArcMap
2. Add the ARCN Sheep Monitoring Data Management ArcToolbox to the Toolboxes palette.
3. Start the 'Pilot tracklog to SQL' tool (Figure SOP11-8, below).
4. Supply the required information:
 - a. Tracklog shapefile you would like to convert to insert queries.
 - b. SurveyID of the Survey in the master database to which the tracklog should be associated.

- c. Pilot's name.
 - d. Aircraft tail number.
 - e. Tracklog source. Defaults to PILOT GPS. May be OBSERVER GPS if the pilot's file is not usable.
 - f. Standard Operating Procedure number describing how the tracklog was collected
 - g. Version of the SOP above.
5. Click OK.
 6. Examine the source tracklog shapefile directory for an SQL script file named the same as the input shapefile with a '.sql' suffix.
 7. Execute the SQL scripts in SSMS inside a transaction.
 8. If there were no errors reported by SSMS COMMIT the insert queries, otherwise ROLLBACK and fix any problems.
 9. Query the PilotTracklogs table in the ARCN_Sheep database and/or examine the waypoints in a GIS to ensure the import process happened correctly.

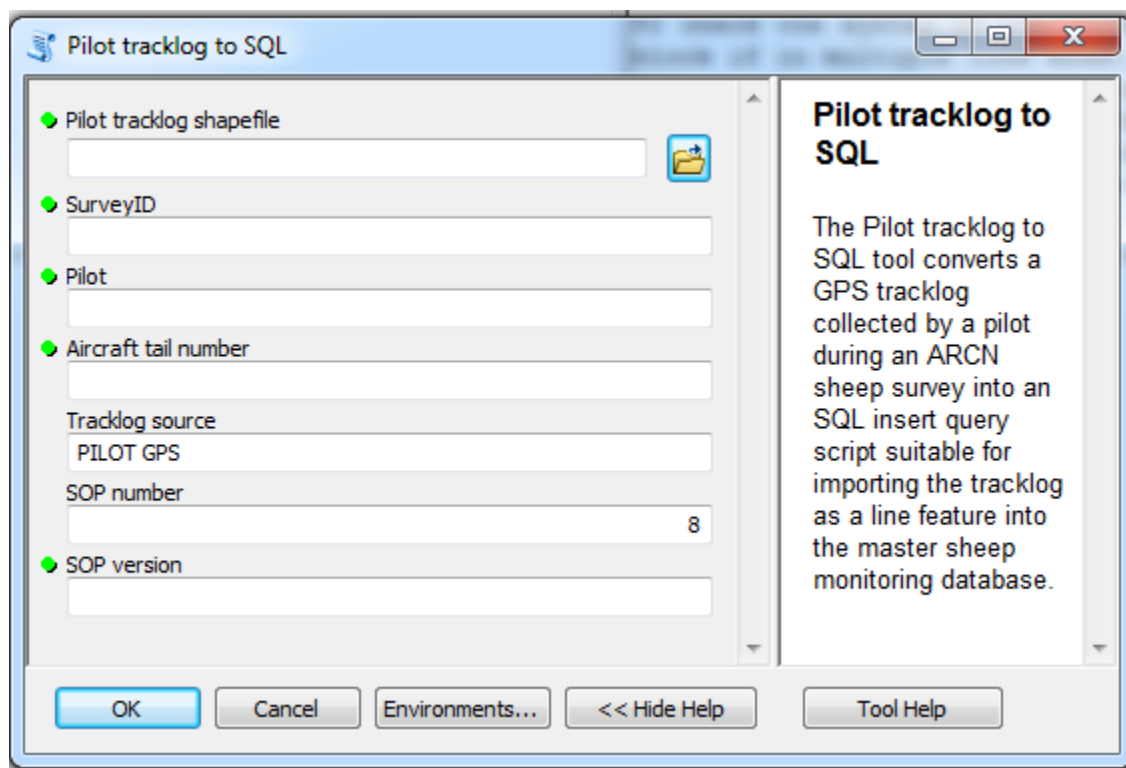


Figure SOP11-8. Pilot tracklog to SQL tool.

SOP 11.7.8 Animal Group Images

The sheep monitoring database was designed to handle sheep group images but as of this writing we determined that the benefit provided by storing images or image metadata in the database would not be worth the significant amount of work to make it happen. Images should be organized, documented and maintained in the sheep monitoring directories on the Network file server (SOP 9, SOP 10). If images are organized and documented according to the SOPs it will be possible to re-trace our steps in using them to extract data. If extra time and resources become available or a stronger need arises for integration of images with database records we will develop such procedures.

SOP 11.8 – SOP Executing Large SQL Files

SQL Server Management Studio automatically validates each query in each script that you open which is memory intensive and may cause 'out of memory' type errors for large files. GPS point .sql scripts map reach hundreds of thousands of queries and hundreds of Mb in size and will crash SSMS. The solution is to run the script from a Windows Powershell. Example:

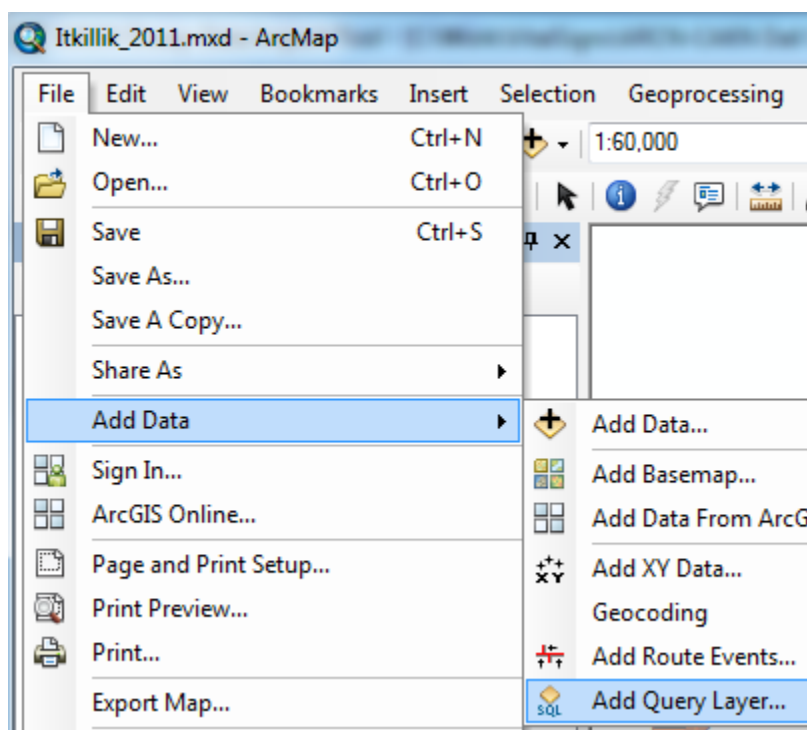
```
sqlcmd -S SERVER\NAME -E -i "C:\GPSPointsLog.sql"
```

SOP 11.9 – Creating Arcmap Query Layers

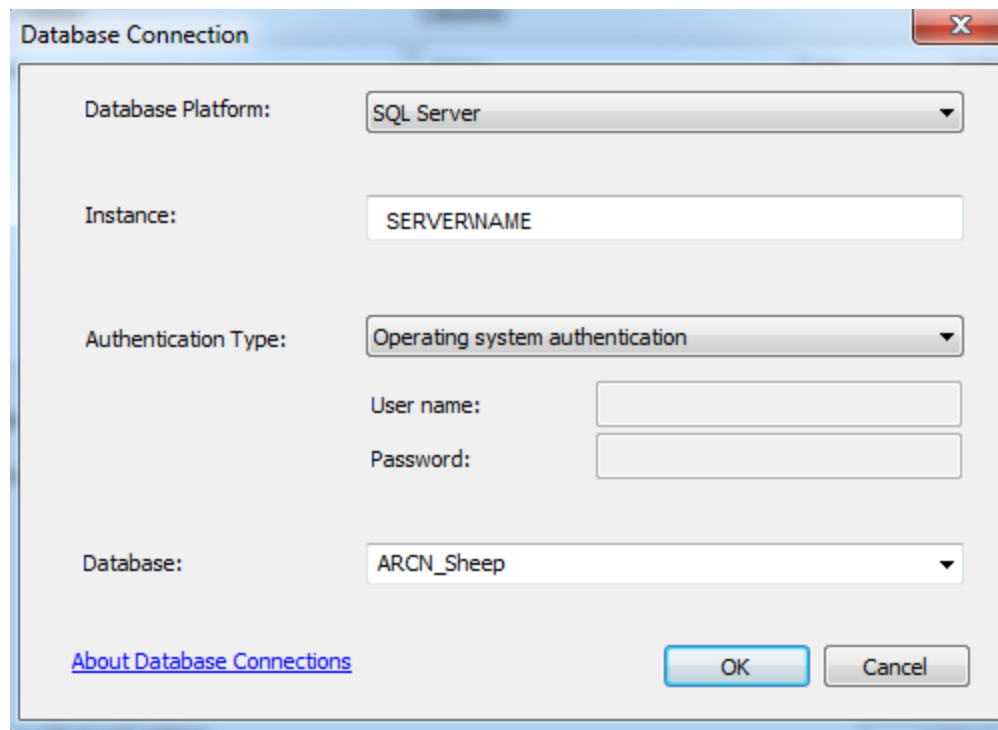
While other GIS software may be used to interact with the spatial data in the sheep monitoring SQL Server database the standard GIS in the NPS is ArcGIS. This is typically done using Query Layers. Consult the ArcGIS help for an overview of Query Layers. This section describes how to load sheep monitoring spatial data into ArcMap using Query Layers. Once a Query Layer has been created you may save it as a layer file.

Procedure:

1. Start ArcMap
2. Click File → Add data → Add Query Layer...



3. If you haven't made a connection to the ARCN_Sheep database do so now supplying the server and database names.
4. Select Operating system authentication from the Authentication Type dropdown box (this option indicates that you will log into the SQL Server using your Windows Active Directory credentials, obviating the need for a username and password).



Database Connection

Database Platform: SQL Server

Instance: SERVERNAME

Authentication Type: Operating system authentication

User name:

Password:

Database: ARCN_Sheep

[About Database Connections](#)

OK Cancel

5. Close the Database Connection window
6. Using the Connection you created above choose the table containing spatial data that you would like to query and drag it down to the Query text area as in the example below which will pull data from the Transect_or_unit_Information table. **Make sure that you request only one geography column per query or ArcMap will return an error or hang.** *Note: ArcMap's query builder is not very good – it confuses geometry columns with geography columns, omits columns from the column chooser and generally is difficult to use. The best thing to do is to craft your spatial query in SQL Server Management Studio and then copy it into the Query text box.*

New Query Layer

Connection: Connection to inpyugams08sql.sde Connections...

List of Tables:

Name
ARCN_Sheep.dbo.LegacyUnits
ARCN_Sheep.dbo.PilotWaypoints
ARCN_Sheep.dbo.Surveys
ARCN_Sheep.dbo.Transect_or_Unit_I
ARCN_Sheep.dbo.TransectImages
ARCN_Sheep.dbo.TransectTracklog

Columns:

Name	Type	Nullable
Aircraft	Text	No
CenterPoint	Geometry	No
CloudCover	Text	No
ElapsedTime	Long Integer	No
Elevation_M	Long Integer	No
Flown	Text	No
FlownDate	Date	No
GeneratedTransect	Geometry	No

Name: Transects

Query:

```
SELECT [TransectID]
,[SurveyID]
,[Elevation_M]
,[Aircraft]
,[ObserverName1]
,[ObserverName2]
,[PilotName]
,[CloudCover]
```

☐ Show advanced options Validate

< Back Finish Cancel

7. Give the query a name. This name will appear in the map's table of contents.
8. Click Validate
9. In this example TransectID is the primary key for the table we're querying and is thus guaranteed to be unique so I unchecked all the other columns in the Unique Identifier Field(s) box. I selected Geometry Type: Line, changed the SRS to WGS84 which has an SRID (or EPSG code) of 4326 (all tables in the ARCN_Sheep database store data in WGS84 SRS so this will apply to any query layers you care to build).

New Query Layer

Unique Identifier Field(s):

	Name	Type	Nullable
☒	TransectID	Text	No
☐	SurveyID	Text	No
☐	Elevation_M	Long Integer	No
☐	Aircraft	Text	Yes
☐	ObserverName1	Text	Yes
☐	ObserverName2	Text	Yes
☐	PilotName	Text	Yes

Spatial Properties

☐ Coordinates include M values. Used to store route data.

☐ Coordinates include Z values. Used to store 3D data.

Geometry Type: Line

Spatial Reference: GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID["WGS_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]] Change

SRID: 4326

< Back Finish Cancel

10. Click Finish
11. Overlay the query layer with its source layer from NPS.gdb to ensure they synchronize and overlap correctly.
12. Optionally, save query layer as a layer file so you can just drag and drop it into ArcMap the next time you need it. The query layer will save all the connection information.

SOP 11.10 – Final Quality Assurance/Control Procedures

Primary tools used for assuring the quality of sheep monitoring data include manual data verification checks, ad-hoc data visualization checks and running programmatic database queries, scripts and stored procedures such as DBQualityReport. Analytical and statistical analyses will also reveal data quality deficiencies and should trigger another round of data quality assurance checks.

Validation and quality control efforts should ideally be completed by mid-November for data from the previous summer's surveys. Once the data have been validated by a project manager, they can be certified and analyzed.

SOP 11.11 – Data Preparation for Analysis

Analysis of the sheep distance data relies on Bayesian modeling techniques and is detailed in SOP 12, Schmidt et al. (2012) and Schmidt and Rattenbury (2013). Please note that the scripts used to run the models in WinBUGS or OpenBUGS (via a program R interface) can be found online as supplemental information for Schmidt et al. (2012) and Schmidt and Rattenbury (2013). Validation is also an iterative process with analysis. If the estimates show anomalies or trends that do not make sense (basically following validation of analyzed data), the data will be re-verified, re-validated and re-analyzed. If results are reported from data that have not been fully validated (e.g., preliminary results before validation of the model estimates), a statement and explanation should be included to caution users regarding data quality.

Program Distance, R and WinBUGS or OpenBUGS are the primary tools for analyzing sheep data. Database views have been created for most analytical formats in the ARCN_Sheep SQL Server database (Table 3, below). Alternatively, formatted data used in analysis can also be produced from the attributes table of feature classes in the NPS.gdb. Use the final, fully QA/QC'ed version of the NPS.gdb to produce datasets from those attributes tables.

Table 3. Analytical data formats, tools and database views.

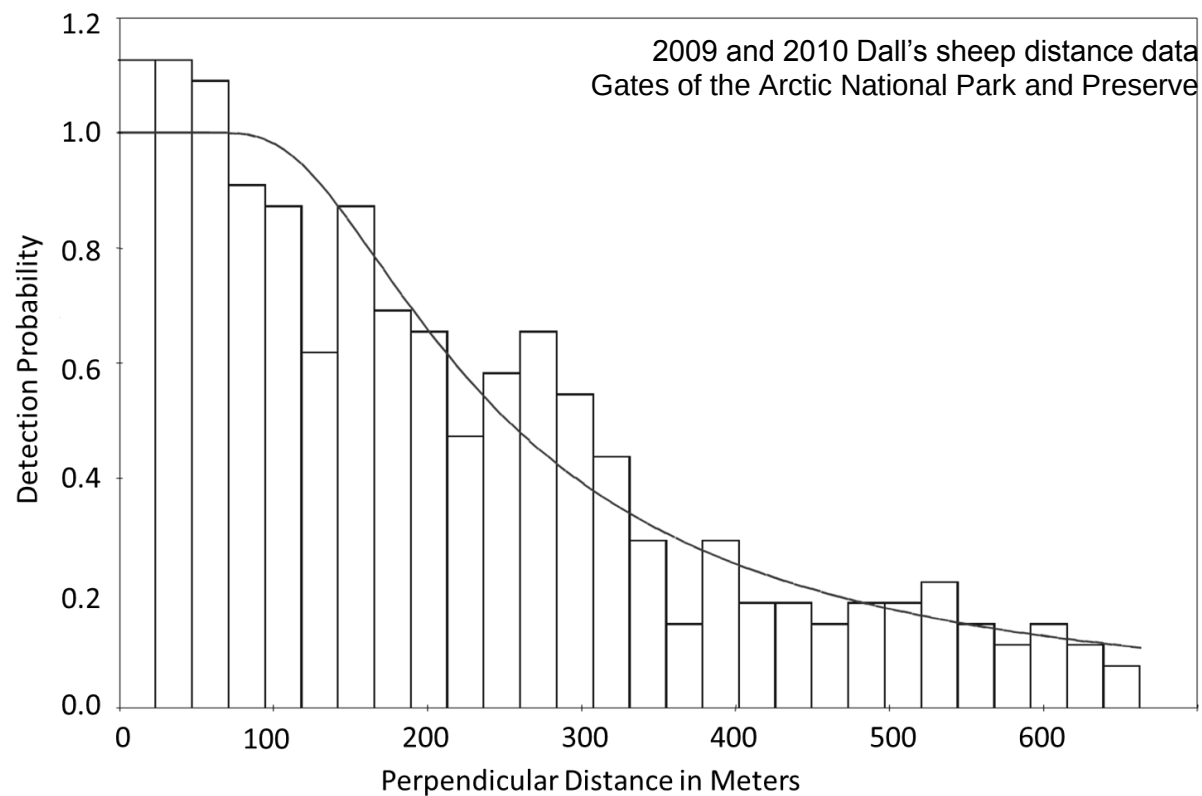
Data Format	Tool	SQL database view or associated feature class/attributes table in NPS.gdb
Group sizes and distances (SOP 11.11.1)	Program Distance	GroupSizeAndDistance from SQL or derived from DB Animals
Group size totals and number per sex-age class for program (SOP 11.11.2)	R/WinBUGS/OpenBUGS	None or derived from DB Animals
Group distances (SOP 11.11.3)	R/WinBUGS/OpenBUGS	GroupDistance or derived from DB Animals
Transect data including survey conditions and adjusted length (SOP 11.11.4)	R/WinBUGS/OpenBUGS	TransectData or derived from Generated Transects (TrnOrig)
Transect center point coordinates in meters (SOP 11.11.5)	R/WinBUGS/OpenBUGS	TransectData or derived from Random Points (TrnPoints)

Examples of the formats from Table 3 are included in each of the following sections. For those files produced for program R/WinBUGS or OpenBUGS, it is important to note that even if transects weren't surveyed, or there were no sheep detected on the line, those transects still need to be represented (i.e., a data row) with accompanying data (e.g., '0' or NULL). You will also need to provide survey area (m²), including for any subareas for which you want separate estimates (e.g., for the western Baird Mountains as part of a larger WEAR survey).

SOP 11.11.1 Group Size and Distance for Program Distance

Program Distance can be used to produce a histogram of your distance data, to look for violation of the assumption of 100% detection on the line, to determine an effective strip width as the farthest 5% of detected groups are eliminated prior to analysis, to determine the shape of the detection function and to produce a preliminary estimate prior to Bayesian modeling if desired (Schmidt et al. 2012). Schmidt et al. (2012) analyzed the 2009 and 2010 distance sampling data from Dall's sheep surveys in Gates of the Arctic National Park and Preserve (GAAR) with Program Distance and determined that a hazard-rate detection function was best suited to those data, and they set an effective strip width at 663 m (see histogram below). That detection function and strip width will be applied to future surveys.). However, it is not unwise to produce

the group distance and size file and plot your data in Program Distance in the event that your data vary considerably from those collected in 2009 and 2010 in GAAR.



This Program Distance file will consist of one sheep group per row with columns for Survey ID (typically the NPS unit code and year), Transect ID, Animal ID (sheep group for that transect), total sheep in the group and group distance to the transect (see sample below of DENA2011 data). These data can be exported from the GroupSizeAndDistance database view from the ARCN Sheep SQL Server database or from the attributes table of the final DB Animals file of the NPS.gdb. The histogram is produced from the distance to transect column; the y-axis is the count of groups within each distance bin on the x-axis. The other columns are simply for data organization. You do not need to enter rows for transects that were not flown or where there were no sheep observed. You may produce similar data files and histograms for each sex-age class as well.

Survey ID	Transect ID	Animal ID	Total	Distance
WRST2010	5	1	2	450
WRST2010	5	2	4	508
WRST2010	5	3	1	442
WRST2010	5	4	1	404
WRST2010	5	5	2	88
WRST2010	5	6	2	0
WRST2010	6	1	4	121
WRST2010	9	1	2	0

Survey ID	Transect ID	Animal ID	Total	Distance
WRST2010	9	2	3	10
WRST2010	9	3	2	17
WRST2010	9	4	3	160
WRST2010	9	5	1	96
WRST2010	9	6	6	0
WRST2010	9	7	12	408
WRST2010	10	1	5	71
WRST2010	10	2	2	237
WRST2010	10	3	19	272
WRST2010	10	4	8	158

SOP 11.11.2 Group Size and Sex-Age Classes for Program R/WinBUGS/OpenBUGS

Program R is the interface for WinBUGS or OpenBUGS where the distance data are analyzed within a Bayesian framework (Schmidt et al. 2012, Schmidt and Rattenbury 2013). All data files destined for Program R should be in the format of one transect per line (row) regardless of whether the transect was flown or there were sheep seen on the transect. The columns for this file type are thus Survey ID, Transect ID, Flown status (Y or N), and one group per column with the maximum number of columns dependent upon the maximum number of groups seen on any one transect during the survey (or among all surveys to be compared in the same analysis session).

You can produce an Excel file of these data from the attributes table of the DB Animals feature class in your final NPS.gdb. The Excel file can then be converted into a matrix text file for program R/WinBUGS/OpenBUGS. Note the different format from the Program Distance file created in SOP 11.11.1. Similar files should be produced for each sex-age class (ewe-like, lambs, < full-curl rams, ≥ full-curl rams), however, if there were no representatives of a specific class in the group (e.g., no rams in a ewe-lamb nursery group), the value for that cell must be 0, not NULL (or blank).

Survey ID	Transect ID	Flown	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Etc.
WRST2010	1	Y								
WRST2010	2	N								
WRST2010	3	Y								
WRST2010	4	N								
WRST2010	5	Y	2	4	1	1	2	2		
WRST2010	6	Y	4							
WRST2010	7	Y								
WRST2010	8	Y								
WRST2010	9	Y	2	3	2	3	1	6	12	
WRST2010	10	Y	5	2	19	8				

SOP 11.11.3 Group Distance for Program R/WinBUGS/OpenBUGS

The Group Distance File for Program R and WinBUGS (or OpenBUGS) should correspond with the Group Total Size File (see SOP 11.11.2) where there is one row per transect and columns for each of the groups up to the maximum number of groups seen on any one transect, regardless of whether the transect was flown or sheep were seen on the transect. The distances should be queried from the Animals database table in the ARCN Sheep SQL Server database. Alternatively, you can produce an Excel file of these data from the attributes table of the DB Animals feature class in your final NPS.gdb. The Excel file can then be converted into a matrix text file for program R/WinBUGS/OpenBUGS. Please see the sample format below.

Survey ID	Transect ID	Flown	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Etc.
WRST2010	1	Y								
WRST2010	2	N								
WRST2010	3	Y								
WRST2010	4	N								
WRST2010	5	Y	450	508	442	404	88	0		
WRST2010	6	Y	121							
WRST2010	7	Y								
WRST2010	8	Y								
WRST2010	9	Y	0	10	17	160	96	0	408	
WRST2010	10	Y	71	237	272	158				

SOP 11.11.4 Transect Data for Program R/WinBUGS/OpenBUGS

Transect data for use with R/WinBUGS/OpenBUGS can be queried and exported from the database view called TransectData in the ARCN Sheep SQL Server database. This view provides sample area (buffer area) per transect, the effective transect length (buffer area divided by effective strip width) and parameter data (e.g., elevation, park or year designation, weather data, etc.) for the model (see sample format table below).

Alternatively transect data for analysis can be produced as an Excel file from the attributes table of the Generated Transects (TrnOrig) feature class in the NPS.gdb.

- 1) Export the Attributes Table for the Generated Transects feature class to a .csv or text file and open and save it in Excel.
- 2) Retain the following columns (fields): Survey ID (NPS unit and year), Transect ID, Pilot, Observer, Elevation (ELEVFT), Flown (Y/N), Cloud Cover, Precipitation, Turbulence Intensity, Turbulence Duration, Temperature (*F), Year and Flown Date. Delete all other columns (fields).
- 3) Insert two columns – one for Buffer Area (m²) and the second for Effective Length (m). Import the buffer areas from the exported Attributes Table of your final buffer shapefile (SOP 11.6). Check to make sure that your buffer areas match the appropriate Transect ID. For the Effective Length column, divide the Buffer Area values by your effective strip

width (e.g., 663m in Schmidt et al. [2012]). Make sure that the Effective Length column is retained as values not formulas for conversion to the text matrix for program R.

- 4) Insert a column for Turbulence. Values in this column depend on your Turbulence Intensity and Turbulence Duration fields. If the value is 'none' for both intensity and duration, assign a '0' to the new Turbulence column. All other values for intensity and duration (no matter what the value) translate to '1' in your new Turbulence column. In other words, we will be analyzing the data on the variable of turbulence present (1) or not (0). You can then delete the Turbulence Intensity and Duration columns if desired.
- 5) You may rearrange the columns as shown below. Note that Observer, Flown Date, Flown (Y/N) and Year are not shown in this example, but transects 2 and 4 were not flown and thus have null values for all columns. Estimates will be produced for all transects and applied to the entire survey area so you should list the intended Survey ID and year for all transects.
- 6) If you intend to produce estimates for subareas within your larger survey area (e.g., for the Itkillik in a year when all of GAAR was surveyed), make another column for that survey area and indicate yes/no or 1/0 whether or not each transect was inside that subarea.

Survey ID	Transect ID	Buffer Area (m ²)	Effective Length (m)	Pilot	Elev. (ft.)	Cloud Cover	Precip-itation	Turbu-lence	Temp (F)
WRST2010	1	12937826.94	19514.07	Barnes	3415	Broken	None	1	49
WRST2010	2								
WRST2010	3	12536661.82	18908.99	Lee	3868	Scattered	None	1	55
WRST2010	4								
WRST2010	5	12688063.58	19137.35	Barnes	3960	Broken	None	1	49
WRST2010	6	11810417.12	17813.60	Williams	5890	None	None	0	45
WRST2010	7	13388231.66	20193.41	Williams	4164	None	None	0	50
WRST2010	8	11771398.58	17754.75	Lee	4522	Broken	None	0	49
WRST2010	9	12581905.97	18977.23	Lee	5720	Broken	None	1	45
WRST2010	10	12763898.64	19251.73	Lee	4772	Broken	None	1	45

If your transect data were exported from the TransectData view in the ARCN Sheep SQL Server database, R will connect directly to the view via the RODBC package or you can export the data to a file. If you produced an Excel file from the Generated Transects feature class, the Excel file can then be converted into a matrix text file for program R/WinBUGS/OpenBUGS.

SOP 11.11.5 Transect Center Point Coordinates for Program R/WinBUGS

Schmidt et al. (2012) considered spatial correlation in the analysis of the 2009 and 2010 data from the GAAR sheep surveys. Probability of presence was assumed to be affected by the presence of sheep detected on neighboring transects. A file containing transect center point coordinates (in meters) was one of the model inputs whereby data from the eight nearest transects surrounding each transect was allowed to influence the probability of presence on the

center transect. This was not included in the analysis of the 2011 data (see Schmidt and Rattenbury 2013), but may be a factor in future surveys due to the clumping nature of habitat use by Dall's sheep and differential distribution of sheep across the landscape.

This Center Point Coordinates file can be produced from the Attributes Table of the Random Points (or TrnPoints) feature class in the geodatabase (see the PROJTD_X and PROJTD_Y fields). However, the Random Points feature class (SOP 2.3.7) is produced before the Generated Transects (SOP 2.3.8), therefore you must match the OBJECTID_1* fields from each or do a proximity match to line up the correct coordinates from the Random Points file to the transect ID from the Generated Transects file. Be sure to generate coordinates (e.g., using the Feature to Point tool in ArcMap) for those transects which are the second (or third) part of multi-part transects (e.g., their Transect IDs are the original [first part] transect number plus 1000 [or 2000], see SOP 2.3.8, Step 7 and Figure SOP2-17). A sample Excel file showing transect center point coordinates is shown below.

Survey ID	Transect ID	Flown	PROJTD_X	PROJTD_Y
WRST2010	1	Y	539152.69	1464545.69
WRST2010	2	Y	548833.08	1465735.18
WRST2010	3	Y	539854.86	1456833.72
WRST2010	4	Y	549354.73	1456084.81
WRST2010	5	Y	558935.28	1459078.55
WRST2010	6	Y	564472.57	1453807.16
WRST2010	7	Y	574865.37	1455995.94
WRST2010	8	Y	538756.77	1446881.77
WRST2010	9	Y	549944.19	1448280.64
WRST2010	10	Y	557526.78	1447691.84

SOP 11.12 – Seasonal Close Out and Dataset Certification

Data certification is a benchmark in the monitoring project data life cycle that indicates that: a) the data are complete for the period of record; b) they have undergone and passed the quality assurance checks; and c) that they are appropriately documented and in a condition for archiving, posting and distribution as appropriate.

Seasonal close out is the best time to bring SOPs up to date and version their IRMA references if new data management tools or methods have been developed.

Procedure:

1. Verify the sheep monitoring protocol and SOPs are all up to date and versioned in IRMA.
2. Verify QA/QC procedures have been performed to the satisfaction of the project leader.
3. Verify deliverables are documented and archived in IRMA with metadata (SOP 11.13).
4. Review data products and IRMA references for sensitive information and isolate them appropriately.

5. The sheep monitoring program should have a single dataset reference in IRMA to which the latest certified snapshot of the monitoring database is attached. If this reference does not exist the PI or data manager should create it. Ensure it is tagged 'sensitive' and 'operations'. The PI and data manager fill out the reference's metadata.
 - a. The PI fills out a data certification form and signs it (Appendix C). This form is scanned and uploaded to the dataset IRMA reference above.
 - b. A human readable copy (SQL script) of the main sheep monitoring database is uploaded to the sheep monitoring program's dataset reference and versioned.
 - c. A copy of the sheep monitoring SQL Server database files are attached to the I&M SQL Server instance in Ft. Collins, CO. This copy is the certified database.

SOP 11.13 – Archiving Field Deliverables in IRMA

Field data files should be archived on IRMA in a private Dataset reference with each deliverable from the data deliverables schedule. (See narrative.)

SOP 11.13.1 Creating a Field Deliverables Dataset Reference in IRMA

1. Create a new IRMA Dataset Reference and title it informationally similar to '2015 Gates of the Arctic Sheep Survey Dataset'.
2. Fill out as much metadata as possible.
3. **Ensure the Dataset is tagged as 'Sensitive' under the Information Resource Evaluation tab with the justification of 'Commercial Value'.** Also tick the 'Operations' justification as we don't want raw, unprocessed data available to the public (certified data will be available minus the sensitive data through the main sheep monitoring dataset in IRMA).

Step 2: Create a Generic Dataset (Required)

Core **Information Resource Evaluation** Units and Geographical Keywords Permissions

Sensitivity Evaluation

☐ Non-Sensitive ☒ Sensitive ☐ Unknown

If Sensitive, Check 1 or More

☒ Operations	☐ Archeological	☐ Endangered	☐ Security
☐ Personal Info	☐ Indian Sacred Sites	☐ Threatened	☒ Commercial Value
	☐ Wells	☐ Cultural	☐ Paleontological

4. Upload the data deliverables from the Deliverables Schedule above, one by one. Tag them with the date and deliverable ID in the filename and especially the holding description. An example would be 'DS08: Survey data sheets'.

Upload New Files

Name	Status	Size	Progress
NPS.gdb.zip	Complete	10 MB	clear

Total: 1 10 MB of 10 MB

Select Files Upload Clear All Rows

1 File Uploaded

File Name: NPS.gdb.zip

Description: DS01: NPS.gdb (ESRI personal geodatabase, 10mb)

5. Remove all permissions except for the PI and data manager (to reduce the risk of accidental publication of sensitive data).
6. Activate the reference.
7. Note the ReferenceCode.
8. Open the main sheep monitoring IRMA reference at <https://irma.nps.gov/App/Reference/Profile/2214983>
9. Navigate to the Products section and click Manage Products.
10. The Manage Products form appears. Use the form's tools and the ReferenceCode noted above to search for, and link to the dataset created above.
11. Save and exit the Manage Products form.
12. Verify the dataset appears as a product of the main sheep monitoring reference.
13. Activate the reference.
14. Enter the IRMA reference code into the Surveys table of the sheep monitoring database for the associated survey.

Your finished reference should be well documented, marked internal, active and sensitive and have a list of holdings that matches the list of data deliverables in the sheep monitoring protocol narrative (e.g., Figure SOP11-9, below)**Error! Reference source not found..** If done well the annual closeout procedure will be little more than a glance at this reference.

IRMA Portal
 Integrated Resource Management Applications

Home ▾ Search ▾ Add & Manage ▾ Topics ▾

Portal > Home > IRMA Applications > Data Store > Search > Profile

Add Digital File | Add Child | View XML | Clone | Add To Collection

Edit Details | Change Type | Manage Link | Inactivate | Quarantine | Activate

Rattenbury K. 2015. 2011 Itkillik Preserve Dall's Sheep Survey Dataset. National Park Service Arctic Inventory and Monitoring Program. Fairbanks, AK. Generic Dataset-2222131.

Generic Dataset

Reference Code: 2222131

Reference Status as of 05/19/2015: Active

Visibility: Internal

Core Information

Holdings

Digital
 DS08-SurveyDatasheets (pdf, 2mb)

Digital
 DS01: NPS.gdb (ESRI personal geodatabase, 10mb)

Information Resource Evaluation

Sensitivity Commercially Valuable - Park-specific, location-specific, and non-public information about a resource that is traded or sold in legal or illegal markets such that unauthorized removal of the resources from a project area has the potential of civil or criminal penalties.

Operations - Important to the internal business operations of the NPS

Figure SOP11-9. Example of a sheep survey dataset IRMA reference. The example is incomplete but shows how annual monitoring deliverables should be archived. Having everything listed in one place will simplify annual closeout and data certification procedures.

SOP 11.13.2 Link the Field Deliverables Reference as Product of the Master Sheep Monitoring Project Reference

The seasonal data deliverables Dataset references in the IRMA Data Store must be linked as a product of the master sheep monitoring Project reference in the IRMA Data Store.

Procedure:

1. Using a browser, navigate to the seasonal deliverables Dataset reference created in the section above.
2. Record the reference's reference code.
3. Navigate to the master sheep monitoring Project reference at <https://irma.nps.gov/App/Reference/Profile/2214983>.
4. Ensure you have editing privileges.

5. Scroll to the Products section.
6. Click 'Manage Products'
7. Click 'Create New Link'.
8. Search for the reference to attach using the reference type and reference code from above.

The screenshot shows the IRMA Portal interface. At the top, the header includes 'IRMA Portal' and 'Integrated Resource Management Applications'. Below the header is a navigation bar with links: Home, Search, Add & Manage, Topics, News, About, Help, and Contact Us. The main content area is titled 'Link This Project to an Existing Reference (i.e., Bundle Products)'. It features a search bar with a dropdown menu set to 'Geospatial Dataset' and a text input field containing '2224484'. A 'Search' button is next to the input field. Below the search bar is a table titled 'Select a Reference to Link'. The table has two columns: 'Code' and 'Display Citation'. The first row shows the code '2224484' and the citation 'Rattenbury KL. 2015. 2011 Wrangell-St. Elias National Park Dall's Sheep Survey Data Deliverables. National Park Service Central Alaska Network Inventory and Monitorin...'. A checkbox is checked next to the code. At the bottom right of the form, there are two buttons: 'Link Selected' and 'Done'. A red-bordered box at the bottom center contains the text 'Reference(s) linked successfully'.

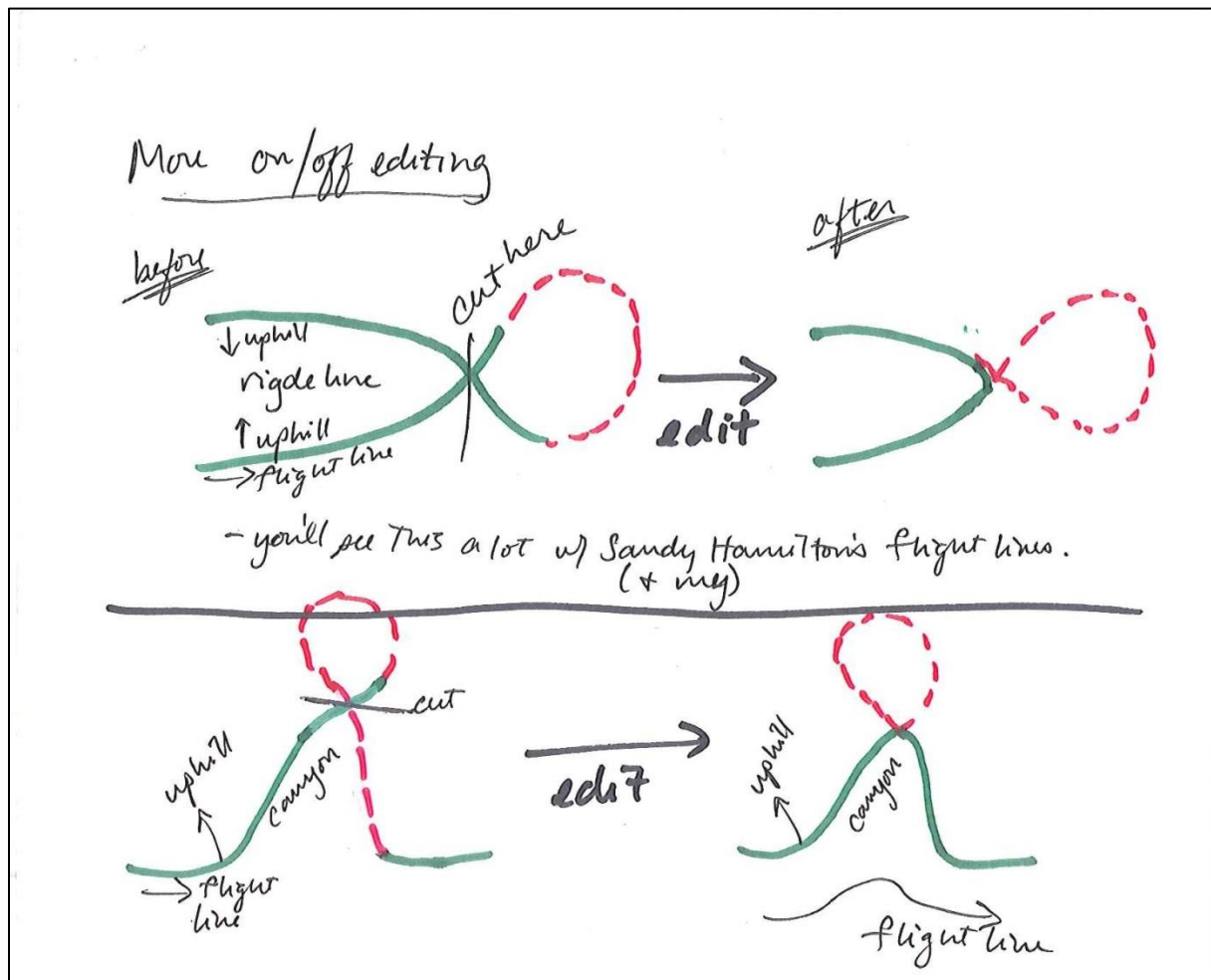
9. Check the correct reference and click Link Selected.
10. Click Done and back out of the remaining forms.
11. Check that the deliverables Dataset exists as a product of the master Project reference.
12. The deliverables dataset reference should not appear on the sheep monitoring web profile on the ARCN website because it was tagged 'Operations' and 'Sensitive'. Verify that it does not appear publicly at <http://science.nature.nps.gov/im/units/cakn/vitalsign.cfm?vsid=44>.

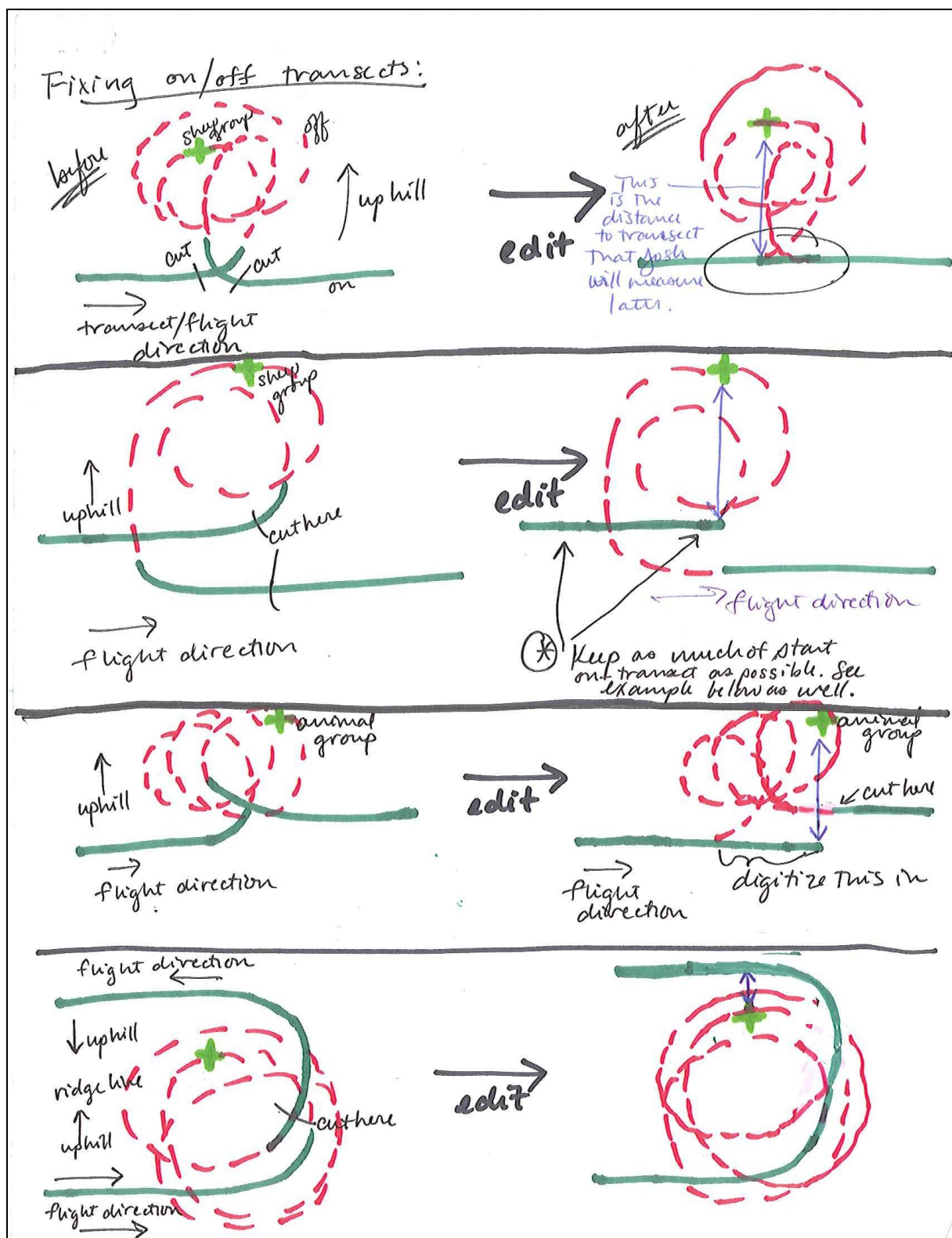
SOP 11.14 – Literature Cited

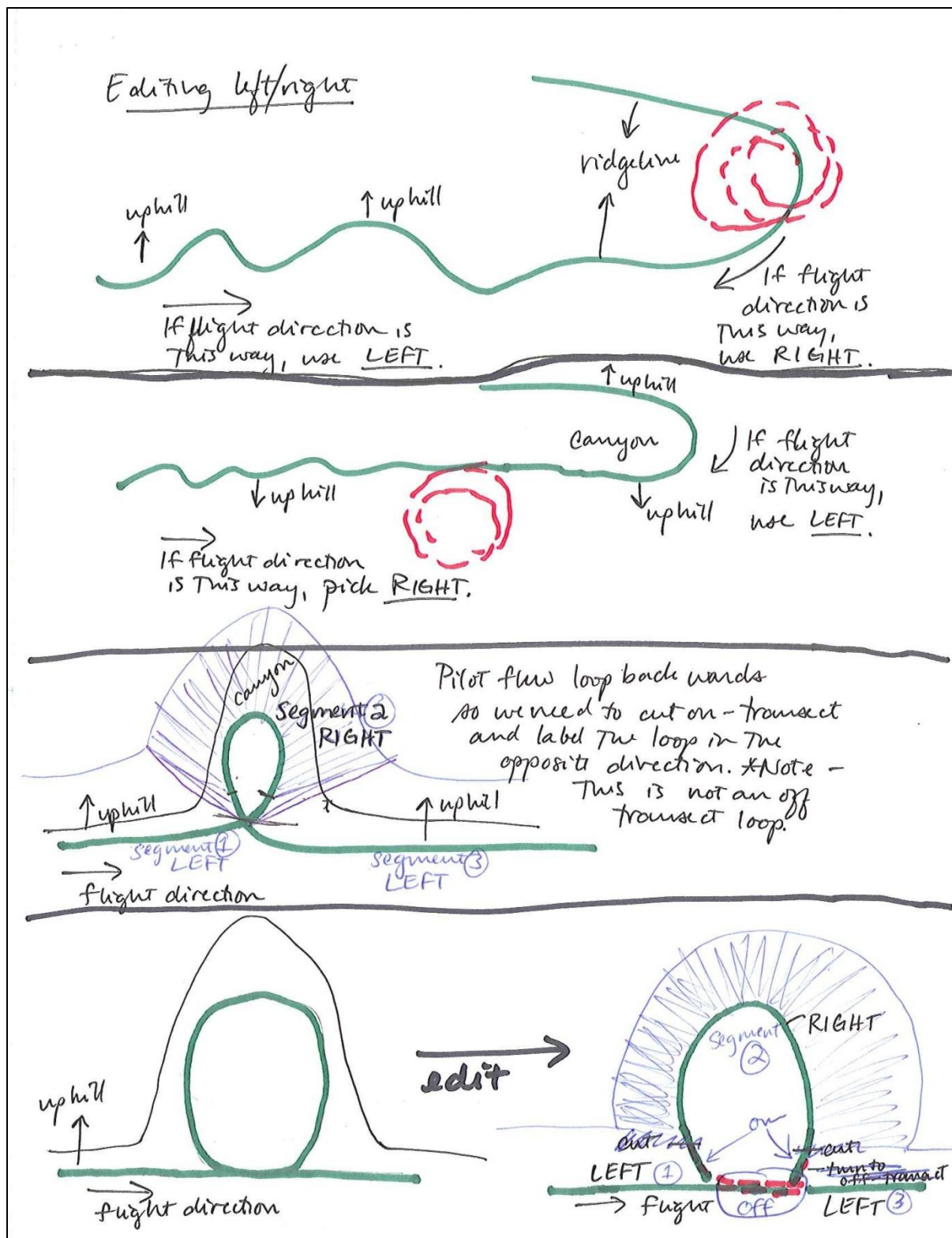
- Schmidt, J. H., and K. L. Rattenbury. 2013. Reducing effort while improving inference: Estimating Dall's sheep abundance and composition in small areas. *Journal of Wildlife Management* 77:1048-1058.
- Schmidt, J. H., K. L. Rattenbury, J. P. Lawler, and M. C. MacCluskie. 2012. Using distance sampling and hierarchical models to improve estimates of Dall's sheep abundance. *Journal of Wildlife Management* 76:317-327.
- Walsh, P., J. Reynolds, G. Collins, B. Russell, M. Winfree, and J. Denton. 2010. Application of a double-observer aerial line-transect method to estimate brown bear population density in southwestern Alaska. *Journal of Fish and Wildlife Management* 1:47-58.

SOP 11.15 – Appendix A: Transect Editing Illustrations

It is not always possible for observers to mark off-transect segments and sheep groups on the field computer in a timely (accurate) manner while on a survey flight due to a number of factors but mainly because their primary function is to detect groups of sheep versus focusing on the computer while on-transect. This Appendix shows edits that observers may complete each night during the survey, but that must be completed as part of the QA/QC process after the survey is completed and before data analysis. These hand-drawn illustrations supplement SOP 11.4.

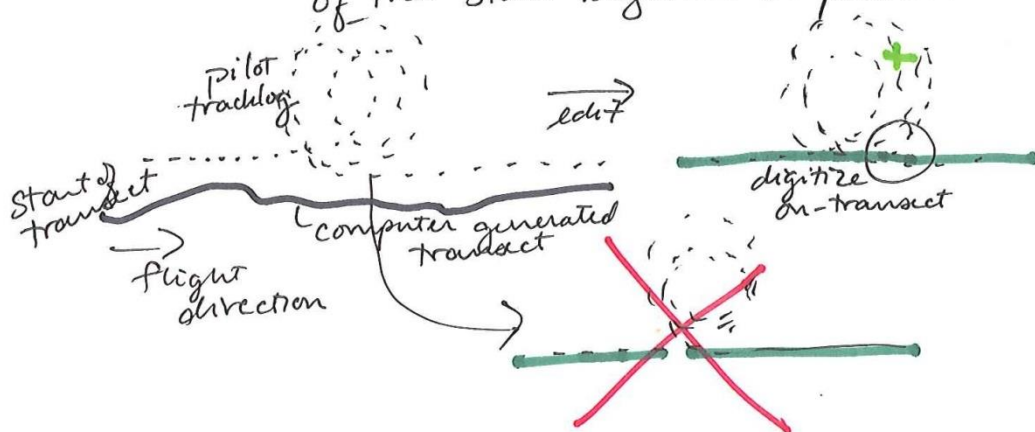




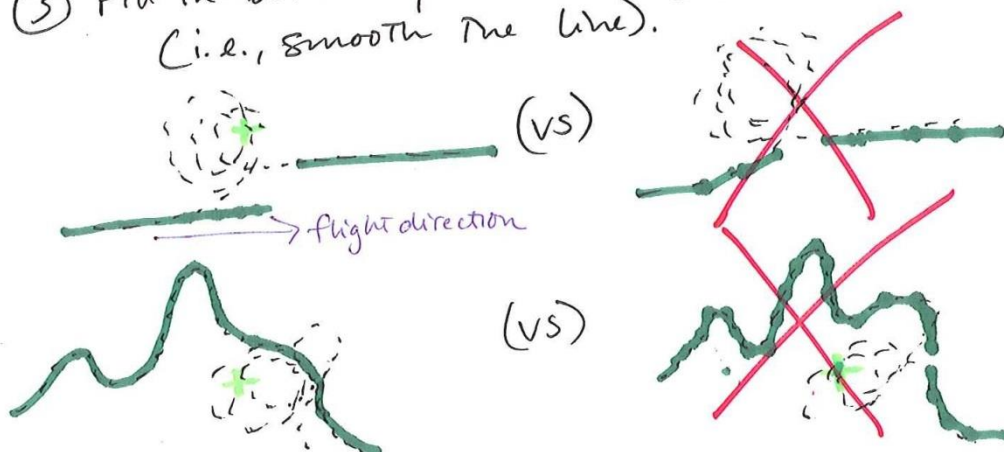


Digitizing on pilot's tracklog

- ① Only digitize on-transect segments.
 – don't worry about off-transect segments
- ② — pay attention to start + end of transect
 based on pilot's waypoints.
 ↪ for on/off cuts – preserve as much
 of the start segment as possible:



- ③ Fill in-between pilot tracklog points
 (i.e., smooth the line).



SOP 11.16 – Appendix B: Alternate Method to Generate Buffers

In 2015 there was a new problem with buffering that affected both the custom NPS Transect Tool and the generic ArcMap Buffer tool requiring additional edits to DB Tracklog segments as part of the buffer generation process. This may be a change in the upgrade from ArcMap 10.1/10.2 to 10.2.2/10.3. Where transects run around a ridgeline, the buffer is only produced for the area between the outgoing and return transects, and is not produced for the area outside of those lines which would be the full buffer for the opposite side of the ridgeline (see Figures SOP11-10a and SOP11-10b, below).

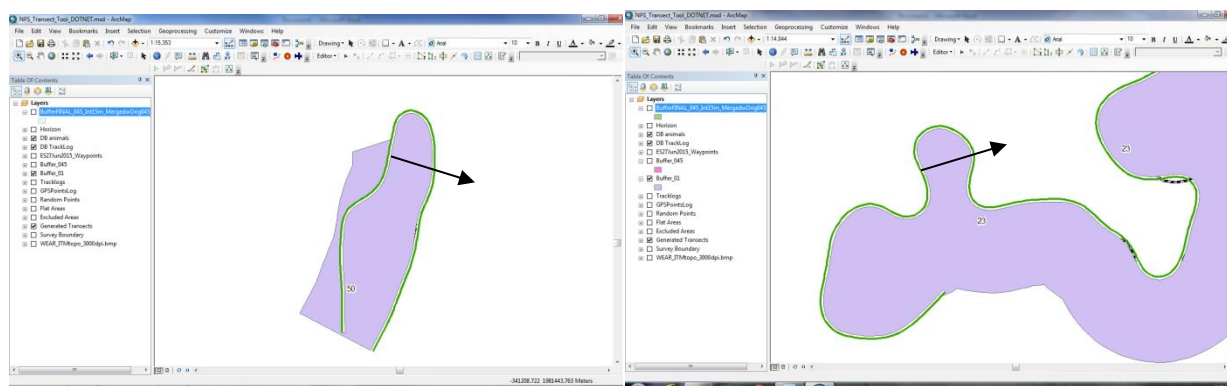


Figure SOP11-10a. Copy for Appendix B – Screenshot examples of incorrect ridgeline buffers produced with the Buffer tool in ArcMap 10.3. The buffer from the left side of these ridgelines should overlap the right-side flight line (green DB Tracklog line) to cover the full effective strip width, indicated by the black arrows. The buffer tool in later versions of ArcMap erases this overlap during processing, resulting in underestimated buffer areas. Compare with Figure SOP11-10b.

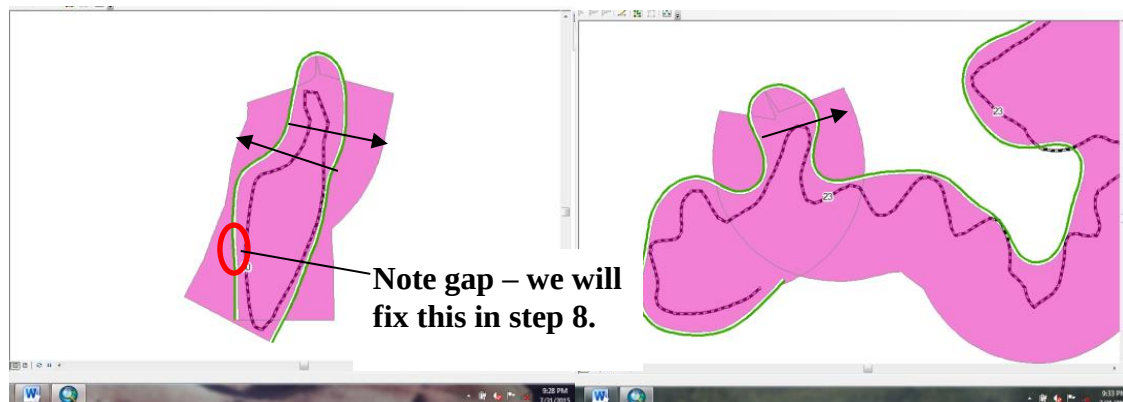
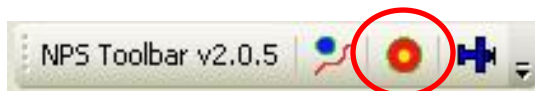


Figure SOP11-10b. Copy for Appendix B – Screenshot examples of corrected ridgeline buffers produced after splitting DB Tracklog and re-running the Buffer tool in ArcMap 10.3. The black arrows indicate where the corrected buffers cover the full effective strip width compared with Figure SOP11-10a.

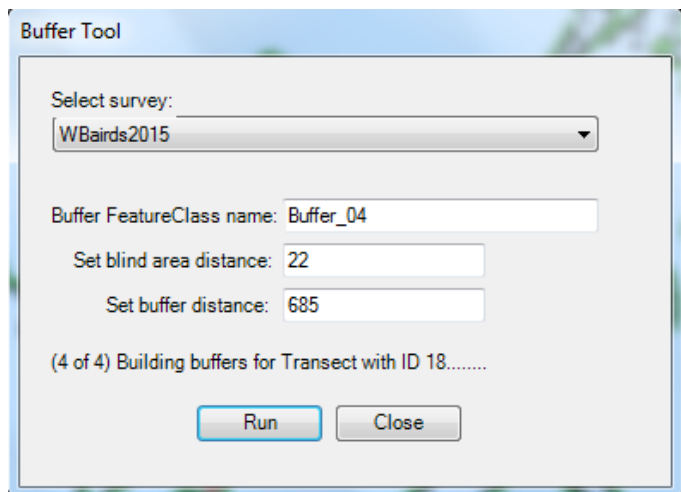
If your first buffer tool attempt (see steps 1 and 2 below) produces buffers like Figure SOP11-10a (incorrect, underestimated buffer areas), follow the instructions that follow here to make buffers. Then go to SOP 11.7. If your first buffer tool attempt looks like Figure SOP11-10b, you can go back to SOP 11.6, step 3.

- 1) Open the FINAL ArcMap project where you combined all geodatabase information from all observers and aircraft for the survey.

- 2) Run the Buffer tool for the first time so that you can check that all buffers are on the uphill side of each transect. Click the Buffer (target-shaped) button on the NPS Toolbar v2.0.5.



In the Buffer Tool window that opens, pick your Survey ID from the drop-down list, name your Buffer feature class, set the distances at 22 and 685 and hit Run.



You may get an error message that the Buffer table doesn't exist. Close the Buffer Tool window, Save and Close your ArcMap project. Reopen your ArcMap project and select the Buffer tool button again. This time your Buffer feature class should be there for you to select. Once it's running you'll see it processing and making buffers in the Buffer Tool window. This may take 5-30 minutes to run. If the Buffer tool doesn't work, see SOP 11.6 steps 5-6 and generate one buffer set (shapefile) each for the left and right on-transect segments from DB Tracklog.

- 3) Review the first buffer set(s) with a topo layer background to make sure all buffers are on the uphill side of their respective transects. For those that are on the downhill side, you will need to change the left or right designation of the associated on-transect segments in the DB Tracklog feature class. Start an Editing session and make sure DB Tracklog is set as the target dataset for editing. Open the Attributes table for DB Tracklog and make the appropriate changes Left or Right in both the PilotDir and Obs1Dir fields. The designation should always indicate the uphill side for the direction that the on-transect segment was flown.
- 4) In an Editing session, split DB Tracklog segments where they go around a ridgeline (Figure SOP11-11, below).

- 5) Close the Editing session.
- 6) Rerun the Buffer tool or follow steps 5-9 from SOP 11.6 if the Buffer tool doesn't work. This time after splitting DB Tracklog at ridgelines, the Buffer tool will create full buffers for each side of the ridgeline (Figure SOP11-10b). This is your "Base Buffer" shapefile.
- 7) Export the base buffer as a new, "Integrated Buffer" shapefile (e.g., Itkillik2012_Buffer_Integrate15m.shp).
- 8) Run the Integrate (Data Management) Tool (you can find this with the Search window in ArcMap) on the Integrated Buffer shapefile you created in step 7. Set the XY Tolerance to 15 m. This will close the gaps shown in Figure SOP11-10b.
- 9) Merge the Integrated Buffer with the Base Buffer. This will create a new, "Merged Buffer" shapefile (e.g., Itkillik2012_Buffer_Int15m_Merged.shp). I like to name each successive shapefile with all of the preceding steps used to make it.
- 10) Run the Dissolve (Data Management) Tool on the Merged Buffer using TransectID as your Dissolve Field. The result will be a new shapefile with one buffer per transect (e.g., Itkillik2012_Buffer_Int15m_Merged_Dissolved.shp).
- 11) Look through your buffers. The Integrate-Merge-Dissolve process is not perfect and you may find you need to do a little vertex editing of your buffers. See Figure SOP11-12, below.

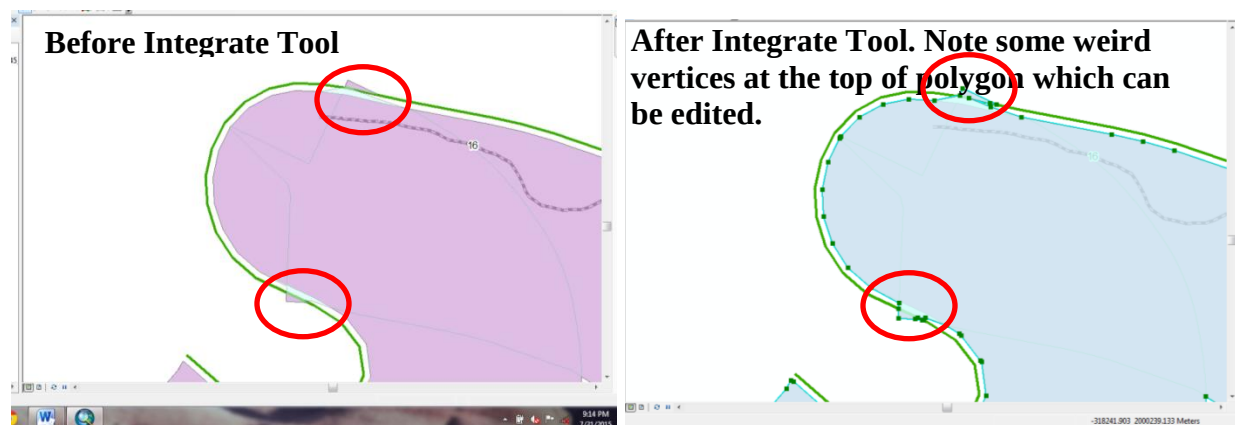


Figure SOP11-12. Screenshots showing where additional buffer editing may be necessary. Note the unusual vertices at the top of the polygon in the After Integrate Tool screenshot. This may be extreme in some places and should be edited but in this example, the area is small so making the edits won't change the resulting area very significantly and is not a critical edit.

- 12) Look for overlapping buffers from different transects. Check the date and elevation. If they were flown on the same day and the difference in transect elevation is < 900', then you will have to erase that portion of the buffer from the transect that was flown second where it overlaps the buffer of the transect flown first. To do this, add a new field into the buffer attributes table called ClipOrder. Then open an Editing session and in the ClipOrder field label the first transect "1" and the second "2". Remember you only have to do this for transects flown the same day that are < 900' difference in elevation. Save your edits. Select and export the #2 buffers to a new shapefile and add that to your ArcMap project. Delete the ClipOrder #2 buffers from the main buffer shapefile. Close the Editing Session. Use the Erase (Analysis) tool on the #2 buffers shapefile with the main buffer shapefile as your erase feature. The result will be a new, adjusted #2 buffers shapefile where those overlaps have been removed. Use the Append tool to add the adjusted #2 buffers back into the main buffer shapefile.
- 13) If any of the buffer areas extend outside of the Survey Boundary, use the ArcMap Clip (Analysis) tool to remove those areas from the buffer shapefile. This procedure will create a new, clipped buffer shapefile.
- 14) Calculate buffer area (m²) for the clipped buffer shapefile. In ArcMap 10, you can use the Calculate Areas (Spatial Statistics) tool, which will produce another shapefile. In ArcMap 10.3, just add a new field to the attributes table called Area_m2, right click on the field and select Calculate Geography to calculate and input area for each buffer.
- 15) Export the final buffer shapefile as a Feature Class to your final survey geodatabase (where the final DB Tracklog and DB Animals are housed) and call it Buffers_Final. See Figure SOP11-13, below.

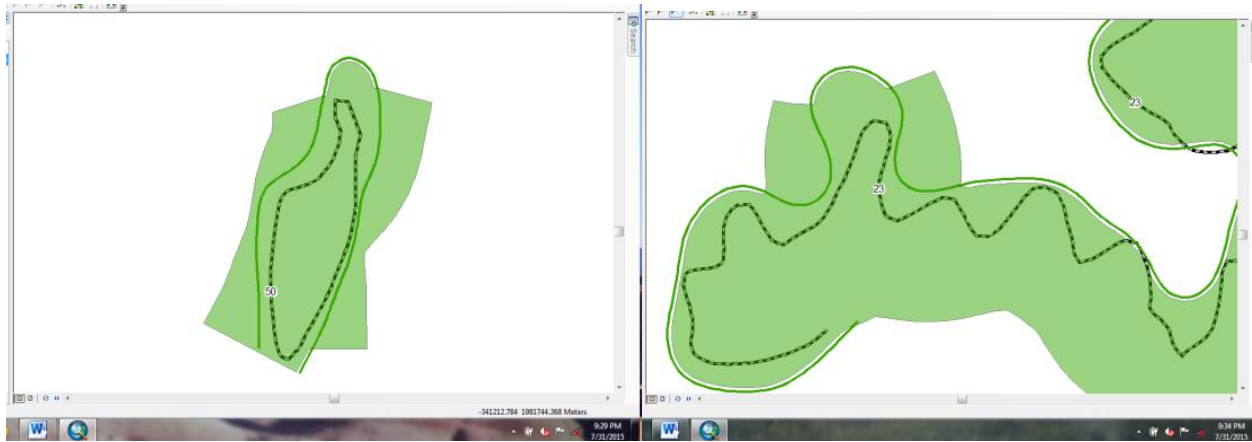


Figure SOP11-13. Final buffers after checking left/right designation and overlapping transects; splitting tracklogs at ridgelines; running the Buffer, Integrate, Merge and Dissolve tools; Erasing the overlap areas; and Clipping the Buffers to the Survey Boundary.

SOP 11.17 – Appendix C: Dataset Certification Form

Data certification is a benchmark in the project data life cycle that indicates that:

- The monitoring data are complete, up to date and ready for analysis and publication
- The data have passed quality assurance checks
- The data are appropriately documented
- All aspects of the monitoring program that generated the data, including the protocol, standard operating procedures and data management systems are in order, documented and up to date.

Certification is not intended to imply that the data are completely free of errors or inconsistencies which may or may not have been detected during quality assurance reviews.

This certification form must be uploaded to the monitoring dataset's IRMA reference in scanned pdf format.

Certification date:

Certified by:

Title:

Affiliation:

Project title:

Range of dates for certified data:

Description and scope of data being certified:

____ This certification refers to data in accompanying files. Check all that apply, and indicate file names to the right:

____ Database file(s): _____

____ Spatial data file (s): _____

____ Geodatabase file(s): _____

_____ Other (specify): _____

Is there any sensitive information in the certified data which may put resources at greater risk if released to the public (e.g., spotted owl nest sites, cave locations, rare plant locations)?

_____ No _____ Yes

Details: _____

Description of data processing, quality assurance measures, and results and summary of quality assurance reviews, including details on steps taken to rectify problems encountered during data processing and quality reviews.