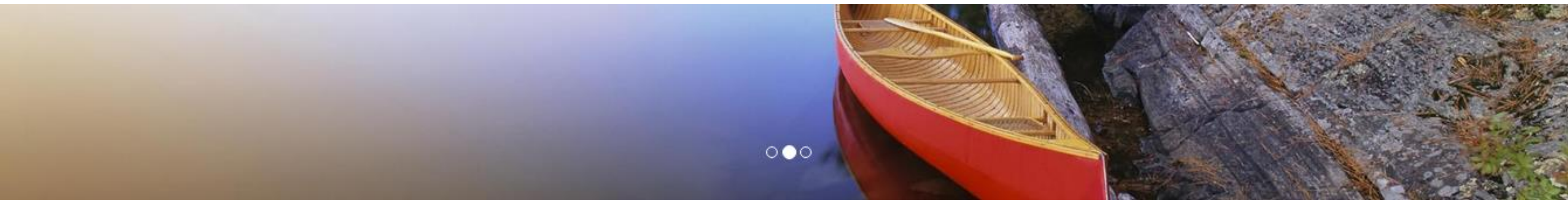
A Data Submission Framework for the National River Recreation Database

A Capstone Project in Fulfillment of the Masters in GIS Degree Program
University of Washington

Katie Messick and John Ritzman
August 21, 2015

Sponsored by Susan Rosebrough, National Park Service, and
Risa Shimoda, River Management Society



1. Recommended Course of Action

Work with the transformed test donor data to check its compatibility with the National River Recreation Database (NRRD). Integrate the test donor data into the NRRD to prepare for public release. Implement tasks labeled “urgent” in Section 5.4 (Next Steps). Distribute Cookbook and toolkit as zip file on ArcGIS Online for prospective data donors to download to allow them to consistently translate their data into the format of the NRRD for integration. Support donors with training in the toolkit and HEM tool. Add to and update toolkit and Cookbook as needed or following any changes to the NRRD schema.

Table of Contents

List of Figures and Tables

2. Introduction

The National River Recreation Database (NRRD) is a cooperative venture between the National Park Service and the River Management Society, in partnership with other organizations including the Bureau of Land Management, United States Forest Service, American Whitewater, ESRI, and the United States Geological Survey. Its primary purpose is to "provide river recreation and management information useful to resource managers and interested publics." (Rosebrough, 2015). The NRRD is currently in its infancy, with a well-established framework and a few test data sources from the National Park Service and other partner organizations. The vision is to create a robust, nationwide online resource where people can "shop" for river recreation experiences and learn about river safety and conservation, with the aim of increasing the awareness of and protection for Wild and Scenic Rivers, whitewater rivers, and water trails.

Another goal of the NRRD is to increase ecological awareness and protection for rivers. "As recreation users become enthused about [river recreation] experiences, we want to educate them on the value of river protection and river ethics to create a community of stewardship and sustainable river use." (Rosebrough, 2015). Recreation can degrade the environment through direct interference with plants, animals and terrain (Liddle, 1980) or indirectly through pollution (Forster, 2005). However, through increased visibility and local and regional economic influence, recreationists also contribute significantly to the conservation of natural areas (Burger, 2000). By bringing local and regional river recreation opportunities to the attention of a national audience, the NRRD hopes to leverage some of those ecotourism dollars to increase river conservation. Table 1 summarizes the social-ecological system that the NRRD represents.

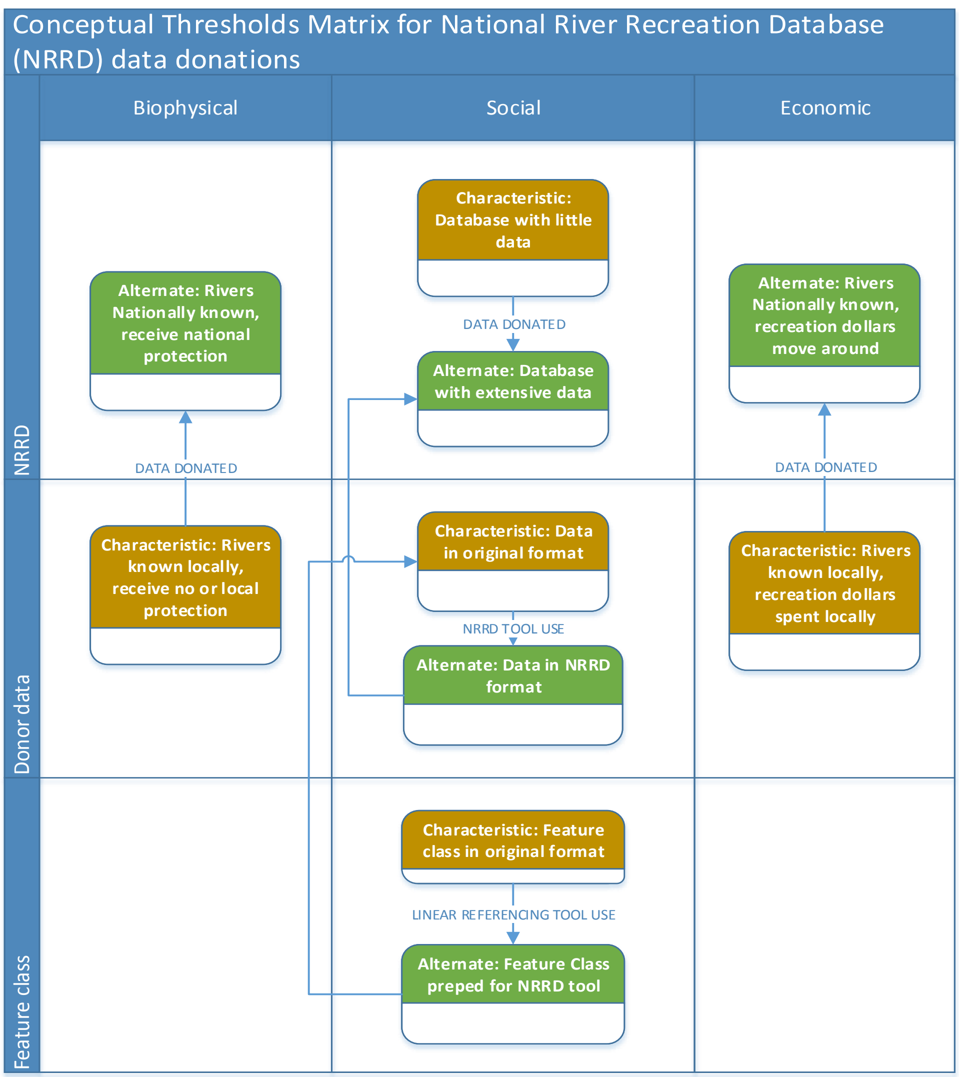
Table 1: Social Ecological Systems Table

Social-ecological Systems Table for National River Recreation Database (NRRD) data donations			
	Biophysical	Social	Economic
NRRD: Wild and Scenic Rivers, water trails and whitewater rivers Nationally	Nationwide the NRRD has the potential to increase protection for rivers with recreational value by winning advocates for conservation from outside of the local or regional area.	One of the NRRD's goals is to "educate [users] on the value of river protection and river ethics to create a community of stewardship and sustainable river use." Increased awareness translates to increased conservation.	The NRRD has the potential to increase the visibility of river recreation locations nationwide, thus channeling ecotourism dollars to areas that might not see them without this exposure. There are costs associated with creating and maintaining the NRRD.
Donor data: Local and Regional water trails, whitewater rivers and Wild and Scenic Rivers	Rivers protected for recreational purposes often provide ecological benefits, including habitat conservation and water quality protection. On the other hand, recreation can contribute to environmental degradation, directly impacting plants, animals and terrain.	One of the challenges of the NRRD project will be attracting willing data donors. The prototyped NRRD toolkits presented here should lower some potential barriers to data donation by making the process simpler.	River recreation at a local or regional scale provides economic support to local communities, river trip outfitters, and other business involved in ecotourism. Local organizations donating data to the HEM will incur costs as the investment of time needed to learn and use the tools successfully.
Feature class: Local and regional river recreation features (e.g. campgrounds)	Riverside campgrounds, boat ramps other access points can be a source of pollution, such as sediment, sewage and hydrocarbons, as well as other forms of environmental degradation. Adequate, well maintained facilities minimize these impacts.	Education about stewardship and river etiquette can lead to better, cleaner recreation facilities and experiences. Good information about and maps of point data such as river rapids and emergency access points can increase safety.	Fees at campgrounds and boat launches help maintain facilities.

Currently there is only one national organization that dispenses river recreation information. American Whitewater is a membership organization with a mission "to conserve and restore America's whitewater resources and to enhance opportunities to enjoy them safely." (American Whitewater, 2015). The website has detailed resources on current conditions on hundreds of whitewater rivers. However, there is no comparable resource for other river recreation experiences, including water trails and Wild and Scenic Rivers. The NRRD will fill this void. In addition, the NRRD will maintain practical information, such as locations of an amenities at access sites and campgrounds.

The NRRD framework consists of an ESRI geodatabase with multiple feature classes representing point and line features such as access sites and river reaches. For the sake of consistency across all rivers, these data are referenced to the National Hydrography Dataset (NHD) (USGS, 2015), which is the standard linear referenced dataset for water courses on a national level. Therefore, all data imported into the NRRD must also be referenced to the NHD. In addition, in order to integrate data from outside sources into the NRRD, those data must be put into the same schema as the NRRD. Since the NRRD will rely on data donors throughout the country to provide their detailed local knowledge of river recreation facilities and opportunities, there must be a standardized way to transform donor data into the format required by the NRRD. Data integration of this type is fraught with potential for errors and mismatches (Flowerdew, 1991). This project aims to provide both standardization and quality control. A toolkit was created consisting of several toolboxes, one to pre-process donor data and provide linear referencing capabilities, and one each for river reaches, access sites and campgrounds, to translate that processed data into the NRRD format. In addition, a "cookbook," or detailed instruction manual, was written to guide data donors through the entire process. In the end, the process is as simple and clear as possible given the complexity of the tasks. Table 2 shows a conceptual thresholds matrix indicating characteristic and alternate (desirable) states.

Table 2: Thresholds Matrix



3. Design and Methods

The toolkit and "cookbook" instruction manual for submitting data to the National River Recreation Database (NRRD) were developed as part of an eight-week capstone project for the University of Washington Professional Masters in GIS program. A proposal was submitted to the project sponsors outlining the steps needed to develop these tools. These steps included:

1. Translate sample donor data into the NRRD format;
2. Create models of the workflow to automate task 1;
3. Write tools based on the models; and
4. Upload tools to ArcGIS online.

The actual workflow was much more fluid, with models and scripts created simultaneously, and step 1 not completed until after all the tools had been drafted. In the end, step 4 was beyond the scope of an eight-week project as well. The sample donor datasets used included a small dataset (six river reaches) from the Bridger-Teton National Forest and a large dataset (156 river reaches) from the Minnesota Department of Natural Resources.

3.1 General Approach

Tools were developed primarily using ESRI's ArcGIS for Desktop application. Donor data and the most recent version of the NRRD were provided as ArcGIS file geodatabases. In addition, the USGS Hydrography Event Management (HEM) tool for linear referencing is an ArcGIS add-in. These dependencies led the project team to choose an "ESRI solution" for the toolkit as opposed to an open-source solution. This decision is covered in more detail in the Discussion section.

ArcGIS provides a robust framework for creating custom geoprocessing tools that can use and extend basic operations such as appending features to new feature classes, adding and deleting attribute fields, and performing more complex spatial analysis. These operations can be chained to perform very specific actions on geospatial datasets. This provided a natural environment for developing a suite of tools to convert geographic features to the NRRD. Two primary environments are available in ArcGIS to develop custom tools, each with its own advantages and disadvantages. The ModelBuilder environment allows the developer to add ArcToolbox functions to a graphical user interface and then connect geographic datasets and other tools to create a full workflow. ModelBuilder tools tend to be more readable and easier to build and edit, but do not provide much functionality outside of chaining pre-existing ArcGIS tools. On the other hand, the Python environment enables the developer to write the execution code of the tool in the Python programming language. ArcToolbox tools can still be called, but manual code must be written to chain these operations. Python script tools tend to benefit from being able to perform operations outside of just ArcGIS tools, but suffer from less automatic error checking compared to ModelBuilder. Figure 1 and Figure 2 show an example of tool development in each environment. Tools of both types can coexist in a single ArcGIS toolbox, and the tools can even call each other.

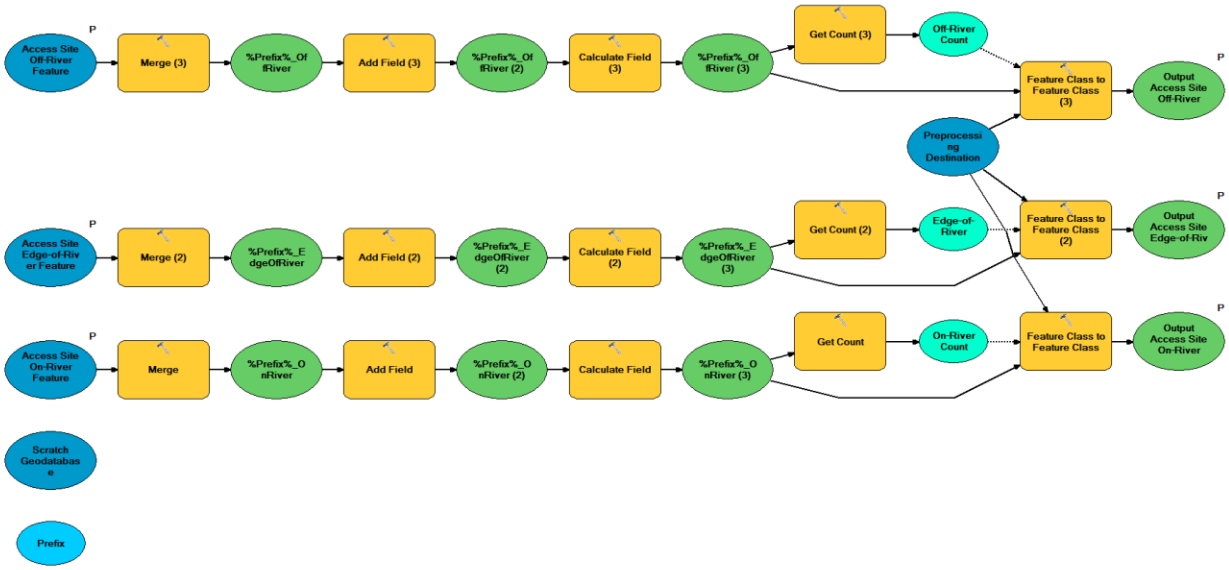
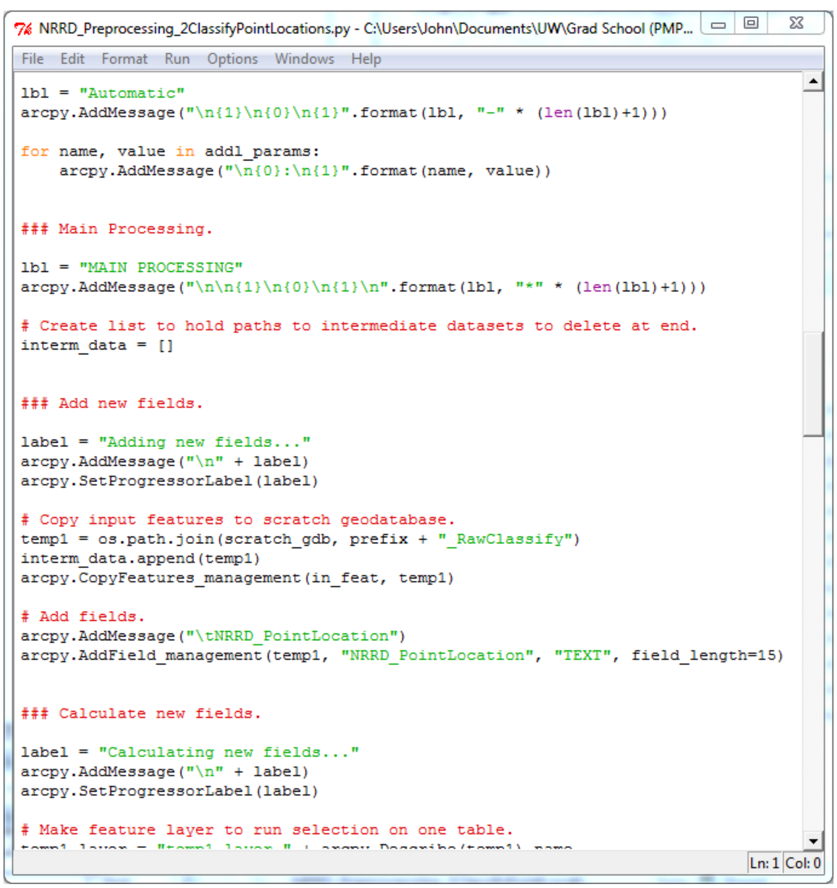


Figure 1: Development in ArcGIS ModelBuilder involved chaining ArcToolbox tools together in a graphical user interface to create a workflow. This was used for tools where ArcToolbox functions alone could perform the desired operations.




```
lbl = "Automatic"
arcpy.AddMessage("\n{1}\n{0}\n{1}".format(lbl, "-" * (len(lbl)+1)))

for name, value in addl_params:
    arcpy.AddMessage("\n{0}:\n{1}".format(name, value))


### Main Processing.

lbl = "MAIN PROCESSING"
arcpy.AddMessage("\n\n{1}\n{0}\n{1}\n".format(lbl, "*" * (len(lbl)+1)))

# Create list to hold paths to intermediate datasets to delete at end.
interm_data = []


### Add new fields.

label = "Adding new fields..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Copy input features to scratch geodatabase.
temp1 = os.path.join(scratch_gdb, prefix + "_RawClassify")
interm_data.append(temp1)
arcpy.CopyFeatures_management(in_feat, temp1)

# Add fields.
arcpy.AddMessage("\tNRRD_PointLocation")
arcpy.AddField_management(temp1, "NRRD_PointLocation", "TEXT", field_length=15)


### Calculate new fields.

label = "Calculating new fields..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Make feature layer to run selection on one table.
```

Figure 2: Development in Python involved manually coding the operations of the tool. This was used for longer tools or tools where additional functions outside of ArcToolbox were required.

The advantages and disadvantages of each environment heavily influenced where each was used. ModelBuilder was always attempted first, and development was switched to Python if needed. ModelBuilder tended to perform best on tools with a mostly linear workflow and where complex spatial analysis functions were not required. This was common with the post-processing tools for populating NRRD attribute fields, where ModelBuilder enabled the rapid prototyping of roughly 20 tools from a common template. After the first tool was made, it could be copied and tweaked to work on a new attribute field. Python was required whenever the workflow became convoluted or when more complex operations were required. This was common with the preprocessing tools, which were designed to address specific issues with donor data that were causing difficulties later in the workflow. In these cases, Python enabled a wider set of operations to be performed in order to accomplish these tools' goals. The final NRRD toolkit thus contains a mix of ModelBuilder and Python tools.

In addition to the two development environments, ArcGIS applications such as ArcMap and ArcCatalog were used to perform a variety of other tasks to investigate donor data and test out the workflow. A custom ArcGIS add-in called the Hydrography Event Management (HEM) tool from the USGS was also used during the linear referencing phase of the workflow to match donor data with rivers in the National Hydrography Dataset (NHD) (USGS, 2015). Several other programs outside of geographic information systems were also critical in this project. Google Drive was used extensively to collaborate. Google Drive enabled the transfer of large GIS datasets and geoprocessing tools as well as editing shared documents. Microsoft Visio was used to develop some of the workflow diagrams and tables in this report. Skype, Join.me, and GoToWebinar were used during some of the conference calls with project sponsors to share screens, especially during the HEM tool training. Finally, Microsoft Word, Excel, and PowerPoint were used to create the cookbook and some of the other deliverables for this project.

3.2 Preprocessing Tools

The preprocessing tools were not originally part of the workflow. Initially, it was imagined that data donors would first run their data through the USGS Hydrography Event Management (HEM) tool to reference them to the rivers of the National Hydrography Dataset (NHD) and then use the custom tools created in this project only afterwards to populate the NRRD-specific attribute fields. However, issues with HEM output and different situations of donor data convinced both the project team and sponsors that a set of preprocessing tools were needed before working with the HEM was advised. These tools would be placed in a separate toolbox and designed around the theme of addressing particular situations that, if left unresolved, would cause significant hang-ups later in the workflow. As such, these tools were less planned than the post-processing tools and were developed reactively as these situations were identified.

Coordinate system mismatch is a ubiquitous issue when combining datasets from different sources. With linear referencing to the NHD (now a middle step between pre- and post-processing) being a complex operation to begin with, this became an issue important to address early in the process. The NHD uses the un-projected North American Datum of 1983, which would be difficult to change due to the number of feature classes that are part of its geometric network, used by the HEM tool. Projection to NAD83 thus became the very first step of the workflow to address issues of coordinate systems as early as possible. This was not an issue with either sample dataset as both were in a NAD83-based coordinate system already. Most datasets in the United States will likely be in a NAD83-based coordinate system or one of its derivatives (e.g. NAD83_HARN) and so will have little trouble with this step. The built-in ArcToolbox tool Batch Project perfectly encapsulates the functionality of projecting multiple feature classes to a single location, so a custom tool was not needed.

However, a structure for preprocessing quickly became important to develop as well. As donor feature classes may not be in the NRRD types of river reach, access site, and campground yet, a simple "stages" idea – using a type/number prefix (e.g. AccessSite_1_StateFieldAdded) to organize intermediate datasets within a root file geodatabase – like with the post-processing tools was not feasible. Instead, several feature datasets (all in NAD83) were set up within a preprocessing geodatabase to organize datasets at different stages. Figure 3 shows this geodatabase in Catalog.

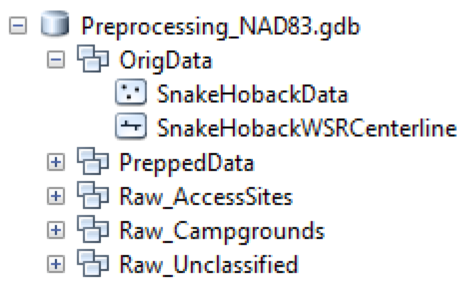


Figure 3: The schema of the NRRD preprocessing geodatabase. Because donor datasets are not already separated and grouped into access sites, campgrounds, and river reaches at this point, a deliberate organization had to be applied to help keep track of intermediate datasets. The OrigData feature dataset here shows the two Bridger-Teton National Forest feature classes.

A feature dataset called *OrigData* holds the initial projected feature classes, identical to the donor data other than coordinate system. *PreppedData* will hold the datasets ready for the second phase of the workflow, linear referencing. The 3 *Raw_* feature datasets hold intermediate datasets for the different point types, with *Unclassified* being for catching points that did not fit into access sites or campgrounds and thus would not be taken further in the workflow (for reference for donors). River reaches did not need an intermediate feature dataset they have only one step in preprocessing in the current workflow, and thus are able to be put directly into *PreppedData*. This structure, along with the cookbook walkthrough, was designed to best help donors keep track of where they are and what they are doing in the preprocessing workflow. Additionally important was a separate scratch geodatabase where the intermediate datasets created *during execution* of each individual tool were written, which were deleted after successful execution.

The custom preprocessing tools themselves actually span from before, to during and after the linear referencing step. This is because, since linear referencing is applied differently to different types of features, several tools were required *after* linear referencing to reunite these feature classes. Because there were then tools before and after linear referencing, the two custom "HEM alternative" tools (described in the following section), were also added to the preprocessing toolbox. The complete set of custom preprocessing tools that resulted from this is shown in Figure 4.

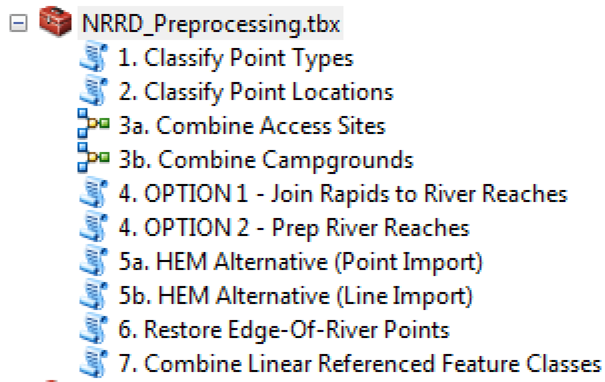


Figure 4: The total set of preprocessing tools delivered with this project. The tools actually cover all the way through linear referencing and afterwards to prepare datasets for post-processing.

The tools are organized as follows. The first 3 tools prepare point data (access sites and campgrounds) for linear referencing. Tool #4 prepares river reaches for linear referencing. Tool #5 provides an alternative method of linear referencing to the HEM for those donors unable or uncomfortable in using it. This tool is described more completely in the next section. Tools #6-7 reorganize feature classes after linear referencing (regardless of method used) to prepare them for the post-processing tools. Most of the tools ended up having to be written in Python rather than left as models due to two small issues ModelBuilder struggled to handle: iterating over a certain set of attribute fields and creating dynamically named, but fixed location, output feature classes. The first issue was important due to handling original donor attribute fields differently than extra attribute fields being added by these tools. The second issue was critical in keeping donor feature classes separate while still keeping the organizational schema of the preprocessing geodatabase.

3.2.1 Classify Point Types

The **Classify Point Types** tool provides SQL queries for data donors to select which features in their point feature class(es) are access sites and which are campgrounds. Figure 5 shows its tool dialog box.

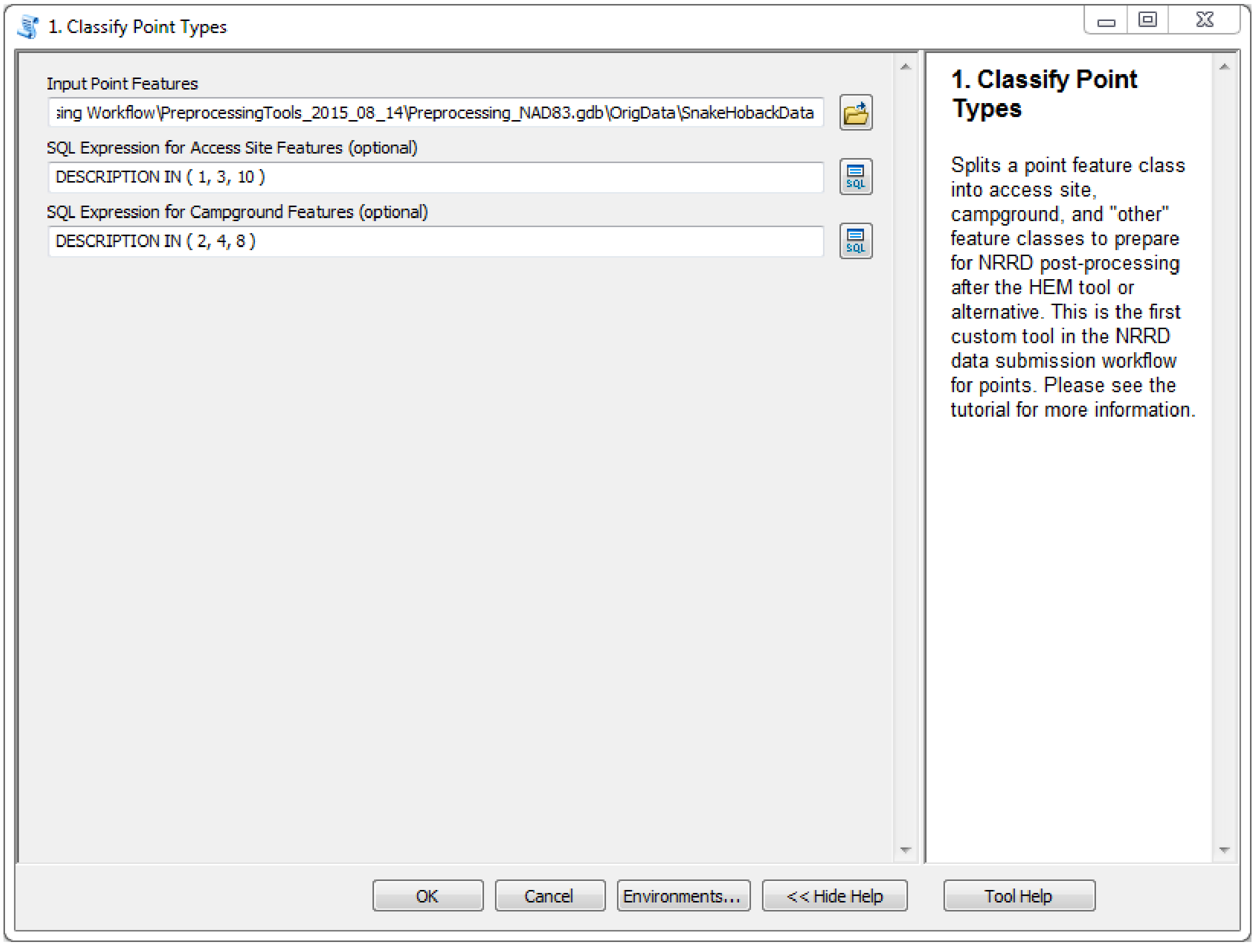


Figure 5: The tool dialog box for the first preprocessing tool. SQL queries enable the donor to split their point feature classes into access sites and campgrounds.

The general format of using SQL queries was nearly ubiquitous throughout the tools created in this project, both pre- and post-processing. This parameter type is dynamic as it allows the donor to classify features based on current attribute fields, and then the tool performs select logic on each class, such as categorizing one as access sites, another as campgrounds, and a third as unclassified. This additional information from each tool was usually captured in an additional attribute field added to the output. In the case of the Classify Point Types tool, this was a new field called “NRRD_PointType”, which would be calculated to “Access Site”, “Campground”, or NULL depending on which SQL query the feature fell into. The feature class itself was then split based on this new field into three separate feature classes, each put into their respective feature datasets in the preprocessing geodatabase. Figure 6 shows the updated preprocessing geodatabase for a donor that has run Classify Point Types.

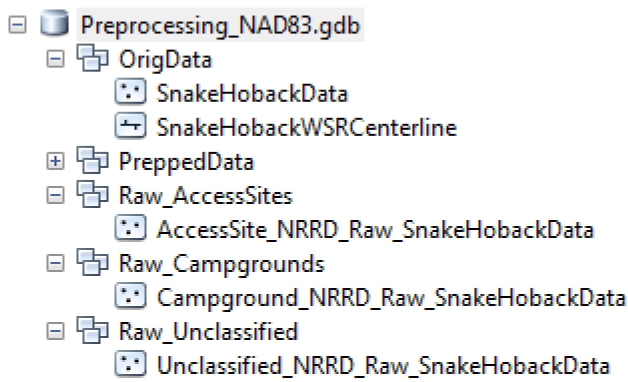


Figure 6: The NRRD preprocessing geodatabase after running the first custom tool, Classify Point Types. Multiple point feature classes can be classified into the two NRRD point types this way and will be combined later to produce complete access site and campground feature classes. Points left unclassified by the donor's SQL queries are kept for their reference.

If the donor has multiple point feature classes (e.g. boat launches and emergency access points), they can each be run through the Classify Point Types tool to be organized based on their NRRD type. The name of each of the output datasets is suffixed with the original feature class name to help keep track of where datasets are coming from. Eventually, all access site feature classes will be combined, as will all campground feature classes. Unclassified points are kept only for reference and quality control in case the donor wishes to see which features did not make it into the NRRD based on their SQL queries. If another NRRD point type is added in the future (e.g. bridges), an extra SQL parameter and output feature dataset is all that would be needed (along with tweaks to the code) to accommodate that new NRRD point type.

Another function that the Classify Point Types tool performs is prefixing original attribute fields with "DONOR_". Figure 7 shows this in the attribute table of one of the outputs of this tool.

DONOR_NAME	DONOR_DESCRIPTION	DONOR_RIVER	DONOR_Notes	DONOR_Public_Data	DONOR_Pic_Name	NRRD_OrigFC	NRRD_PointType
Taco Hole River Access Point	10	Snake	<Null>	YES	<Null>	SnakeHobackData	Access Site
Lunch Counter River Access Point	10	Snake	<Null>	YES	<Null>	SnakeHobackData	Access Site
West Table Boat Ramp	1	Snake	<Null>	YES	<Null>	SnakeHobackData	Access Site
Kahuna River Access Point	10	Snake	<Null>	YES	<Null>	SnakeHobackData	Access Site

Figure 7: The attribute table of Bridger-Teton National Forest's point data after having been run through the Classify Point Types tool. As much as possible, all NRRD tools attempt to help donors keep track of where data comes from. This tool addresses this by prefixing all original fields with "DONOR_" and new fields with "NRRD_".

This functionality was deemed necessary after field name conflicts were encountered in the later post-processing tools. Adding a NRRD field called "State" would fail if a field called "State" was already present in the feature class. Instead of building functionality to handle this into each subsequent tool, it was decided that it would be simpler to prefix original attribute fields with "DONOR_" at the start of the workflow. This would also help data donors keep track of which fields were theirs originally and which were being added throughout the workflow. As only NRRD fields (the ones being added throughout the workflow) would be kept in the final submission datasets, a "DONOR_" prefix would also help indicate which fields would be dropped at the end of the workflow. Rather than building a separate tool to perform this prefixing, this function was instead just rolled into the first tool for points (#1) and for lines (#4). A final additional field, "NRRD_OrigFC" was also added in these tools to keep track of the origin of features. This can be used to help donors identify where individual features came from even much later in the workflow, aiding QA/QC.

One final aspect of the first tool that is also followed by all preprocessing tools is using ArcGIS tool execution messages to help with QA/QC. Figure 8 shows an example of this.

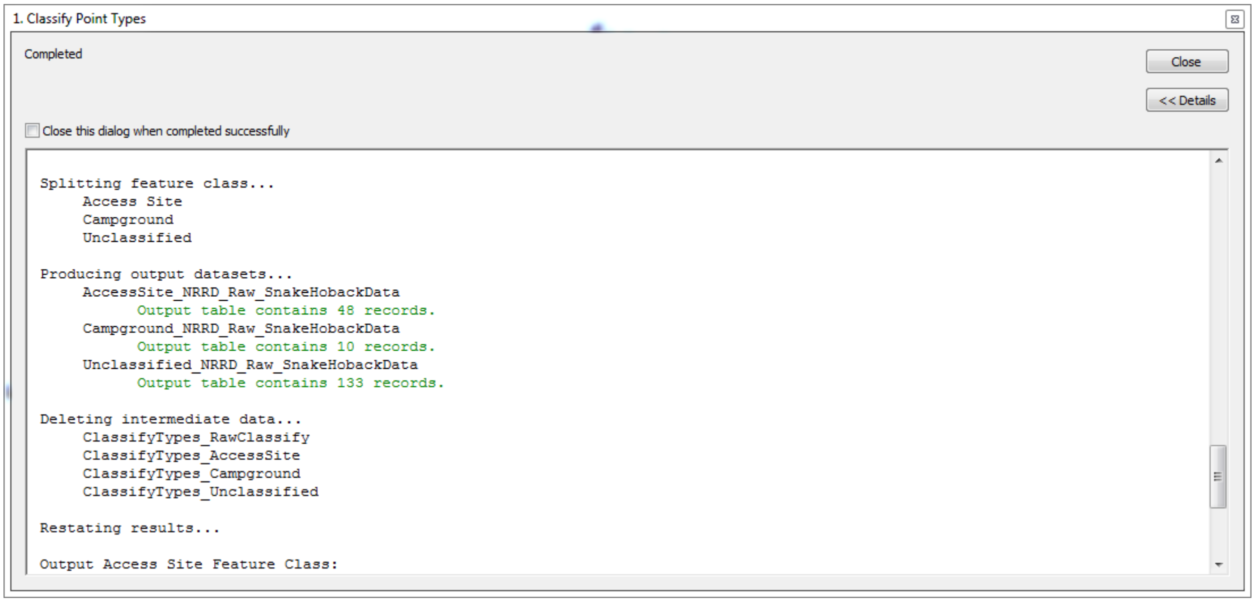


Figure 8: Tool execution messages within ArcMap also help the donor with QA/QC. Here, the number of features falling into each NRRD category help the donor determine if their SQL queries select the intended features. This helps cut down on subsequent work that may have to be redone if the queries did not function as intended.

SQL queries can easily introduce errors, especially if working with a large number of features. As the donor can see the selections taking place, tool messages can help them determine if a query selected features incorrectly (e.g. if there were 0 access sites when there should have been some). This helps cut down on subsequent work that has to be redone later. Donors can always check the output of each tool to determine if queries executed properly on individual features as well.

3.2.2 Classify Point Locations

The **Classify Point Locations** tool is also worthy of some specific explanation due to its links to the linear referencing part of the workflow later. Figure 9 shows its tool dialog box.

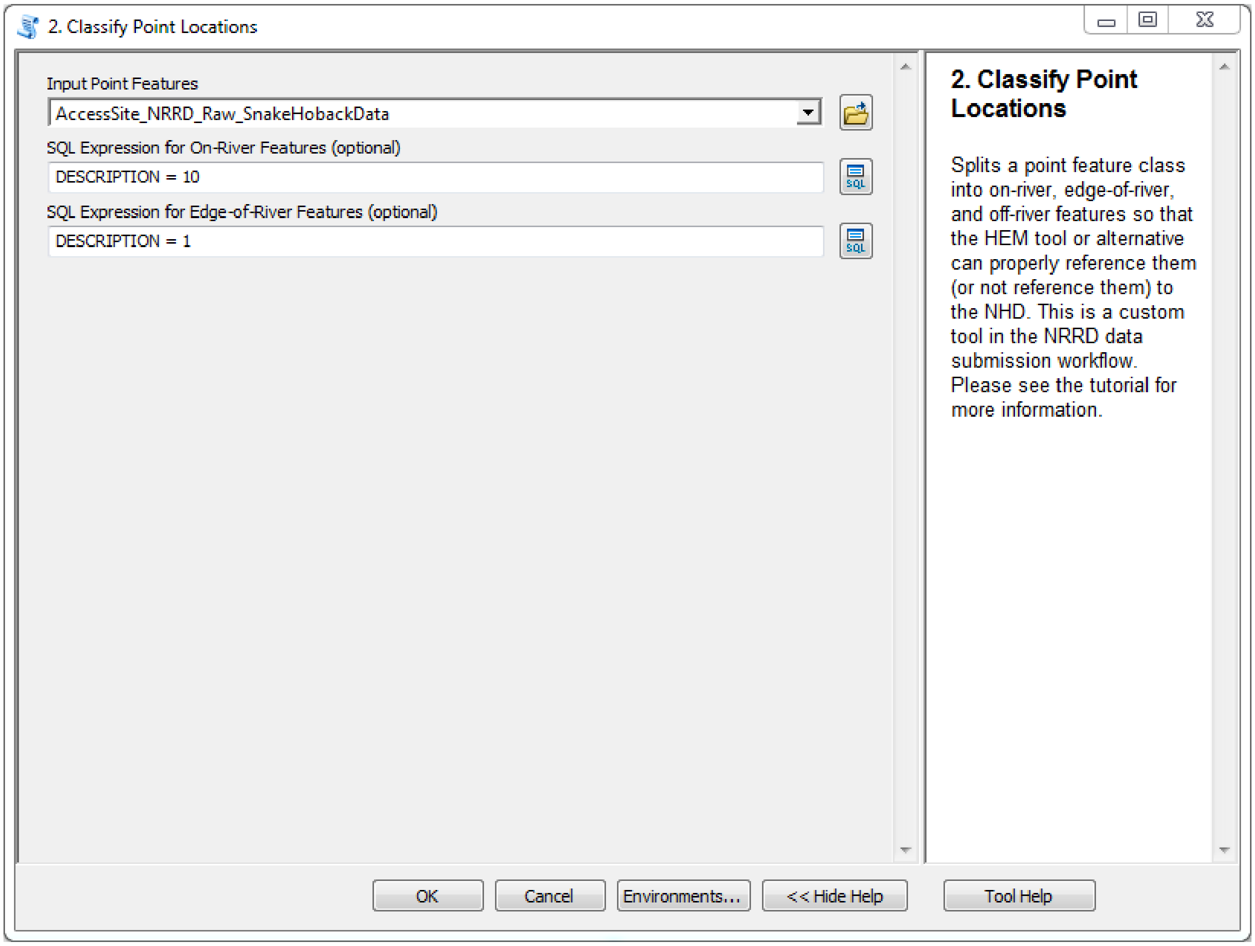


Figure 9: The tool dialog box for Classify Point Locations. Here, donors decide how significantly their points will be linearly referenced based on three "generic locations". Later in the workflow, on-river points will be snapped (moved) to NHD rivers, edge-of-river points will be related to rivers but not snapped, and off-river points will not be referenced at all.

The Classify Point Locations tool gives the donor the ability to determine how significantly their points are linearly referenced to the NHD later in the workflow. While it was determined essential to snap donor river reach features to NHD river reaches, the same was not the case for point features. Because some point features exist farther from rivers, but may still be relevant to recreation along them (e.g. a highway pullout observation point or a campground), not all points should be snapped to their nearest NHD flowline. Instead, it was decided that the donor should decide between three levels of linear referencing for their point features based on three “generic locations”. On-river features, which include phenomena such as rapids, should be completely moved to the nearest location on an NHD river. Edge-of-river features, which include boat launches, swimming areas, etc., should be related to the nearest NHD river reach (in their attribute table), but not moved. They thus maintain their original locations, which may be important for people to find on maps. Finally, off-river features are those far enough away from rivers that it does not make sense for them even to be related to the nearest NHD flowline. This might be the case for some campgrounds. These features are not linearly referenced at all. SQL queries like in the previous tool allow the donor to classify their features into these three locations. The only physical change to the feature classes at this point is an extra attribute field is added called

“NRRD_PointLocation” that holds this value. This tool can be run on both access sites and campgrounds from the previous tool, with the output looking similar to Figure 10.

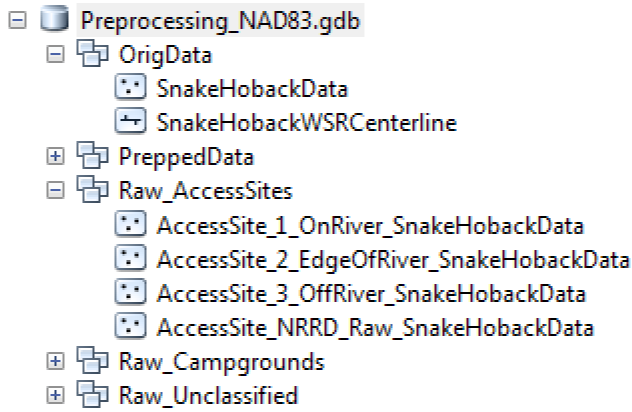


Figure 10: The three location types of each access site or campground feature class each receive their own feature class in the same feature dataset. If there are multiple raw access site feature classes (from multiple source feature classes), the prefixes on these new feature classes help organize them by location (e.g. on-river) rather than by origin, useful for the next tool.

3.2.3 Combine Point Features

There are two tools for step 3, **Combine Point Locations**, one each for access sites and campgrounds. Additional copies of this tool could be added if new NRRD point types are defined in the future. This tool performs two functions. First, in cases where there are multiple origin feature classes for a given point type (e.g. access sites), this tool will combine them for each location. For example, if there was both *AccessSite_1_OnRiver_SnakeHobackData* and *AccessSite_1_OnRiver_SnakeHobackOtherPoints*, this tool would combine them into *LinRefPrep_AccessSite_OnRiver*. The three output feature classes for this tool are all put in the *PreppedData* feature dataset, indicating that these points have been prepared for linear referencing. The three access site (or campground) locations cannot be combined yet because linear referencing must occur differently for each, so this tool may appear to “do nothing” in cases where there is only one origin feature class.

However, this tool does still need to be run because of its second function, adding and populating a field called “NRRD_AutoID”. This field is simply an incrementing ID field similar to ObjectID, but that will not change during creation of new datasets. Donors will likely already have their own permanent ID field, but “NRRD_AutoID” was created in order to support linear referencing functions more easily, so it is added anyway. It also ensures that there will be no duplicate numbers in this field. The 3 outputs of this tool for access sites are shown in Figure 11. The additional attribute fields added to points as part of preprocessing are shown in Figure 12. Feature classes remain separated at this stage by point type (access site, campground) and location (on-river, edge-of-river, off-river), yielding a maximum of six point feature classes with the current NRRD schema. This completes the preprocessing currently required for point features, which are now prepared for linear referencing with the HEM tool or alternative.

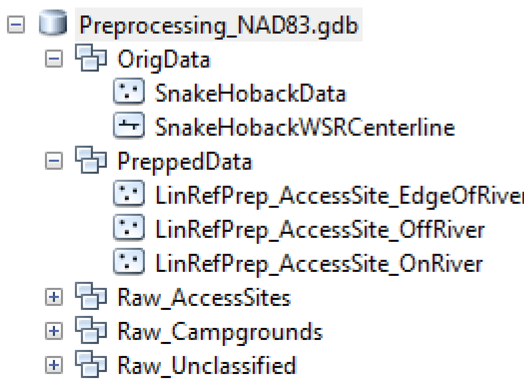


Figure 11: Access sites after being combined using Combine Access Sites. This merges in cases where there are multiple origin feature classes for access sites, but keeps locations separate because they must be linearly referenced differently.

NRRD_OrigFC	NRRD_PointType	NRRD_PointLocation	NRRD_AutoID
SnakeHobackData	Access Site	On-River	1
SnakeHobackData	Access Site	On-River	2
SnakeHobackData	Access Site	On-River	3

Figure 12: The 4 additional attribute fields currently added as part of preprocessing for point features. Features are classified by Point Type (intended destination in the NRRD) and Point Location (intended linear referencing method). Each combination has its own feature class in PreppedData. Original Feature Class and Auto ID help keep track of features during linear referencing and later in the workflow.

3.2.4 Prep River Reaches

There are also two tools for step 4, **Prep River Reaches**. In this case, however, only one option needs to be run on donor river reaches to prepare them for linear referencing. There were fewer preprocessing operations in general that needed to be run on river reaches, so a single step was created instead of multiple, as with points. There was, however, one issue that needed to be addressed, but would not apply to all donors, so two “options” of the tool were created. The issue was that some donors may have point data that represent rapids with associated difficulty class (Class I through VI). Stream difficulty in the NRRD, however, is represented by two attribute fields in river reaches, “Difficulty” and “Difficulty_Outlier”. It was decided that rapids point data should be converted into the “Difficulty_Outlier” field because rapids represent an increase in whitewater difficulty compared to the rest of the stream. The first option of tool #4 allows donors to join the difficulty of rapid point features to the line data of river reaches. For data donors without rapids point data, the second option just leaves this part out and calculates only the mandatory “NRRD_OrigFC” and “NRRD_AutoID” features. Donors will need to run one, but not both, of the two options on their river reaches depending on whether they have rapids they wish to join.

The tool dialog box for **4. OPTION 1 – Join Rapids to River Reaches** is shown in Figure 13.

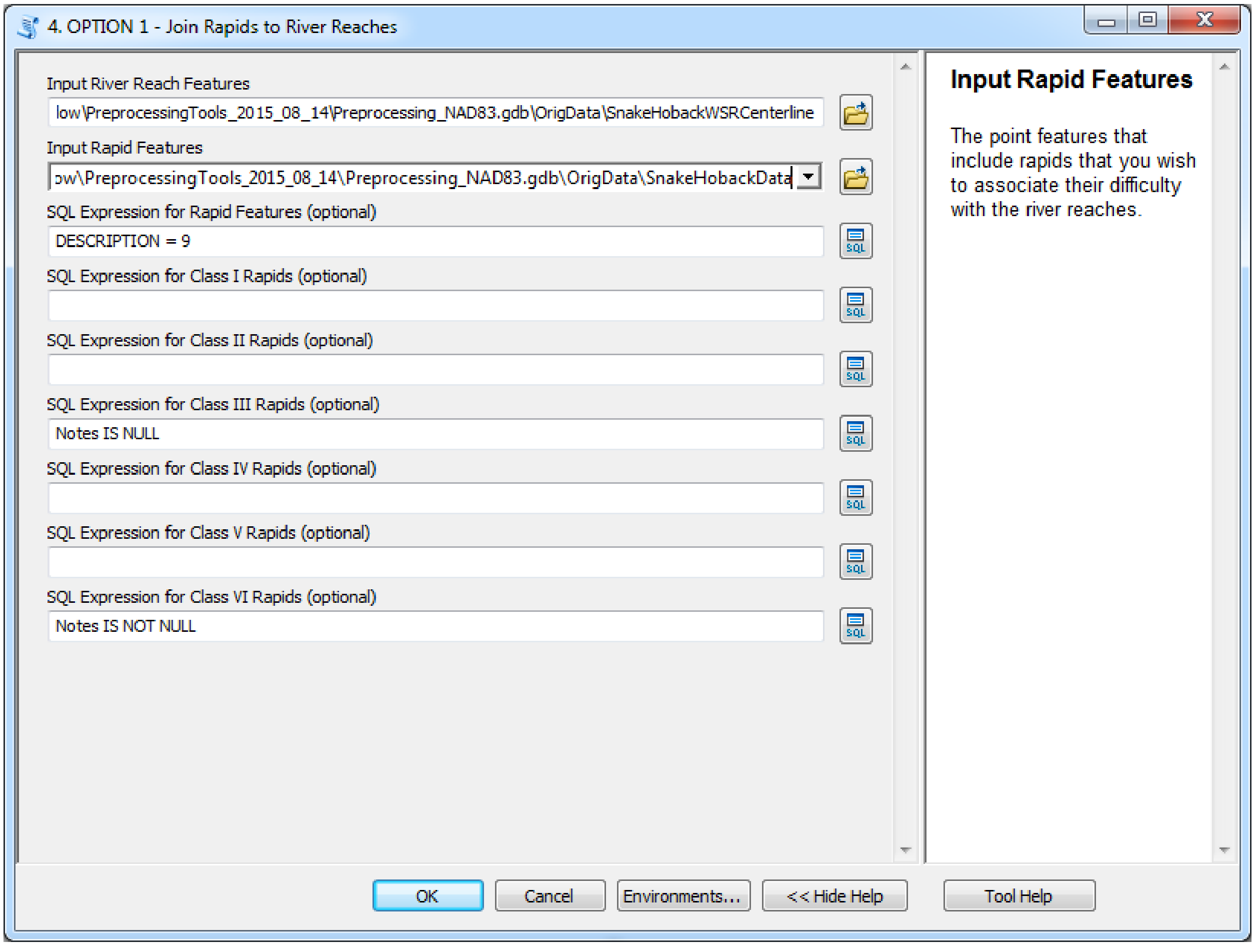


Figure 13: The tool dialog box for Join Rapids to River Reaches. Donors select the target river reaches and the source rapid points, then classify their rapids into stream difficulty (Class I through Class VI). The highest difficulty rapid out of the rapids nearest each river is then joined in a new field called NRRD_JoinRapid.

The donor selects the target river reaches and the source rapid points, then classifies their rapids into stream difficulty (Class I through Class VI). This is similar to some of the early point preprocessing tools. ArcToolbox's Spatial Join was the method employed to perform the actual join. Each rapid point is joined to the river reach that it is closest to. Then, using Summary Statistics, the highest difficulty out of all the rapids associated with each river reach is selected and only that difficulty is added in a new field called "NRRD_JoinRapid". The two standard new fields of "NRRD_OrigFC" and "NRRD_AutoID" are also calculated and the output is placed in *PreppedData* as no other preparation is required for linear referencing. If a donor does not have rapids, they can use the second option, Prep River Reaches, to just calculate "NRRD_OrigFC" and "NRRD_AutoID" and place the output in *PreppedData*.

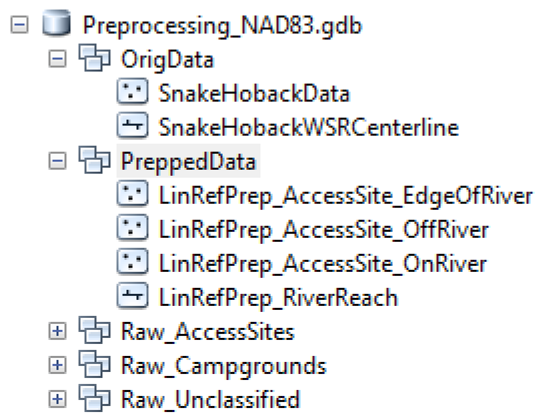


Figure 14: River reaches after being run through either option of tool #4. The only difference is the presence of the "NRRD_JoinRapid" field if the first option was chosen. No other preprocessing is required at this time for river reaches.

Both point and line data are now prepared for linear referencing with the HEM tool or alternative.

3.2.5 HEM Alternative

The **HEM Alternative** tools are described in more detail and compared to the USGS's Hydrography Event Management (HEM) tool in the next section. However, two tools remain to run *after* linear referencing that are still considered part of preprocessing. These are described below.

3.2.6 Restore Edge-of-River Points

The **Restore Edge-of-River Points** tool moves points back to their original locations before being snapped to NHD flowlines. Edge-of-river points are features that the donor indicated were worth locating along a river reach, but that still had a latitude-longitude location that was worth preserving. This classification was done in the Classify Point Locations tool. For example, a campground might be located along a river and so would benefit from being referenced to that river, but snapping the point to the centerline of the river removes the information about which side of the river the actual campground is on. This tool forms the final step in this process of moving the points back to their original locations after having received the Reachcode and Measure values from linear referencing. This can be run on the output of either the HEM tool or alternative tool.

The tool dialog box for Restore Edge-of-River Points is shown in Figure 15.

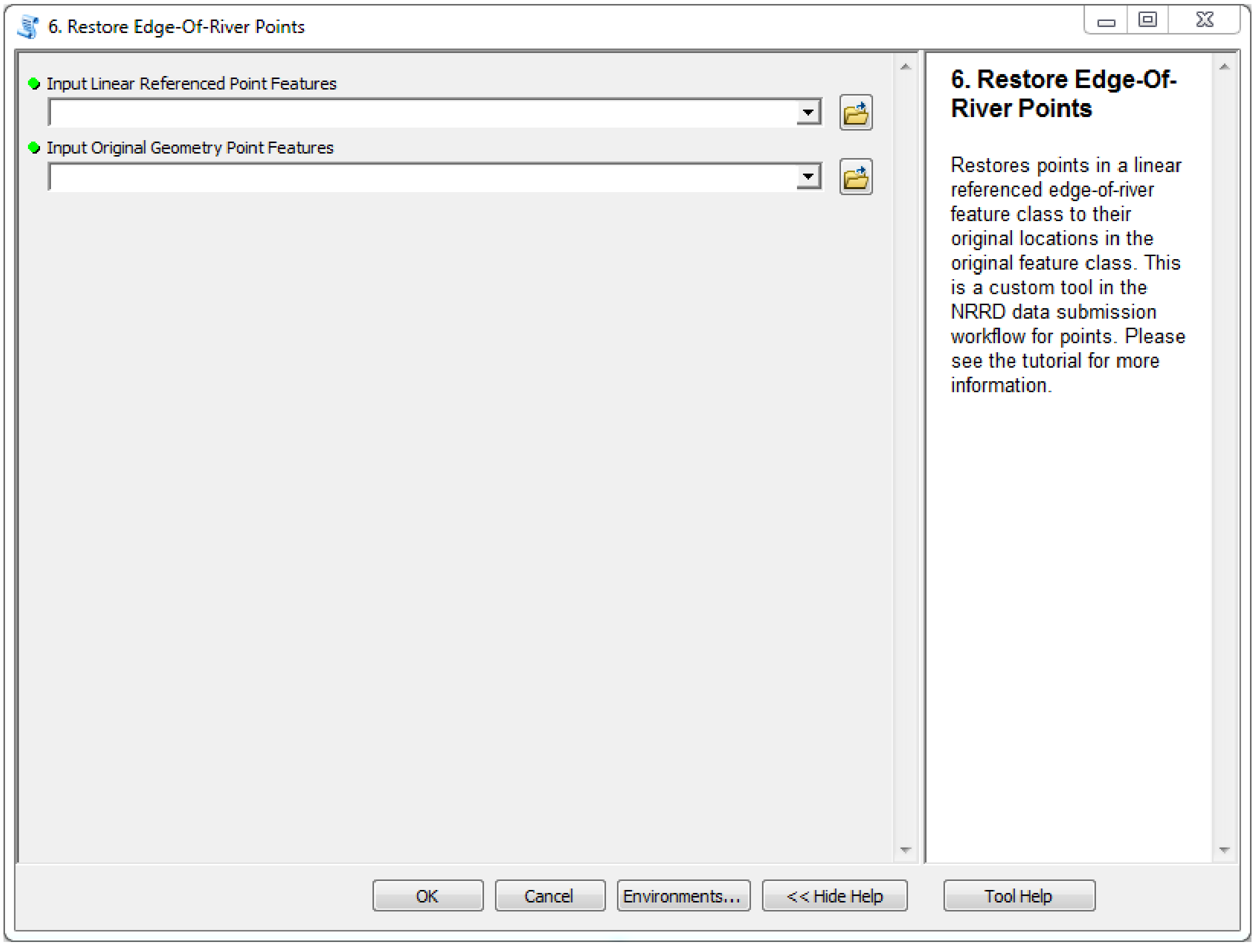


Figure 15: The tool dialog box for Restore Edge-of-River Points. This tool allows the donor to move linearly referenced point features back to their original locations while still preserving the attribute link to the closest river and measure.

The donor needs only to specify where the linearly referenced edge-of-river points are, which can be in different geodatabases depending on the linear referencing method used, and the original edge-of-river points, which are in the *PreppedData* feature dataset of the preprocessing geodatabase. The tool then creates a blank table with only the "NRRD_AutoID" field, adds the features from the original feature classes, then joins the rest of the fields from the linearly referenced feature class with Join Field, using the "NRRD_AutoID" as the join field. This is one reason why "NRRD_AutoID" always needs to be created during preprocessing. The output is placed in a new, blank geodatabase called *NRRD_BeginPostprocessing.gdb*. This output is shown in Figure 16.

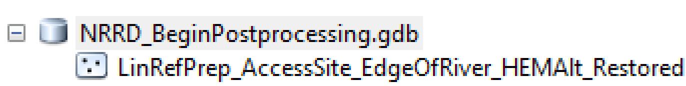


Figure 16: Output of the Restore Edge-of-River Points tool for access sites (in this case, linearly referenced using the HEM alternative tools). The new geodatabase will mainly be used to hold datasets that are ready for the NRRD post-processing tools.

3.2.7 Combine Linear Referenced Feature Classes

One final tool, **Combine Linear Referenced Feature Classes**, completes the preprocessing workflow as developed in this project. This tool merges all the versions of a NRRD feature class that are still separate and puts the output into *NRRD_BeginPostprocessing.gdb*. For example, for access sites, on-river points were linearly referenced; edge-of-river points were linearly referenced and then restored to their original locations, but with new attributes; and off-river points were not linearly referenced at all. These feature classes are now in three different geodatabases with names consisting of several suffixes. This tool merges these into the coherent *NRRD_AccessSite* feature class, ready for post-processing.

The tool dialog box for Combine Linear Referenced Feature Classes is shown in Figure Figure 17.

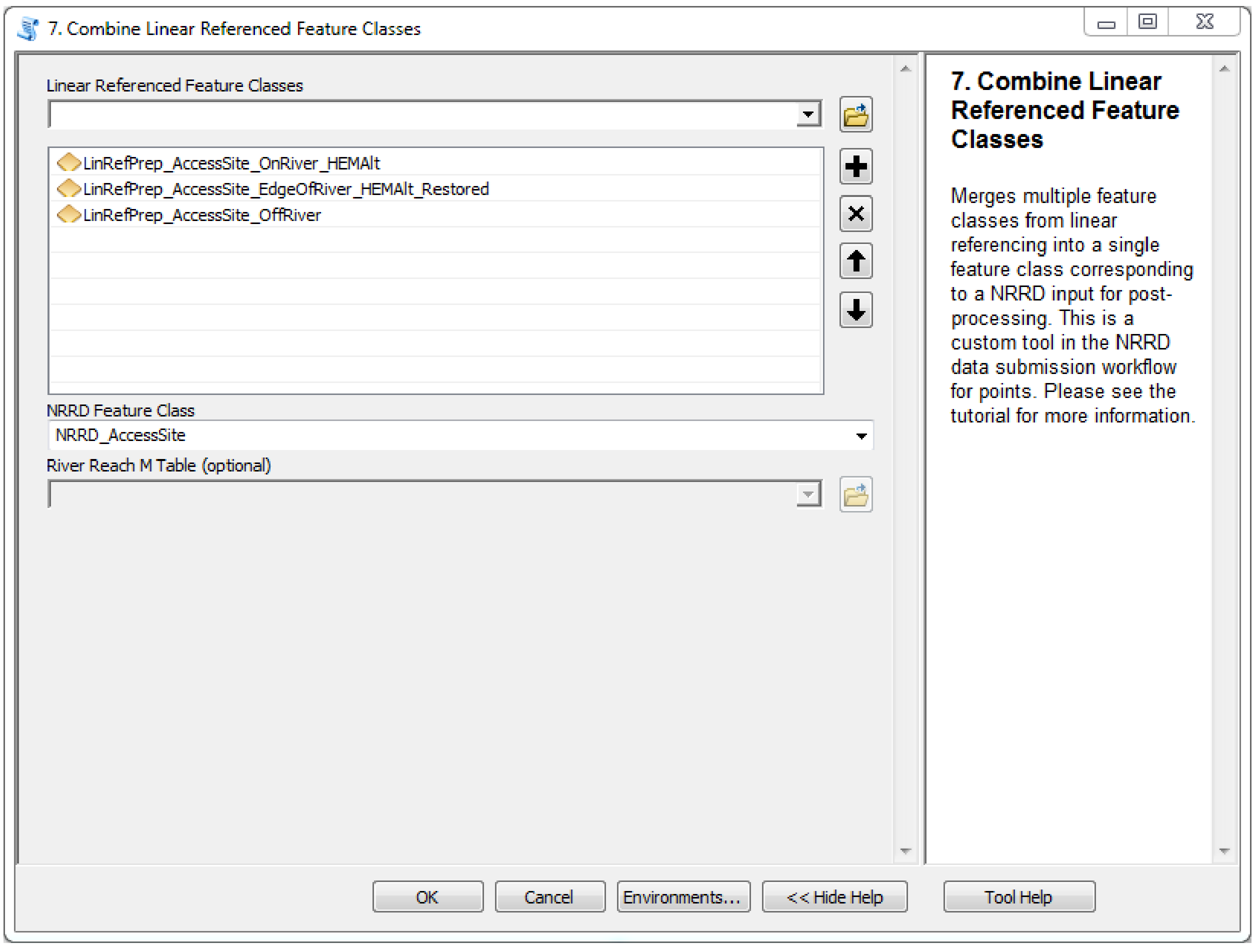


Figure 17: The tool dialog box for Combine Linear Referenced Feature Classes. This tool finally combines all subsets of NRRD feature classes into a single one, renames it, and copies it to the NRRD_BeginPostprocessing.gdb geodatabase. River reaches also requires the copying of an additional table, entitled the "M Table", that contains measures for each NHD river reach.

The cookbook instruction manual helps the donor locate which feature classes are actually meant to comprise the "final" output of preprocessing for river reaches, access sites, and campgrounds. The feature classes are then merged and copied to *NRRD_BeginPostprocessing.gdb*. In addition, the "NRRD_AutoID" field is recalculated after merging so that numbers remain unique. This breaks

NRRD_AutoID's connection with previous feature classes since the number has changed for each feature, but it is necessary for this attribute to be potentially used as a unique identifier later in the workflow. However, all "DONOR_" and "NRRD_" fields, along with the HEM linear referencing fields, are preserved in this tool's output. This gives the donor access to all detail from the original feature classes going into post-processing, along with additional information such as "NRRD_OrigFC" and "NRRD_PointType." Figure 18 shows the output of this tool for access sites and river reaches. Campgrounds would be similarly named.

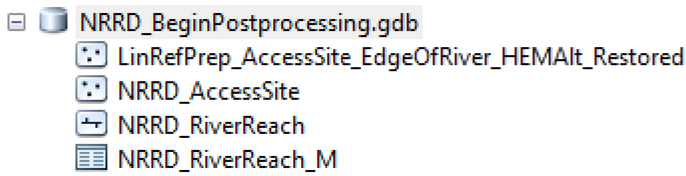


Figure 18: Output of the Combine Linear Referenced Feature Classes tool. Using a separate geodatabase to store the feature classes ready for post-processing, as well as naming them more simply, helps indicate to the donor what feature classes they should take into post-processing. River reaches also includes a support table called the "M Table" for measurements.

The donor only needs to run this tool once for each NRRD feature class they are preparing. If preparing river reaches, an additional "M Table" is also required to be copied, which the cookbook will also help them locate. These final feature classes have been passed through both preprocessing to handle specific issues, and through linear referencing to align them or relate them with the NHD. The short names of the resulting feature classes, along with the separate geodatabase, help indicate that these feature classes are ready to be moved into the final phase of NRRD preparation.

3.3 Linear Referencing

Taking a step back, linear referencing, as one of the three phases of the NRRD data submission workflow, deserves its own section because of both how critical it is to combining multiple datasets across the country and how much work went into making it as straightforward as possible despite its complexity. The National Hydrography Dataset (NHD) provides a nationally verified "drainage network with features such as rivers, streams, canals, lakes, ponds, coastline, dams, and streamgages" (USGS, 2015). In the case of the National River Recreation Database, the plan is to linearly reference both line and point data, such as floatable river reaches and access sites, into line and point "event" features along the NHD. The USGS encourages this by stating that "[the] use of events is a key characteristic of the NHD by allowing vast amounts of scientific information to be linked to the NHD while keeping the design simple and by making advanced analysis techniques possible" (USGS, 2015).

Event feature classes are identical to normal ArcGIS feature classes except that features must be on top of NHD flowlines (rivers, streams, etc.) and also carry information about which stream they are on (Reachcode) and how far from the nearest downstream confluence they are (Measure). In the case of line events, a "From Reachcode" and "To Reachcode" are required, as well as "From Measure" and "To Measure", in order to pinpoint the start and end locations of the line. Floatable river reaches, usually just referred to as NRRD "river reaches" in this paper, would thus be represented as lines "on top of" the base NHD flowlines, usually covering more than one stream segment. Point data such as access sites would be represented by single points located along NHD flowlines.

While the stance on using strict events was lessened for point data in order to encompass more types of data (e.g. points farther from rivers, but still relevant), it still made sense to require linear referencing for floatable river reaches. What follows is a description of the two methods supported in the NRRD data submission workflow for linearly referencing points and lines to the NHD. The first, the USGS's own Hydrography Event Management (HEM) tool, is the recommended option, while the second, a set of "HEM Alternative" tools developed during this project, is available for donors who have an incompatible version of ArcGIS for the HEM or who feel more comfortable using other tools.

3.3.1 The USGS Hydrography Event Management (HEM) Tool

The United States Geological Survey (USGS) provides the Hydrography Event Management (HEM) tool for linearly referencing features to the National Hydrography Dataset. The tool functions as an ArcGIS add-in and custom toolbar. As this tool was the only robust way of performing linear referencing at the start of the project, it was one of the major determiners of pursuing an ArcGIS, rather than an open-source, workflow (see Discussion section for more information). Figure 19 shows the HEM toolbar.

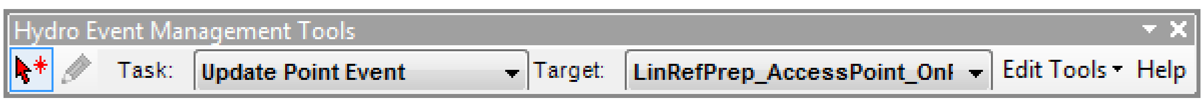


Figure 19: The toolbar of the HEM tool in ArcMap. Entire feature classes can be snapped to the NHD as well using this tool.

The HEM tool allows donors to provide an input feature class and a reference tolerance (maximum distance to move), and have their features be snapped to the NHD to become events. Each point feature is snapped to the nearest point along an NHD river reach and each line feature is snapped to the route along the NHD that most closely matches the input feature. The HEM tool also provides a set of editing functions to snap features that were either outside the reference tolerance from a river or that were snapped to the incorrect river (e.g. if a smaller river ended up being closer to an access site than the river that the access site is actually on).

The post-editing functionality is one of the greatest benefits of the HEM tool over the alternative tools because it automatically updates the Reachcode and Measure attributes within the feature class being edited. This is not the case with the alternative tools. However, the HEM tool has many functions and is easy for new users to become overwhelmed. Two 2-hour training sessions were set up with Michael Tinker of the USGS to introduce the project team to the HEM tool and to some standard workflows. This investment presents a significant challenge when there could easily be hundreds of organizations wanting to submit data to the NRRD. As such, the approach taken was to find the minimal necessary workflow to submit data to the NRRD and provide careful instructions in the cookbook on that particular workflow.

Other shortcomings of the HEM include long processing times (multiple hours is not uncommon for a feature class import) and compatibility issues. One major issue is that a HEM tool version does not currently exist for the latest version of ArcGIS (10.3 as of this writing). This means that donor organizations that have upgraded would be unable to use the HEM tool at all. Because of the time investment required to learn the tool and the incompatibility with ArcGIS 10.3, a set of "HEM alternatives" using only built-in ArcToolbox tools and Python was also created.

3.3.2 The HEM Alternative Tools

The HEM alternative tools are provided in the *NRRD_Preprocessing* toolbox. Figure 20 shows these tools again.

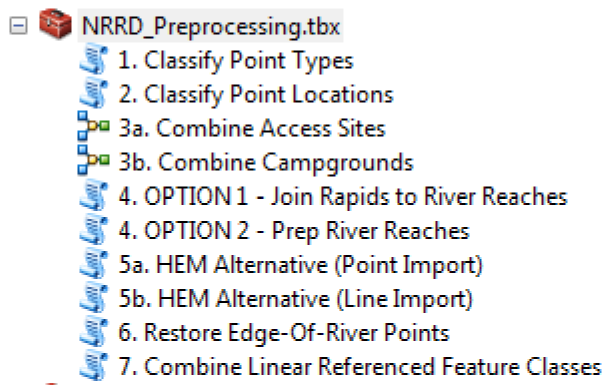


Figure 20: The HEM alternative tools are provided in the NRRD_Preprocessing toolbox. A different tool is provided for points and lines. Additional functionality could be added in extra tools in the future.

The current two tools cover point imports and line imports. This mimics the HEM functionality of linearly referencing entire feature classes at once ("Import to Events"). Figure 21 shows the tool dialog box for HEM Alternative (Point Import), which also mimics the "Import To Events" dialog box in the HEM tool. The two HEM alternative tools also both output to a copy of a HEM output feature class for points or lines; this gives the output identical fields to the HEM.

The HEM alternative tools primarily employ the Linear Referencing toolbox in ArcToolbox, including Locate Features Along Routes and Make Route Event Layer. A start to these tools was provided by the National Park Service as they had a working HEM alternative for lines in ModelBuilder. This was extended in Python to handle points and to provide more specific support for NRRD feature classes, including using the "NRRD_AutoID" field to help track features throughout the process.

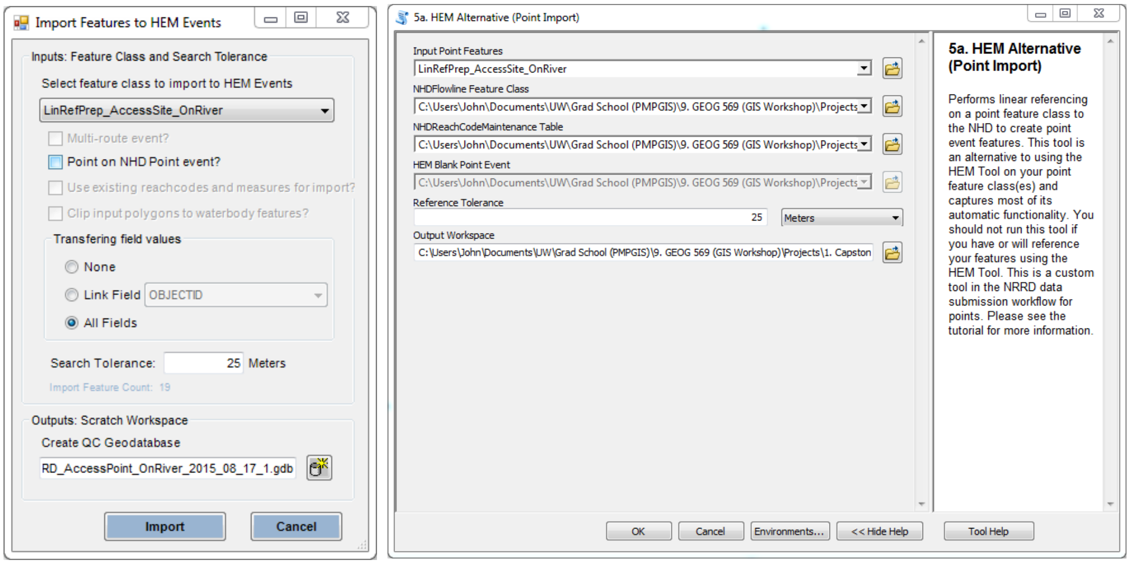


Figure 21: Dialog boxes for both the HEM tool (left) and HEM Alternative (Point Import) tool (right). Both tools take a point feature class from PreppedData and move features to the nearest point on an NHD river reach. Both tools also take a reference tolerance (maximum move distance) and output to a separate geodatabase.

Working on the alternative tools, however, really demonstrated how much the actual HEM tool accomplishes on its own. One issue is editing after the initial snap. A donor can use ArcMap's built-in editing functionality to move points in the output feature class that might have been incorrectly moved, but that does not update the "Reachcode" and "Measure" fields as the HEM edit session does. To update these fields, a second tool would have to be made for both points and lines. While this would mostly entail just taking a portion of each "import" tool to make an "update" tool, there was not enough time in the nine weeks to do this, so this was left as a potential gap in the workflow if a donor needs to edit snaps.

More importantly, the HEM tool is also much better than the alternative at line referencing. Points are more straightforward to match than lines, which often have to cross multiple features in the NHD to attempt to match the original line. The ArcToolbox functions will create a separate feature for each NHD feature along a given trace, meaning that a single input line can become multiple features in the output. Normally, the Dissolve tool could be used to recombine these features into a single line. However, this problem is exacerbated by the reference tolerance. If too low, there can be gaps between new lines where the input line diverged too much from the NHD flowlines. If too high, there can be "stubs" where NHD rivers that are *not* the intended route can be traced as well. Figure 22shows examples of both stubs and gaps encountered while linearly referencing river reaches. Both of these errors created difficulties when trying to dissolve line features back together. In the actual HEM tool, there is an option called "Multi-Route Event" that indicates that this dissolving should take place.

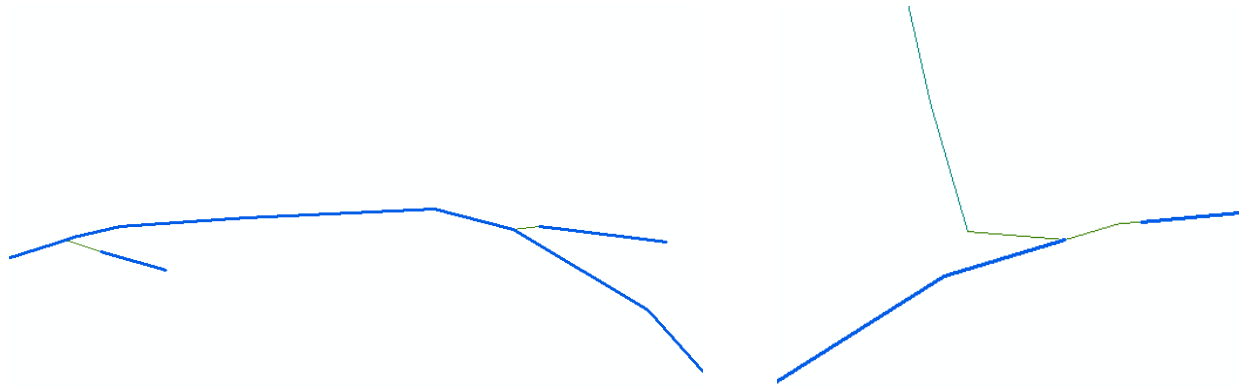


Figure 22: Examples of stubs (left) and gaps (right) encountered during linear referencing of river reaches without the HEM tool. Stubs tended to appear when the reference tolerance was too high and gaps when it was too low. Both presented difficulties in dissolving raw features back into full lines.

These situations were handled by a mixture of other ArcGIS functions such as Snap to fix gaps and by judgment calls based one whether features connected on both ends. This led to satisfactory results with the Bridger-Teton National Forest data, but many remaining errors with the much larger Minnesota dataset. Ultimately, it was decided that the amount of extra functionality to truly handle these cases was probably the inspiration for making the HEM tool in the first place, meaning that more work would necessarily tend towards the HEM tool anyway. As such, the HEM alternative tools would be left in as an option for small datasets or ones that already used the NHD (leading to less errors to begin with), but that the HEM tool would still be the recommended approach in most cases.

To help donors gauge how well the HEM alternative tools performed on their river reaches, two additional datasets are created with the required output of the referenced river reaches and M table. First, part of the original tool created by the National Park Service, a Topology dataset is created that points out any disconnected ends of lines. By comparing these points to the input lines, any unintended gaps in the output river reaches can be identified. In addition, the Snap tool from ArcToolbox, used to fix gaps, can cause lines to diverge from the NHD slightly to snap to each other. To identify these divergences, which violate the "event" specification of the output, the Erase tool is used to show where output line features do not follow the NHD. These divergences are usually small (shorter than the reference tolerance). Additional QA/QC datasets could be created for line linear referencing, but time constraints stopped it at two. Figure 23 shows the output of the HEM alternative tool used on Bridger-Teton National Forest river reaches, along with the two QA/QC datasets.

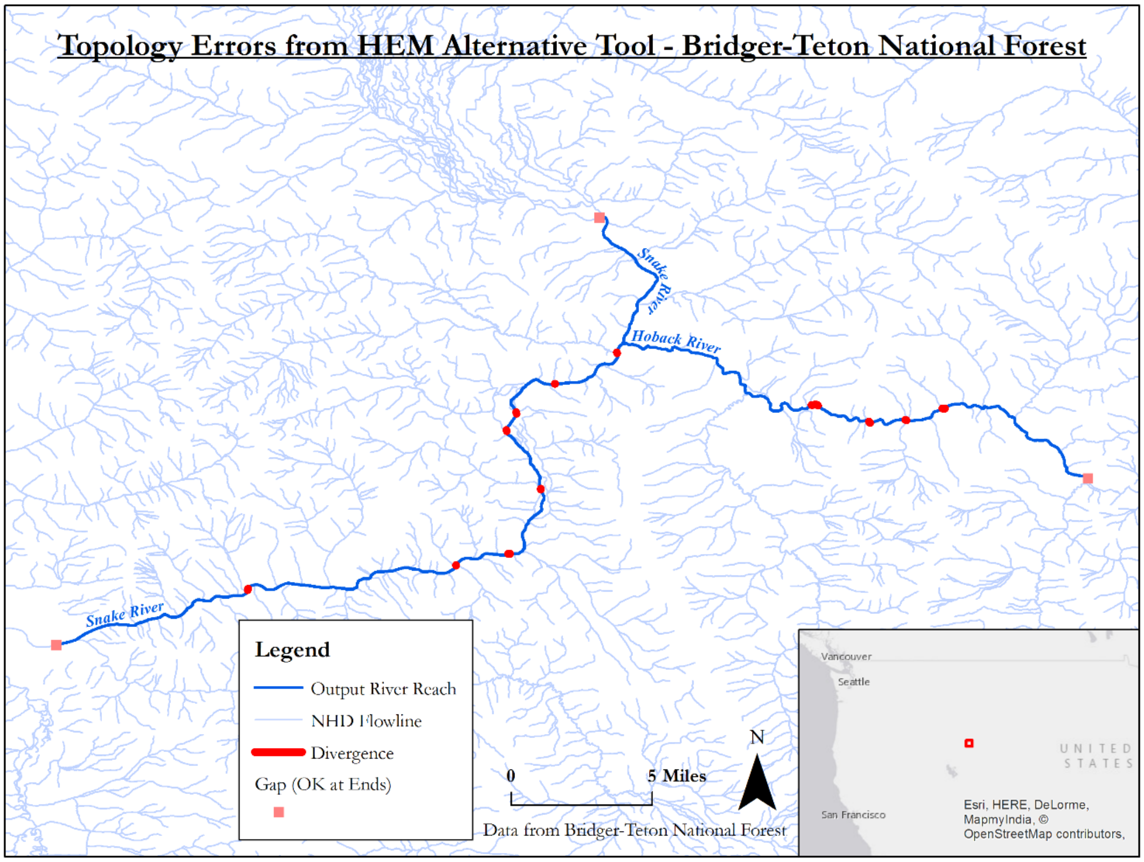


Figure 23: Output of the HEM Alternative (Line Import) tool. The number of NHD flowlines in the map shows the difficulty of selecting a continuous route to match input features. The tool outputs two additional datasets, Divergence and Gap, to help donors see where errors might have been introduced in producing a continuous dataset.

The output of linear referencing is usually in a separate geodatabase. The results are then passed to the remaining two preprocessing tools in the previous section to finish preparation for post-processing.

3.4 NRRD Post-processing Toolkit

Once donor data has been preprocessed and linear referenced, it can be run through a selection of relevant tools to transform the data from the donor's format to the NRRD format. The toolkit consists of three toolboxes, one for each feature class in the NRRD: River Reaches, Access Sites and Campgrounds, plus a series of geodatabases to hold intermediate and final data. **Error! Not a valid bookmark self-reference.** shows the tools the toolkit. Each individual tool populates one or more fields in the corresponding NRRD attribute table. The toolkit is designed so that the donor need only run the tools that transform the data they have and can skip any tools that are meant to populate fields for which the donor has no comparable data. For example, if a donor has no information about campgrounds, they can skip the entire Campgrounds toolbox, or if a donor has no data about parking at access sites, they can skip the "Access Site – 2. Parking" tool and move on to the next one.

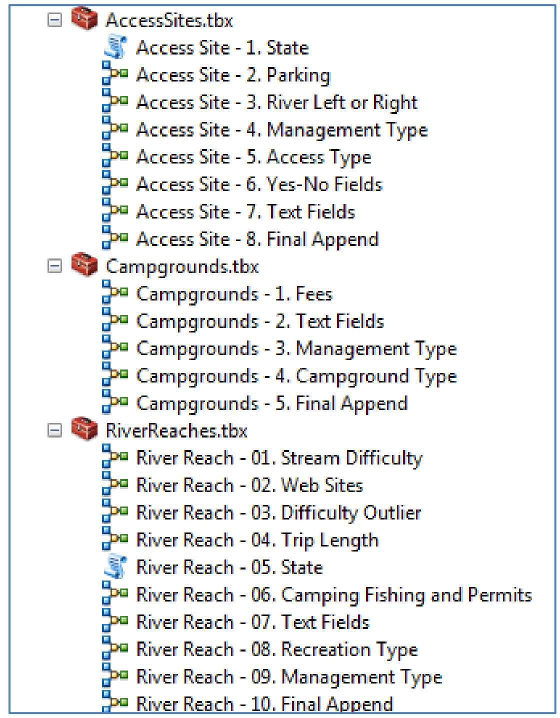


Figure 24: NRRD post-processing toolkit

In addition to the toolboxes, the toolkit holds a folder called "scratch" that houses a series of geodatabases for holding the NRRD templates and intermediate data sets created by the tools. Figure 25 shows this file structure. These include "NRRDgdb_Scratch.gdb" for holding temporary files, "NRRDgdb_Template" which holds blank feature classes with the NRRD schema, "NRRDgdb_Stages.gdb" which holds the output of each tool until they're ready to be combined, and "NRRDgdb_Final" which holds the final output of each tool.

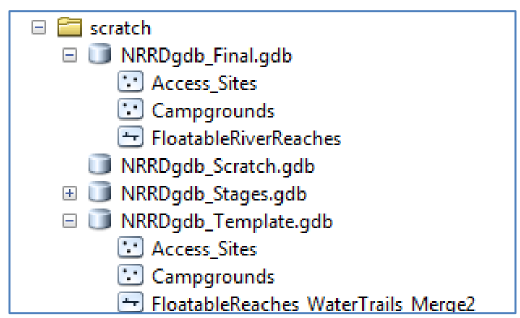


Figure 25: Scratch folder holding necessary geodatabases

3.4.1 Model setup

With the exception of the two "State" tools, all of the tools were built using ArcGIS Model Builder functionality, and are set up with a common format. Examples of each tool model are shown in Appendix A. Figure 26 shows how each model is set up with input parameters (in this case for the "Access Site – 2. Parking" tool). All of the tools use the same steps to begin with, as follows:

1. The user specifies the Pre-processed Access Sites feature class – this should be the output of the preprocessing tools, located in the "NRRD_BeginPostprocessing.gdb" geodatabase.

2. The user then enters SQL expressions into the dialog box to point to data in the specified feature class that corresponds to each value in the NRRD field (in this case the Parking field has a domain with three possible values, “overnight,” “day-use,” or “no parking.”)
3. Once the inputs are entered, the tool prepares for the selection process, and
4. Copies all the features to a new, temporary file, located in the Temporary Geodatabase (“NRRDgdb_Scratch.gdb” as seen in Figure 25).
5. As a time-saving measure, a variable (called “Boolean”) is set as a precondition to the select commands, and is set to false. This prevents the select tools from running right away (they run

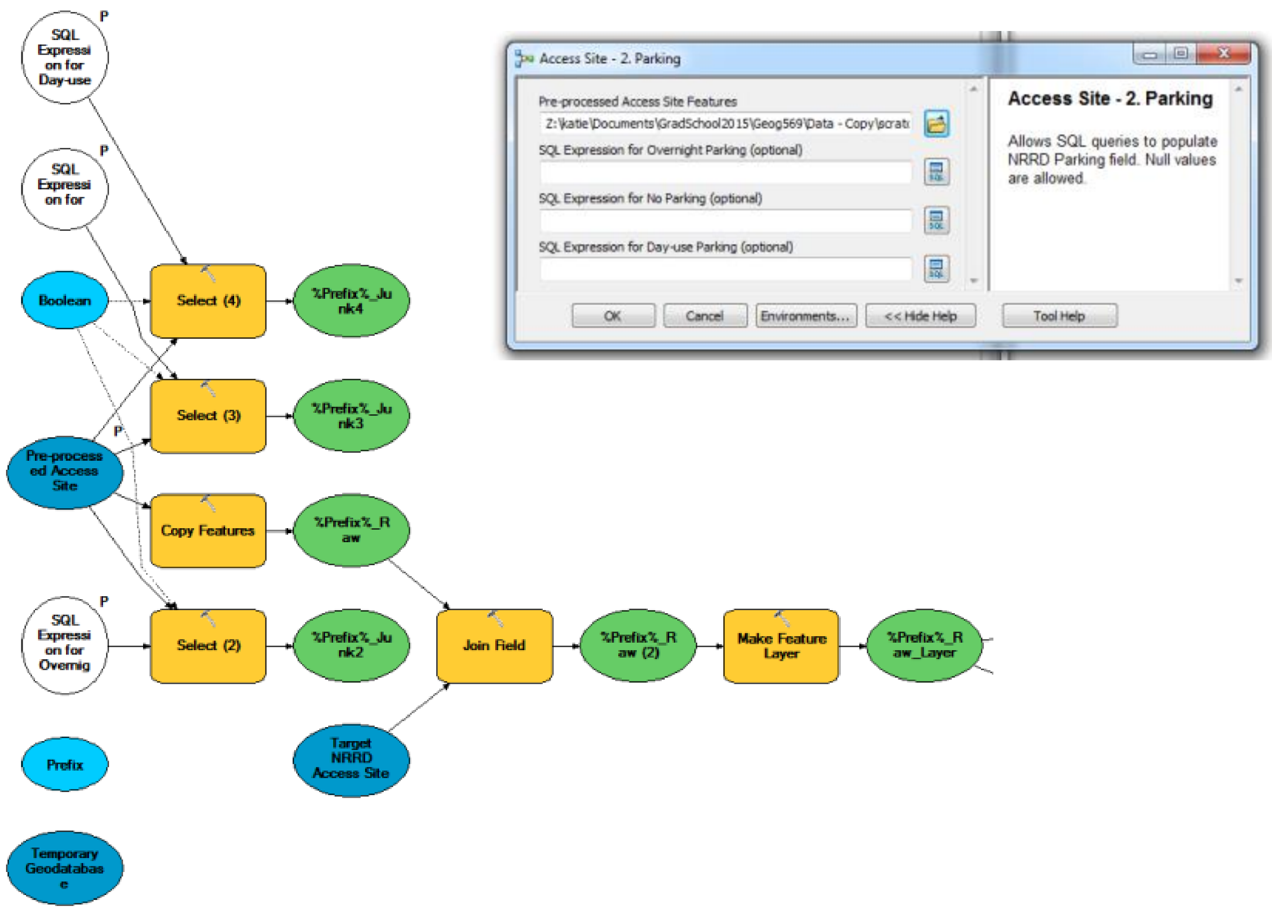


Figure 26: Start of a model showing user inputs and the creation of a feature layer on which all operations are run

later in the model), and shaves about 30 seconds off of the tool run time for some of the longer tools.
6. The “Prefix” variable is assigned the name of the toolbox (“AccessSite,” “Campground” or “RiverReach”) and is added to the name of every intermediate step in the model in order to aid in tracking files.
7. Next, the copied features (from step 4) are joined to the appropriate feature class from the “NRRDgdb_Template.gdb,” and
8. The output is saved as a feature layer. This layer is now ready to accept the input entered by the user.

3.4.2 Populating fields

Once the feature layer is set, each tool populates a field or set of similar fields using the user's input. The tools fall into three basic categories: tools that populate fields with coded value domains; tools that populate yes/no fields; and tools that populate text fields.

3.4.3 Coded value domains

River Reach - 01. Stream Difficulty tool
River Reach - 03. Difficulty Outlier tool
River Reach - 04. Trip Length tool
River Reach - 08. Recreation Type tool
River Reach - 09. Management Type tool
Access Site - 2. Parking tool
Access Site - 3. River Left or Right tool
Access Site - 4. Management Type tool
Access Site - 5. Access Type tool
Campgrounds - 3. Management Type tool
Campgrounds - 4. Campground Type tool

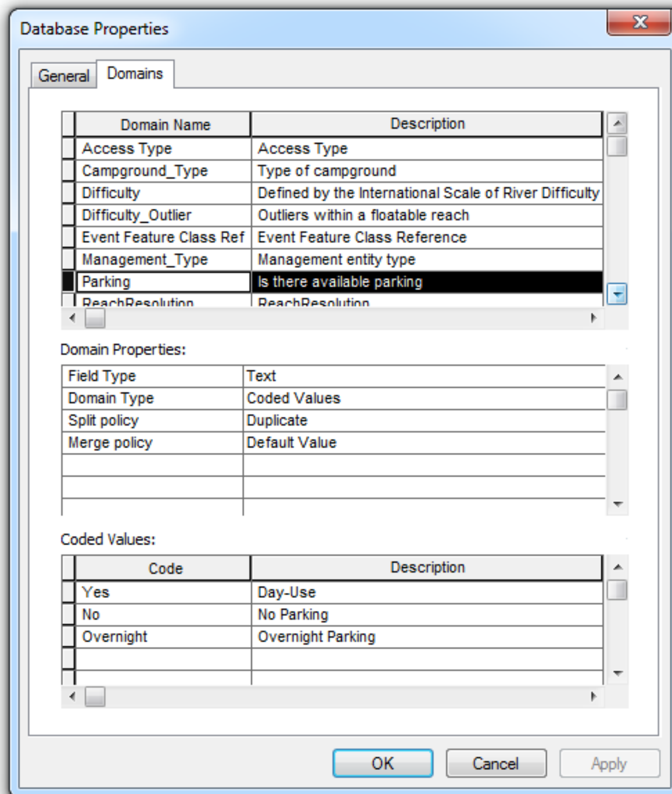


Figure 27: Database properties showing the coded values for the Parking field

Several of the fields in the NRRD feature classes have coded value domains, restricting the values in that field to a finite set. For example, Figure 27 shows the database properties for the "NRRDgdb_Template.gdb" geodatabase with the coded value domain for the Access Site "Parking" field highlighted. In this case, the field will be populated by the codes "Yes," "No," or "Overnight." The dialog box in Figure 26 shows where the user enters expressions that point to each of these values in their own data. The tool then iterates through each possible code, calculating the value. Figure 28 shows one such iteration. The steps are as follows:

1. A precondition is set on the "Select Layer by Attribute" tool requiring it to check to see if an SQL expression has been entered for the given coded value. This precondition is a Python code block reading "True if """ %SQL Expression for Day-use Parking% """.strip() else False," as seen in Figure 29.

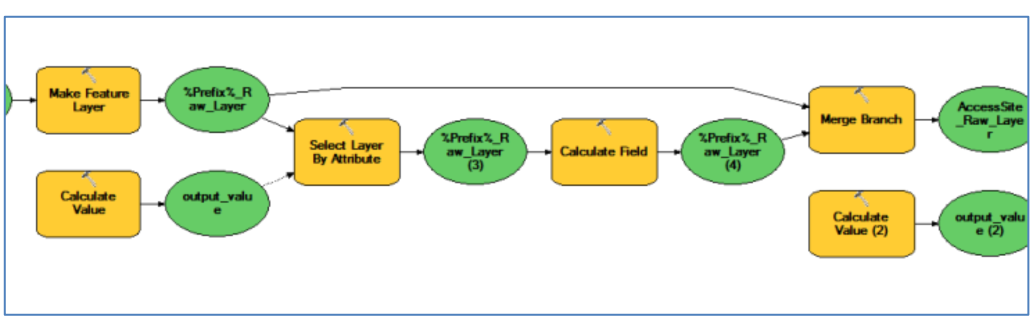


Figure 28: One iteration of the Parking model adding values to the field

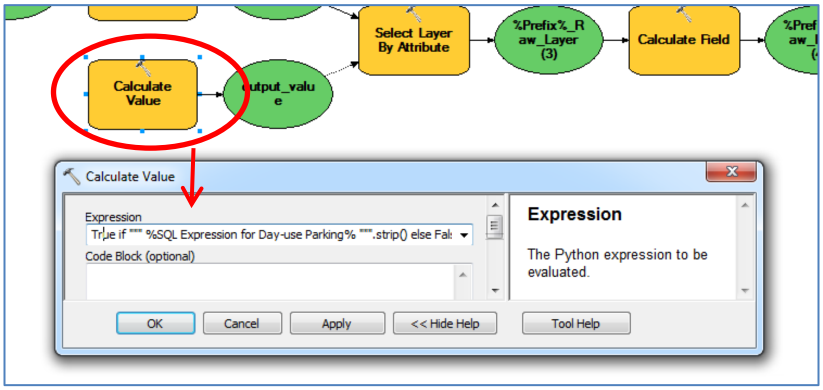


Figure 29: Calculate Value tool

2. If the condition is False (if there is no expression entered in the dialog box for the given code), the model skips this iteration and proceeds directly to “Merge Branch.” “Merge Branch” is a model-only tool that combines the output of two possible paths back into one.
3. If the condition is True (if there is an expression entered in the dialog box for the given code), the “Select Layer by Attribute” tool is run, as seen in Figure 30. This tool makes a new selection from the layer created during setup, and uses the SQL expression set by the user for the given code.

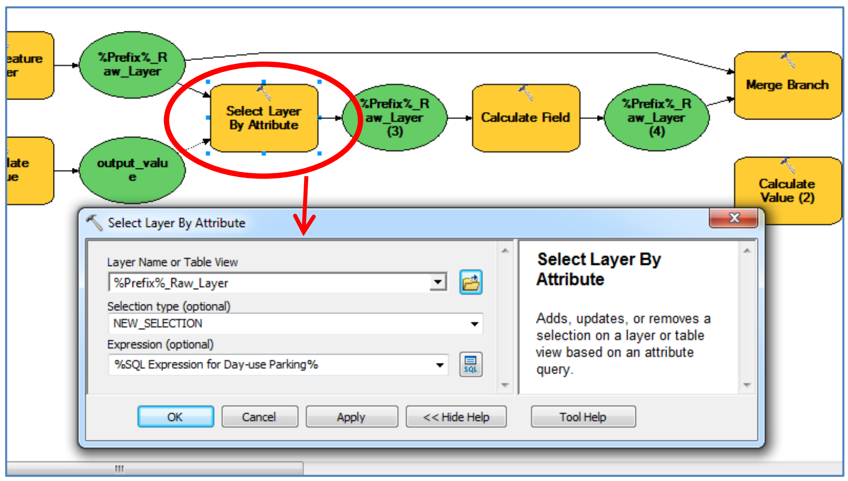


Figure 30: Select Layer by Attribute tool

4. Next, the “Calculate Field” tool is run on the selection, and the NRRD field “Parking” is populated with the code for the given value, as seen in Figure 31.
5. Finally, the model merges with the other possible branch using the “Merge Branch” tool, and creates an output layer that is ready for the next iteration. The model then executes the same sequence of steps for each one of the codes in the coded value domain.

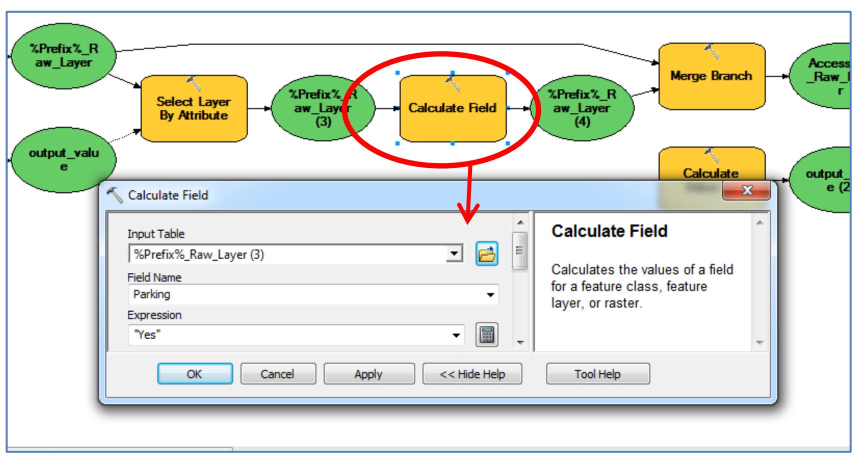


Figure 31: Calculate Field tool

3.4.4 Yes/no fields

River Reach - 06. Camping, Fishing and Permits
River Reach - 08. Recreation Type tool
Access Site - 6. Yes-No Fields tool
Campgrounds - 1. Fees tool

Each of the NRRD feature classes has at least one yes/no field. Since these are relatively simple fields to populate, multiple fields can be covered by one tool. There are six yes/no fields in the River Reach feature class attribute table, one each for whether or not fishing or camping is allowed and whether or not permits are required; and one each for whether or not the river reach is part of a water trail, whitewater run, or Wild and Scenic River (one reach could conceivably be all three). These fields are populated using two tools, because while null values are allowed for camping, fishing and permit fields, null values are not allowed for recreation type fields. There are also six yes/no fields in the Access Sites feature class, with options to designate the presence or absence of put-ins, take-outs, fees, emergency access, observation areas and restrooms. Since all allow null values, the six fields are populated using one tool. The Camping feature class has only one yes/no field, for whether or not there are fees.

The “River Reach – 06. Camping, Fishing and Permits,” “Access Site – 6. Yes-No Fields” and “Campgrounds – 1. Fees” tools are all structured the same way. Figure 32 shows the dialog box for the “River Reach – 06. Camping, Fishing and Permits” tool as an example. The user navigates to the input feature class, then enters an SQL expression to point to yes and no data for each of the three fields. These fields correspond to the SQL statement variables in the set up portion of each model. The set up is identical to the set up for all other tools. Once the tool set up is run and the temporary feature layer has been created, the tool iterates through each of the SQL expressions in the same matter as the coded value domain tools. The “calculate field” tool simply populates the selected records with “yes” or “no” instead of with a code.

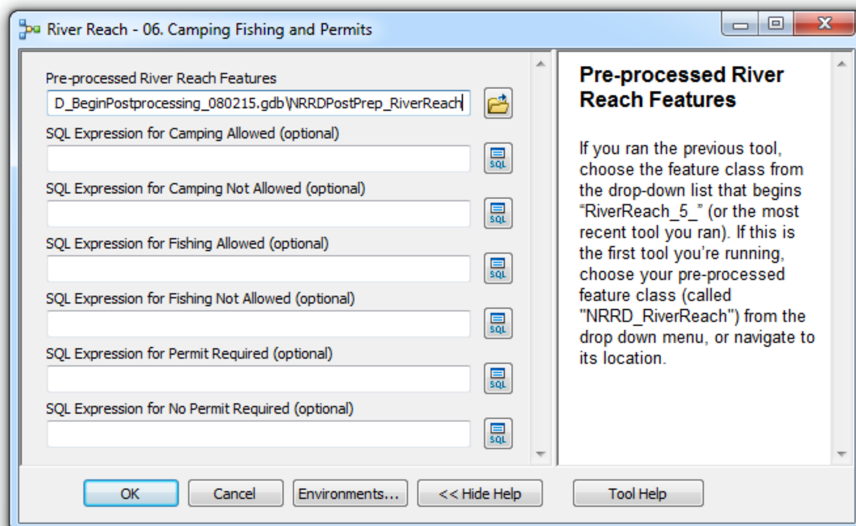


Figure 32: Example of a yes-no field dialog box

Since the “River Reach - 08. Recreation Type tool” has required values, the model functions somewhat differently.

1. To begin with, the user need only enter SQL expressions that indicate “yes” values for each of the three recreation types, as seen in Figure 33. Since null values are not allowed, anything that is not a “yes” is automatically populated with “no.”

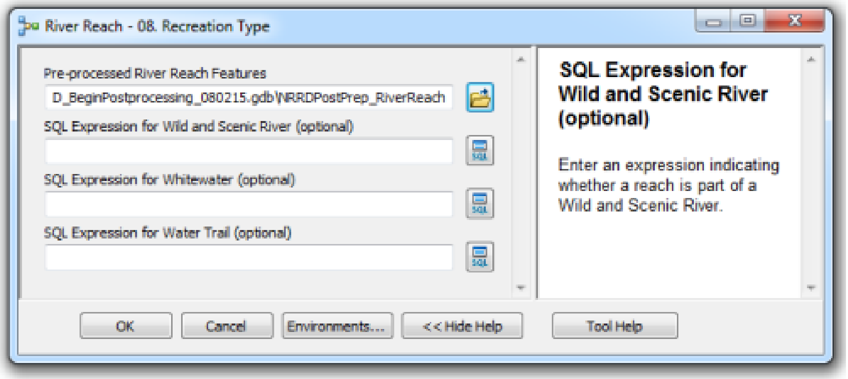


Figure 33: Recreation Type tool dialog box

2. The model begins in the same way as previous tools, creating a temporary feature layer on which to run operations.
3. The difference is that an extra set of tools is applied to each iteration in the model. After the initial field calculation on the records corresponding to the entered SQL statement (setting the value to “yes” for a given recreation type), the model performs a “switch selection” operation as seen in Figure 34, and then sets the value of those records to “no.”
4. The model otherwise operates as previous models.

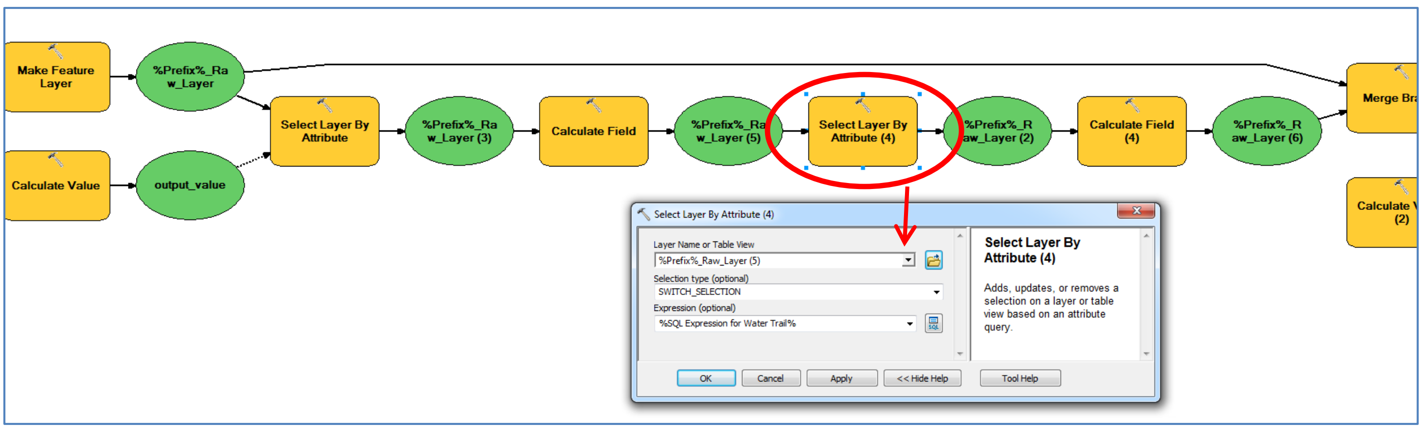


Figure 34: One iteration of the Recreation Type model showing the Switch Selection tool

3.4.5 Text fields

River Reach - 02. Website tool

River Reach - 07. Text Fields tool

Access Site - 7. Text Fields tool

Campgrounds - 2. Text Fields tool

The tools that populate text fields are the simplest models. There are two text fields tools in the River Reach toolbox, one for website URLs and one that populates seven free form text fields: description, upstream point, downstream point, section name, management, address, and regulations. The tool "Access Site – 7. Text Fields" also populates seven free form text fields: directions, river name, restrictions, additional management, site name, and two web link fields. The tool "Campgrounds – 2. Text Fields" populates four free form text fields: notes, website link, management, and campground name. Figure 35 shows an example of the dialog boxes for this type of tool. These tools differ from the other two types in that they use the field calculator instead of the SQL expression builder to allow users to identify which fields in their data contain the information relevant to the corresponding NRRD field. The tool iterates through each of the fields in much the same manner as the previous tools, except it does not make a selection, but simply calculates the value of the given field, as seen in Figure 36.

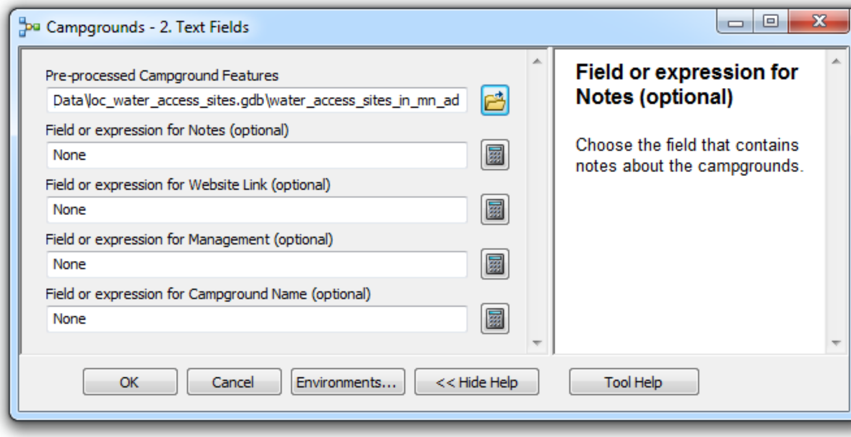


Figure 35: Example of a Text Fields tool dialog box

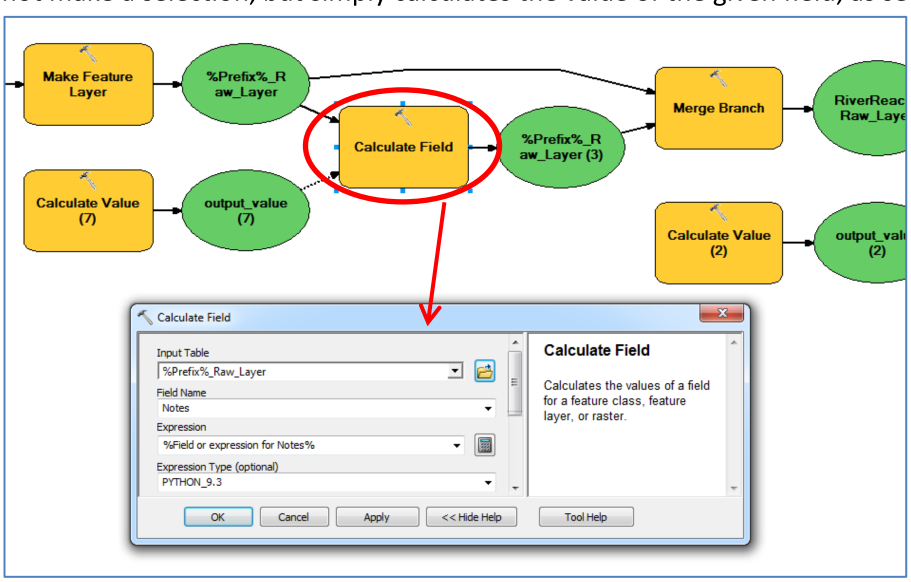


Figure 36: One iteration of the model for the Text Fields tools.

3.4.6 Model wrap-up

All of the models conclude in the same way. After the fields have been populated and the final iteration has run the "Merge Branch" tool, the "Select Layer by Attribute" tool is run one more time to clear the last selection (this step is missing from the text fields tools since there would be no selection to clear). Then the final layer is copied using the "Feature Class to Feature Class" tool into the "NRRDgdb_Stages" geodatabase, and that feature class is added to the map document. Figure 37 shows this part of the model.

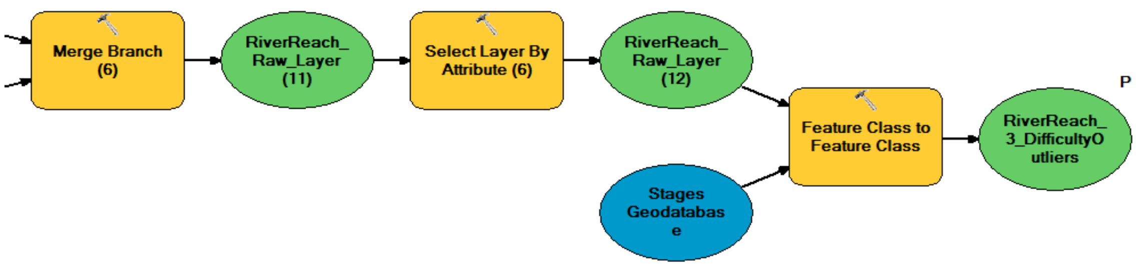


Figure 37: Tail end of each model, creating the final output

The tools are set up so that they can be run (or not) in any order. The key is that in each tool box, each tool adds fields to the output of the previous tool, so it is important to select as input the output of the most recent tool used. The model tools are numbered to make this easier to track.

3.4.7 State tools

Access Site - 1. State tool

River Reach - 05. State tool

The state tools are a special case because they have a coded value domain with 52 values (50 states plus Washington D.C. and Puerto Rico), which makes the method used for other coded value domains impractical. Instead, the tools were written using a Python script. The dialog box for "Access Site – 1. State" is shown in Figure 38. The user has

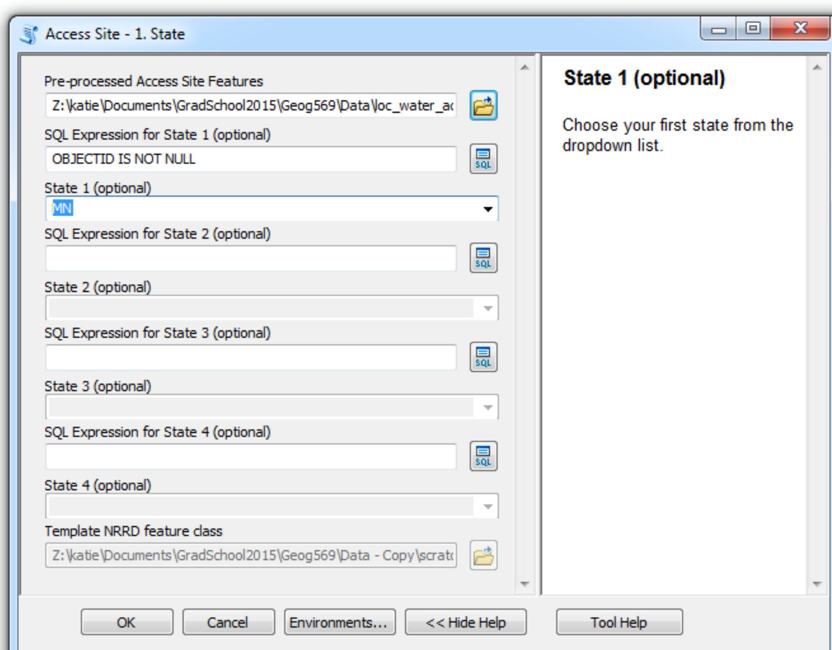


Figure 38: Dialog box for the Access Sites state tool

the option to designate up to four states (this was deemed adequate for most users). After choosing the input feature class, the box "SQL Expression for State 1" is automatically populated with the expression "OBJECTID IS NOT NULL," which for most users would select all records (since all records should have an Objectid). Therefore, if all the user's access sites are in one state, they need only choose the state from the dropdown box. Notice that the boxes for states 2, 3, and 4 are grayed out. They become active only if an SQL expression is entered in the preceding box indicating that some of the access sites are located in a second (or third or fourth) state.

The River Reach state tool ("River Reach – 5. State") populates up to two state fields to accommodate river reaches that cross state boundaries. Since most river reaches will be in only one state, the dialog box for the State 2 field is hidden unless specifically activated by clicking on the drop-down arrow circled in Figure 39. The State 2 field can hold up to four states as well (for example, if the dataset contains river reaches that start in Idaho and end in either Oregon or Washington).

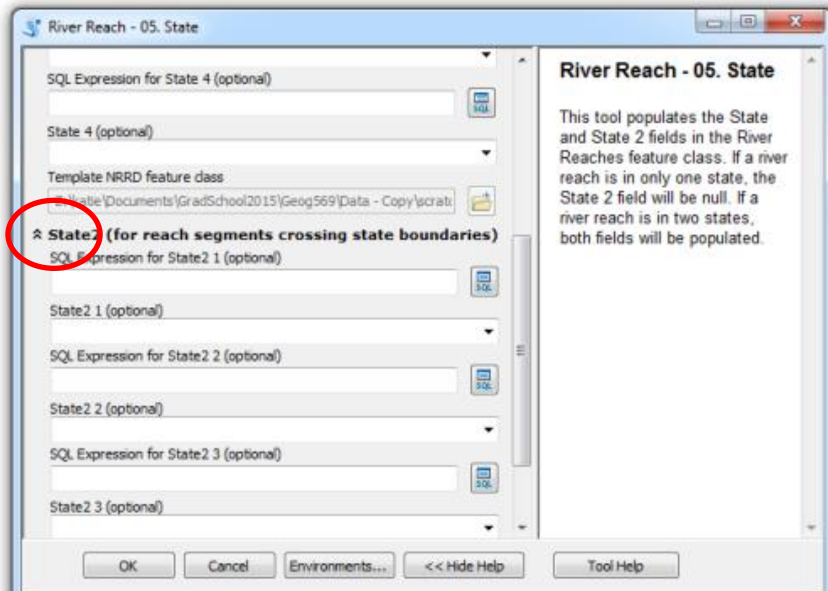


Figure 39: Dialog box for the River Reaches state tool

To make the tools, a new script was added to the toolbox, and parameters equal to each of the boxes in the dialog were added to the script properties. For the "State #" boxes, a value list with each of the possible codes in the domain was added, providing a drop-down list for users to choose from. This is shown in Figure 40. The Python script was then added to the "Validation" tab. The script iterates through the possible values in a manner similar to the way the Model Builder tools work, but much more efficiently. In addition, scripting in Python allowed the addition of a few bells and

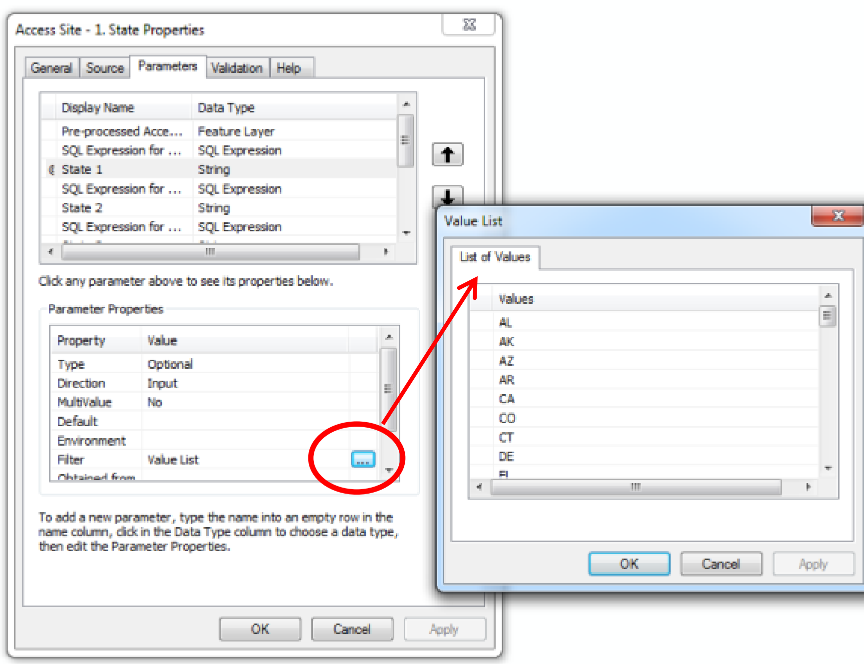


Figure 40: State tool properties showing the list of State values

whistles, including the grayed-out State boxes and the hidden boxes for the State 2 field in the River Reaches tool.

3.4.8 Final Append

*Access Site - 8. Final Append **REQUIRED***

*Campgrounds - 5. Final Append **REQUIRED***

*River Reach - 10. Final Append **REQUIRED***

In each of the three toolboxes, the last tool that needs to be run is called "Final Append." If even one of the tools in a toolbox is used, this tool must be run to complete the workflow. The dialog box for these tools has only one box: the user must simply point to the output of the last tool run for that toolbox. This tool takes the output, strips it of all extraneous fields, and copies it to a new feature class in the "NRRDgdb_Final.gdb" geodatabase. This last geodatabase holds the completed, transformed data, and should be submitted to the NRRD for approval. Figure 41 and Figure 42 show the model and dialog box for the Final Append tools.

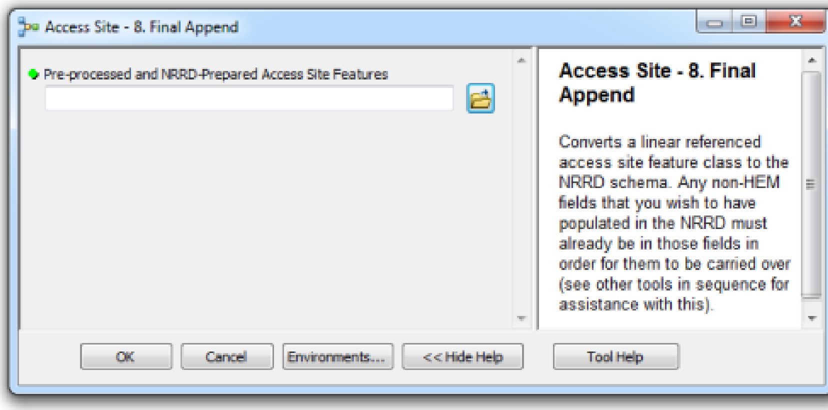


Figure 42: Final Append tool dialog box

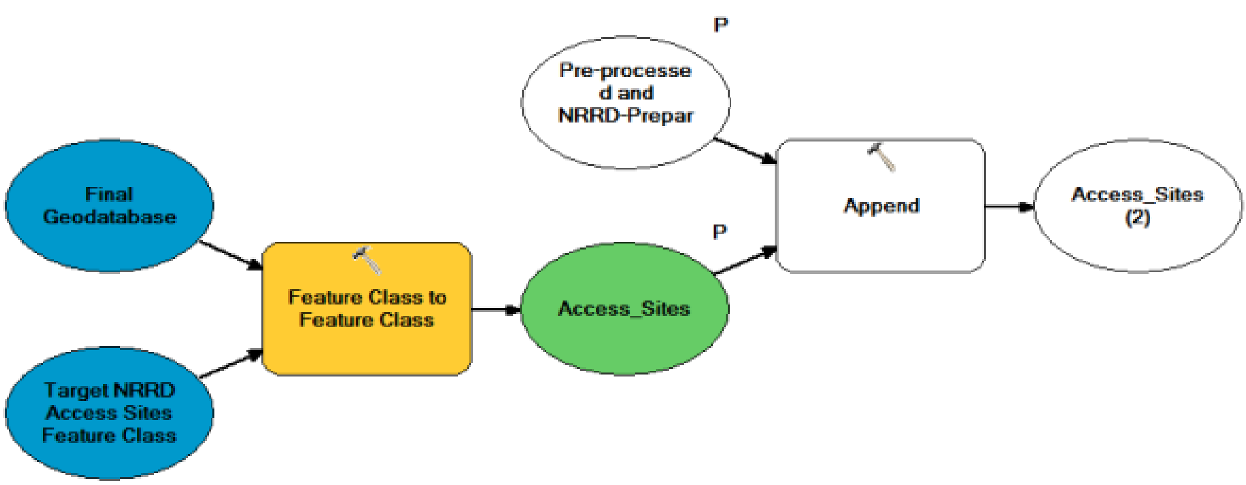


Figure 41: Final Append tool model

3.5 NRRD Data Submission Cookbook

In order to make the NRRD data submission process as straightforward as possible, an instruction manual known as the "Cookbook" was prepared. The Cookbook steps through every stage of the process with clear instructions for how to proceed. Flowchart decision diagrams are inserted at key points to aid the user in identifying which steps/tools are needed for their particular situation. The Cookbook is presented in its entirety in Appendix B. Cookbook contents include the following.

3.5.1 Part A: Introduction

This section introduces the Cookbook and discusses required tools and skills needed to successfully prepare data for submission to the NRRD. Minimum skills and requirements include:

1. Access to ArcGIS version 10.2 and/or 10.3
2. Ability to navigate to folders and geodatabases in ArcCatalog
3. Ability to use tools in ArcToolbox
4. Ability to edit line and point features in an ArcMap editing session
5. Ability to use the field calculator to set the content of fields in attribute tables
6. Ability to compose SQL expressions in a selection dialog box

3.5.2 Part B: Initial Data Preparation

This short chapter walks the user through how to download National Hydrography Dataset (NHD) flowline feature class for their area of interest, then how to re-project data to match the NHD geographic coordinate system. This first step is necessary because all donor data must be linear referenced to the NHD flowlines.

3.5.3 Part C: Preprocessing Point and Line Data

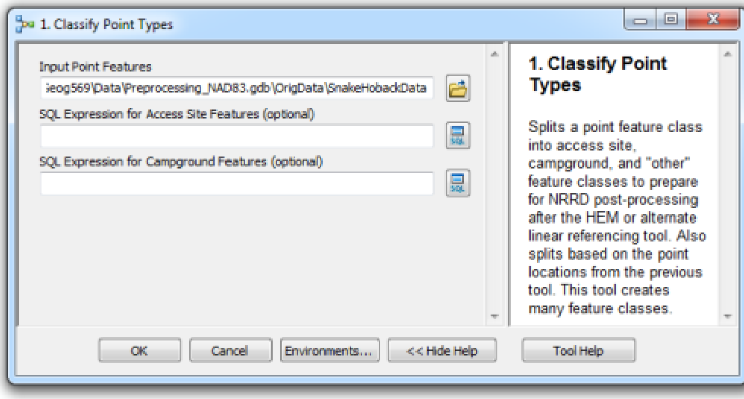
This chapter steps through the preprocessing tools in the NRRD preprocessing toolkit. It walks the user through how to use each tool to prepare their data for linear referencing, with the end result being up to four feature classes in the necessary format for referencing. Figure 43 gives an example of the instructions for one of the preprocessing tools.

IF YOU HAVE MULTIPLE POINT FEATURE CLASSES CONTAINING DATA ON ACCESS SITES AND CAMPGROUNDS, YOU SHOULD RUN TOOLS 1 AND 2 ON EACH OF THOSE FEATURE CLASSES.

1. *Classify Point Types tool*

This tool separates your point features into three categories: access sites, campgrounds and "other." The NRRD only makes use of access site and campground data, so other types of point data (such as river miles) will not be retained. Note that rapids will be dealt with in tool 4, "Join Rapids to River Reaches." This tool adds a field identifying each feature as one of the three point types. THIS TOOL ALSO ADDS FIELDS TO YOUR POINT FEATURE CLASS THAT ARE NECESSARY FOR THE FUNCTIONING OF FUTURE TOOLS, SO YOU NEED TO RUN THIS TOOL EVEN IF YOU HAVE ONLY ONE TYPE OF POINT IN YOUR DATASET (E.G. ONLY ACCESS SITES).

1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "1. Classify Point Types."



2. For "Input Point Features," choose your point feature class from the dropdown menu or navigate to its location. *Tip: make sure your feature class is in the "orig" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase.*
3. In the next two boxes, click the SQL icon to the right of each box and create an expression that indicates which point features correspond to each of the NRRD feature types. If you have no features of a particular type, leave that box blank. *Tip: if you have only one type of point feature, use the IS NOT NULL buttons in the SQL dialog box to select all records (e.g. Access_Site_Name IS NOT NULL).*
4. The output of this tool is three feature classes in the "Preprocessing_NAD83.gdb" geodatabase. One, prefixed "AccessPoint_NRRD_Raw…," is in the "Raw_AccessPoints" feature dataset. Another, prefixed "Campground_NRRD_Raw…," is in the "Raw_Campgrounds" feature dataset. The third, prefixed "Unclassified_NRRD_Raw" is in the "Raw_Unclassified" dataset. You will use these as inputs to the tools that follow. The tool should also add these feature classes to your map document, so you should be able to access them from the dropdown list.

Figure 43: Example of instructions for a pre-processsing tool

3.5.4 Part D: Linear Referencing

This chapter provides information on the two possible linear referencing tools and then walks the user step by step through the process of using their chosen tool option. It begins by detailing everything needed to run the tools on point and line feature classes, as seen in Figure 44. Table 3 compares the two tool options.

D. Linear Referencing

At this point all of your data is prepped and ready to run through the linear referencing tool of your choice.

For line feature classes:

1. If you already use NHD flowlines for your river reach data, or if your river reach geometry is extremely close to the NHD flowline geometry, or if you do not have access to ArcGIS 10.2, you may opt to use the Alternate Linear Referencing tool, which is a pared-down version of the HEM tool and generally runs faster. However, you may have to update some fields in your data by hand. If your line data does not closely match the NHD and you use the Alternate tool, you may have many features to edit by hand, so budget time for that.
2. If your river reach geometry does not correspond closely to the the NHD flowline geometry, and if you have access to ArcGIS 10.2, we recommend that you choose to use the Hydrography Event Management (HEM) tool, which is more thorough and has a built in process for updating all fields in the data.
3. Run your chosen linear referencing tool on "LinRefPrep_RiverReach," located in the "PreppedData" feature dataset.

For point feature classes:

You will need to run a linear referencing tool on the following feature classes (if you have them) from the "PreppedData" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase:

1. "LinRefPrep_AccessPoint_OnRiver" and "LinRefPrep_AccessPoint_EdgeOfRiver."
2. "LinRefPrep_Campground_OnRiver" and "LinRefPrep_Campground_EdgeOfRiver."
3. Note you do NOT need to run it on the feature classes with the suffix "_OffRiver."

Figure 44: Introduction to linear referencing

	HEM tool	Alternate tool
Where to get the tool	download from USGS website	included with NRRD toolkit
ArcGIS version	10.1 or 10.2	10.3
Tool type	ArcGIS toolbar	tool in NRRD toolbox
Tool speed	slow to very slow	fast to slow
Complexity	complex	fairly simple
QA/QC	built-in	manual, time consuming
Output	identical	

Table 3: Comparison between the two linear referencing tool options

3.5.5 Part E: Post Linear Referencing Clean Up Tools

This part of the Cookbook instructs users in how to restore the original geometry of edge-of-river point features using the "Restore Edge of River Points" tool, and then recombine all point and line data into three feature classes with data that corresponds to each of the three NRRD feature classes of Access Sites, Campgrounds and River Reaches, respectively, using the "Combine Linear Referenced Feature Classes" tool.

3.5.6 Part F: NRRD Data Import Toolset

The bulk of the Cookbook walks the user step by step through the three toolboxes in the NRRD toolset, one each for Access Sites, Campgrounds and River Reaches. Each subsection begins with a table, such as Table 4, that shows the schema of the NRRD feature class and describes the data that belongs in each field.

FieldName	Type	Length	Description
Name	String	75	Name of campground
Managing_Entity	String	50	Managing Entity
Website	String	250	Website for campground
Fee	SmallInteger	2	Yes/No
Type	String	100	Developed or Dispersed
Notes	String	300	Notes

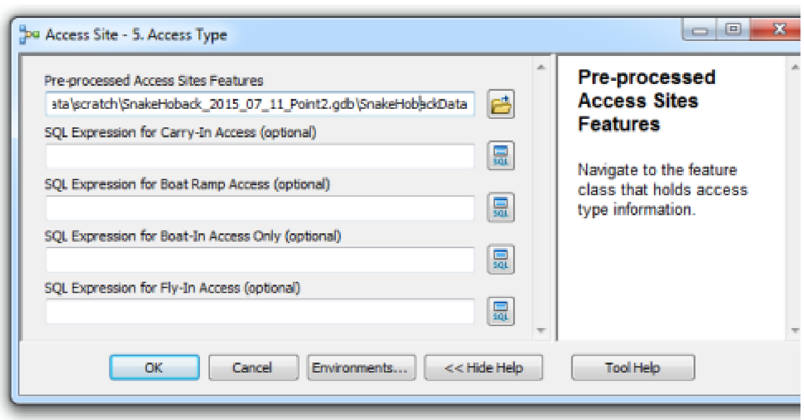
Table 4: Example of tables describing fields in each feature class

The Cookbook then goes through each tool in detail, instructing the user on how to fill out the dialog box, and is accompanied by a screenshot of the tool. Figure 45 shows an example of the instructions for one tool. This section also features workflow diagrams to aid the user in choosing the relevant tools for their data submission.

Once the user has walked through all of the relevant tools in Part F, the donor's data will be in its final geodatabase and ready for submission.

Access Site - 5. Access Type tool

Use this tool to designate access site type as "Carry-In," "Boat Ramp," "Boat-In," and/or "Fly-In."



1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 5. Access Type." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "AccessSite_4_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the access type indicated. If you have no sites with a given access type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Figure 45: Example of instructions for a post-processing tool

4. Results

This section describes the results of our design for a toolkit that will aid data donors in preparing their data for integration into the National River Recreation Database (NRRD). The original goal was to create a set of tools that would translate donor data into the schema of the NRRD after the donor had independently run their data through the Hydrography Event Mangement (HEM) tool. However, while working through the process it became clear that an additional set of tools was needed to prepare data for linear referencing and to organize linear referenced data into the format needed to run the NRRD tools. Finally, the complex nature of the HEM tool and a perceived resistance to its use by some potential donors inspired the creation of an Alternate Linear Referencing tool, which is far simpler in structure, but which requires a commensurate increase in attention to detail when conducting quality control checks on the linear referenced data.

Project sponsors provided the team with two sample datasets to work with while developing the tools: a small data set from the Snake and Hoback Rivers in the Bridger-Teton National Forest, and a much larger state-wide data set of water trails from the Minnesota Department of Natural Resources. The team generally chose to use the Snake-Hoback data to test various parts of the process because of its smaller size and therefore increased processing speed. However, attempts were made to include the Minnesota data whenever possible.

All of the tools work as designed. The Snake-Hoback data was successfully run through all of the tools, including the Alternate Linear Referencing tool, resulting in a set of three feature classes that can be imported into the NRRD. However, the Minnesota data posed some extra problems. The 156 Minnesota water trail line features were hand-drawn by the Minnesota DNR from aerial photographs, making their line features more accurate than the National Hydrography Dataset (NHD), to which all lines are required to be linear referenced. Likely as a result of these discrepancies, the HEM tool ran for over 30 hours to process the Minnesota line data and resulted in an error-ridden output with hundreds

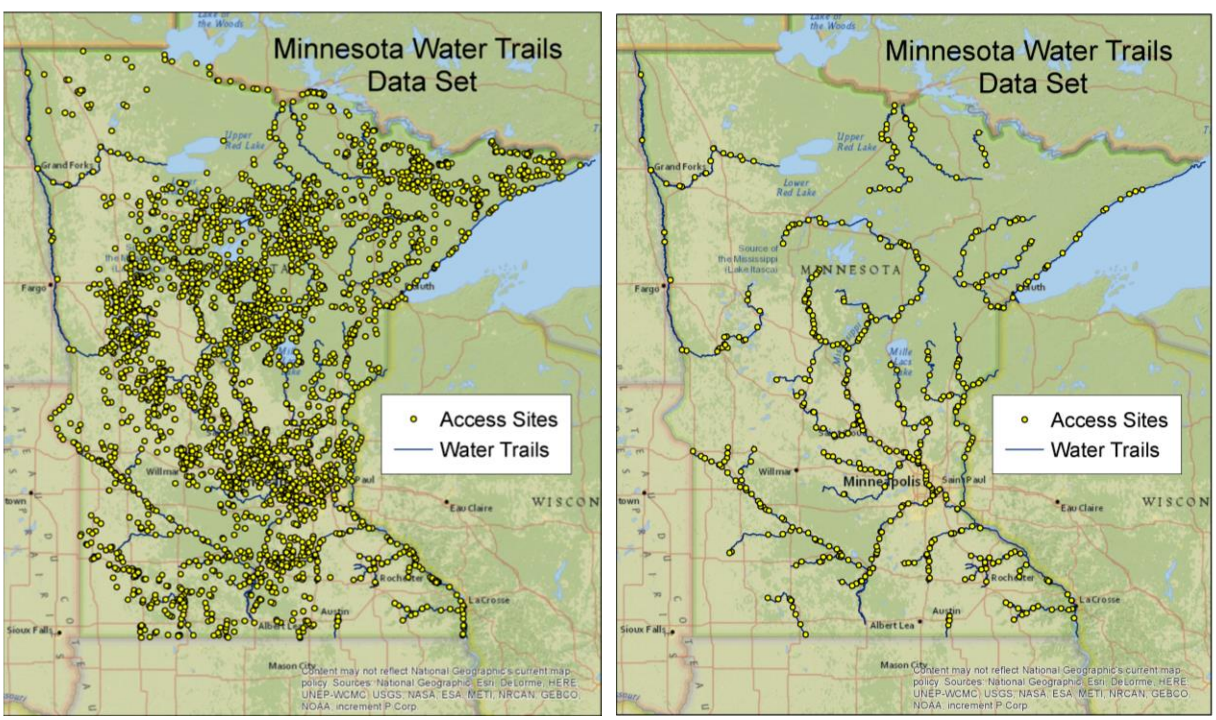


Figure 46: Minnesota data before and after running it through the preprocessing tools

of lines in need of manual editing during the quality control process. In addition, the Minnesota data stewards expressed a strong preference for maintaining their original, more accurate line geometry rather than having it snapped to the NHD. After consulting with the project sponsors, it was decided that in some cases the line data need not be linear referenced before adding it to the NRRD. This will result in more limited functionality of the NRRD for those line features, but removes what may otherwise have been an insurmountable barrier to donors' willingness to donate data. Figure 46 shows the Minnesota data before and after running it through the preprocessing tools. Only those point features that were associated with water trails carried over for the rest of the data submission process. Since the Minnesota data did not include campgrounds, the final output was two feature classes prepared for integration into the NRRD.

The final project deliverables include the following, as seen in the catalog tree screenshot in Figure 47.

1. Empty geodatabases to hold the intermediate quality control feature classes created by the linear referencing tools
2. A folder called "LayerFiles" that holds the layer file needed for the QC process on the Alternative Linear Referencing tool
3. A "scratch" folder of geodatabases to hold the template, intermediate, and final data sets when running the NRRD tools
4. A folder that holds all the tool scripts
5. A geodatabase called "NRRD_Begin_Postprocessing" that holds the final output of the preprocessing tools
6. A geodatabase called "preprocessing_NAD83" that holds the original and intermediate feature classes created by the preprocessing tools
7. A set of preprocessing tools, in the NRRD_Preprocessing.tbx ArcGIS toolbox
8. Three ArcGIS toolboxes with tool sets for converting donor attribute data to the NRRD schema

The contents of the toolboxes are discussed in detail in the previous section.

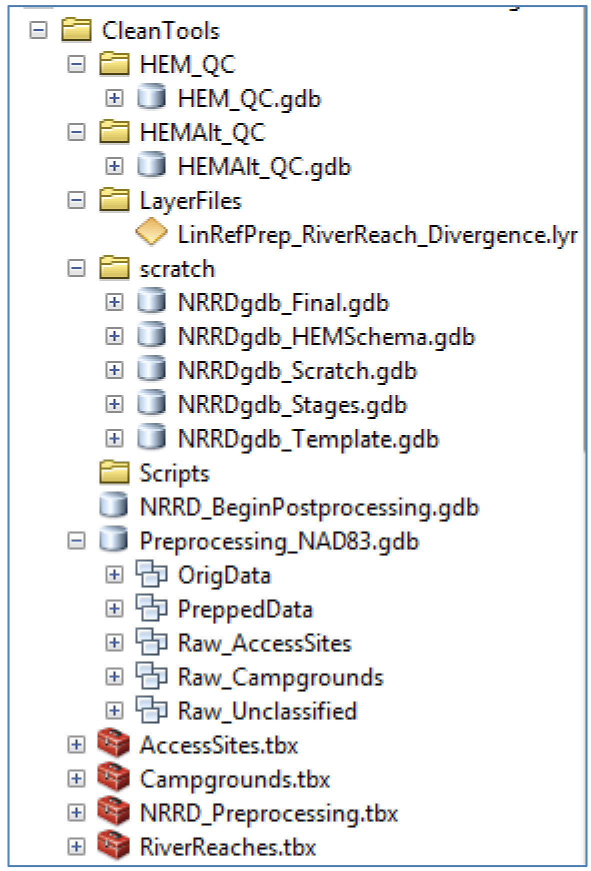


Figure 47: The complete NRRD toolkit

5. Discussion

The toolboxes and "cookbook" instruction manual created during this project attempt to provide the "data integration process in which we have confidence" that the National Park Service and River Management Society described in their initial proposal (Rosebrough, 2015). More than 20 custom tools and a 40 page cookbook walk potential data donors through three phases of data preparation for the National River Recreation Database (NRRD): preprocessing, linear referencing, and post-processing. This workflow was tested on two sample donor datasets, a smaller (6 river reaches) one from the Bridger-Teton National Forest in Wyoming and a larger (156 river reaches) one from the State of Minnesota Department of Natural Resources. More important than the tools created, however, is the framework provided for changing tools and creating more tools of similar styles, which we hope will help the tools and documentation evolve alongside the NRRD throughout its lifespan to provide lasting benefits to NPS, RMS, donor agencies, and the public. This section covers an assessment of the state of the tools at the end of the nine-week project, justifications for particular decisions, challenges encountered, and recommended next steps.

5.1 Conclusions

This section will assess the effectiveness of the current toolbox and tutorial in providing support in each of the three phases of data preparation for the NRRD: preprocessing, linear referencing, and post-processing. The tools and workflow themselves are covered in more detail in the Design and Methods section. In nearly all cases, working with the much smaller (6 river reaches vs. 156) Bridger-Teton National Forest data produced better results than working with the State of Minnesota data. This is to be expected given the wider variation in data from the State of Minnesota and the error-prone nature of linear referencing.

5.1.1 Assessment of Preprocessing Tools

The preprocessing tools were designed to quickly address situations that could cause significant hang-ups with either of the later two phases. Two of the issues addressed immediately are coordinate systems and field name conflicts. The first step of the workflow has donors project their datasets into the North American Datum of 1983 (NAD83) coordinate system. This is the same coordinate system as both the National Hydrography Dataset (NHD), used in linear referencing, and the NRRD itself. This was not an issue with either test dataset as both were in NAD83-based coordinate systems already. Most datasets in the United States will likely be in a NAD83-based coordinate system or one of its derivatives (e.g. NAD83 HARN) and so will have little trouble with this step. If needing to change datums, donor GIS specialists will need to know which geographic transformation to use.

Potential field name conflicts are dealt with by prefixing original attribute fields with "DONOR_". As no NRRD fields, present or future, are likely to begin with this prefix, it represents probably the simplest solution to the problem. However, because versions back to ArcGIS 10.1 are being supported, the Alter Field tool, introduced in ArcGIS 10.2, could not be used to change the names of fields in-place. Instead, copies of the original fields are added with the prefix, populated with the same values, and then the original fields are deleted. Several issues were encountered with this that donors may run into. First, any field that has been designated as "Required" by the donor cannot be deleted and so remains in the table and may cause name conflicts later. It may be possible to solve this issue by passing through a feature layer, where one can designate only selected fields to be saved. Secondly, if using a subtype field, only the codes (integers) will be brought over, not the descriptions, because there can only be one subtype field defined at once. A subtype field is also required, so this is one type of original field that cannot be deleted using Delete Field. This was an issue encountered with the Bridger-Teton National Forest point data.

Point data classification in preprocessing proceeds by using SQL queries to give each point a type (access site, campground, etc.) and generic location (on-river, edge-of-river, off-river). This begins a theme that will be common throughout almost all the tools, especially in post-processing: SQL queries. While Structured Query Language can be clumsy and error-prone, its ubiquity and use in several key ArcGIS tools, including Select and Select Layer By Attribute made them an indispensable part of this workflow. Figure 48 shows an example SQL query on Bridger-Teton National Forest point data.

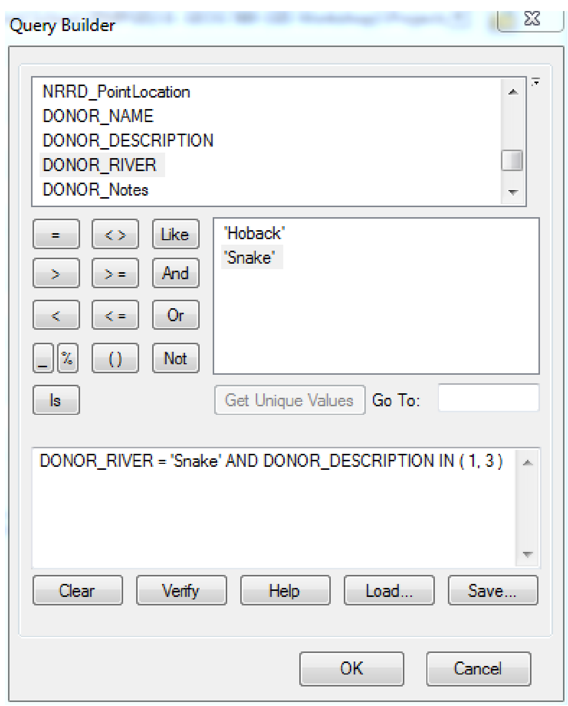


Figure 48: An example SQL query, selecting boat launches (1) and emergency access sites (3) along the Snake River in the Bridger-Teton National Forest point data. SQL parameters can be difficult to write and are error-prone, but are common in many of the NRRD tools as the most succinct way of selecting certain features to perform an operation on.

As a result, a GIS specialist comfortable with SQL queries is likely an indispensable participant in an organization's data donation effort to the NRRD.

The generic point locations of on-river, edge-of-river, and off-river represent the current level of allowing donors to determine how much their features are moved during linear referencing. Data donors tend to be more sensitive about having the geometry of their features changed than having the attributes changed, so this step helps alleviate some of that trepidation. Eventually, more categories of linear referencing methods might crop up, forcing this particular step to change to accommodate that.

One final aspect of the preprocessing tools worth mentioning is the joining of rapids difficulty (from points) to river reaches. This actually represents a first attempt at addressing a much larger issue, attribute representation in different geometry types. Donors could easily have attributes such as stream difficulty or emergency access represented as points or lines, where the NRRD has it the opposite. With a vector representation of space (discrete points, lines, etc.), figuring out which feature in the new geometry type to transfer this attribute to can be difficult. The Join Rapids to River Reaches tool uses ArcToolbox's Spatial Join tool to connect rapids (points) to their nearest river reach (line), then calculates the highest difficulty rapid and saves only that value to a new attribute field in river reaches. This may present problems if rapids are present that are not along floatable river reaches, as these would still be joined to the nearest floatable river reach. A maximum distance threshold parameter might help alleviate this issue. Overall, however, data donors will still have to be very aware of the state of their own data when performing an operation like this.

Overall, the preprocessing tools attempt to perform minimal alterations to the donor feature classes while achieving maximum resolution of potential errors later in the workflow. New information is added in new fields; original fields are maintained with "DONOR_" prefixes; and feature classes are grouped into their respective NRRD destinations. The preprocessing tools tend to be more complex and tailored to a specific situation than the post-processing tools. They also are less structured as they were built in a more reactionary fashion than the planned and organized post-processing tools. In time, as more donor organizations try out the process, more situations requiring preprocessing tools might be encountered, which could lead to a more standardized structured (e.g. additional tools like Join Rapids to River Reaches). Existing preprocessing tools will also probably need less frequent editing than post-processing tools due to small changes in the NRRD schema, but might require less frequent, larger edits when a change is made that affects that tool's particular issue.

5.1.2 Assessment of Linear Referencing

The linear referencing phase of the NRRD submission workflow is probably the phase with the most difficulties and shortcomings at the time of writing. Most critically for donors, the HEM alternative tools do not perform well enough to inspire confidence if the HEM tool is not an option for an organization. The options for linear referencing tools in this workflow are described in detail in the Design and Methods section. In the end, linear referencing, especially for line features, is a difficult task that is not even completely solved by using a custom add-in (the HEM tool) requiring many hours of development. Instead, close collaboration with data donors, including looking at their specific data, by GIS specialists very familiar with the HEM tool is likely to lead to the best results for the time invested. This is described in more detail in the Business Case and Implementation Plan section.

The HEM tool performs very well given a trained user. Training is thus critical to giving potential data donors the best experience with this phase of the workflow. This project attempts to address this issue somewhat by providing a very careful, step-by-step instruction manual (in the cookbook) on exactly which HEM functions to use when in order to submit data to the NRRD. This cuts down on having to locate and test each function as much. When the HEM runs well, it is by far the most recommended option for linear referencing, especially for lines. Unfortunately, development on the HEM tool, as a third party add-in from USGS, is usually having to play “catch-up” to the most recent ArcGIS version. At the time of writing, ArcGIS 10.3 does not have a working HEM version for it, meaning organizations that have upgraded quickly will find themselves having to use the less functional HEM alternative tools.

However, the HEM still runs into considerable difficulties working with river reaches not already aligned with the NHD. This is the Minnesota dataset. Where the 6-river reach Bridger-Teton National Forest dataset referenced within minutes, the 156-river reach Minnesota dataset took an entire night, over 24 hours, on a newly-purchased high-end 2015 laptop running ArcGIS 10.2.2. This is a significant processing time challenge as organizations might perform nightly computer maintenance and updates, stopping any running tasks. Breaking the dataset into chunks capable of running during an 8-hour day would make it last at least 3 days. If any error occurs in the process, this could mean an entire week spent trying to perform this one operation – in fact, an error did occur during one test, which caused the output geodatabase to become corrupted immediately after the import completed its 24-hour processing. While most datasets will not be as large as the State of Minnesota river reach network, some will, and this will present difficulties, especially considering that the HEM alternative tools also perform worst on these types of datasets.

The HEM alternative tools are lacking several critical functions right now, but might represent a start for something resembling the HEM tool, but with more straightforward and user-friendly functions. With some organizations completely unable to use the HEM tool, developing a decent set of alternative functions is not a waste. Currently, the two HEM alternative tools only replicate the import function of the HEM tool. There need to be at least two additional tools to update Reachcode and Measure attributes for points and lines that are manually edited after the first import tool is run. This would give the HEM alternative tools the bare minimum functionality needed to replicate the HEM for the NRRD submission workflow. Currently, the point import tool also completely leaves out points that are outside the reference tolerance, requiring donors to rerun the tool with a higher tolerance to keep all points. Instead, the tool (like the HEM) should still copy points to the output even if they have not been moved. Finally, the greatest challenge is line linear referencing due to ArcGIS’s Locate Features Along Routes tool’s returning lines in a separated format. This either forces donors to do some significant QA/QC on their datasets to remove gaps and stubs or forces the developer to code in some “judgment calls” about which features are stubs vs. legitimate and which gaps are errors vs. legitimate. Currently, two judgment calls exist in the HEM Alternative (Line Import) tool, but this part of the process will likely never achieve the accuracy of the HEM tool itself. Like the HEM tool, the HEM Alternative (Line Import) tool works best when lines are already along the NHD so the reference tolerance can be kept low. This cuts down on stubs. Ultimately, however, the HEM tool is the only solution in cases where the alternative tools fail.

Linear referencing is a necessary step in the NRRD submission workflow that was determined repeatedly to be worth doing. Successful linear referencing means that the whole NRRD is using the same stream network, cutting down on errors and mismatch between different jurisdictions. This process, especially for river reaches, will probably always be imperfect, so it is important that donor GIS specialists are familiarized with the HEM tool and/or alternatives and supported throughout this often frustration step. However, in cases where the donor stream network is of higher detail than the NHD, this can also be an opportunity to improve the NHD, so NPS and RMS should facilitate collaboration between donor organizations, the USGS, and state NHD stewards in these cases. This could represent a longer-term solution to the issues of linear referencing than merely requiring donors to run it.

5.1.3 Assessment of Post-Processing Tools

The post-processing tools of the NRRD submission toolkit comprise the most structured and well-documented portion of the workflow. While they are not without considerations, the post-processing tools represent a good end state for each of the other two phases to strive for in terms of clarity, consistency, sequence, and documentation. These tools add and populate fields from the final NRRD schema that have not already been calculated during preprocessing and linear referencing. These include fields like "Access_Type," "Description," and "Website". The post-processing tools are described in more detail in the Design and Methods section.

The strength of the post-processing tools lies in their organization and planning based on the NRRD schema. In making the post-processing tools, all of the NRRD-specific attribute fields (those not calculated by the HEM tool) were split into three categories: fields with coded value domains (a set list of allowed values), freeform fields, and unclassified fields or those with potential other considerations. This was done for river reaches, access sites, and campgrounds to produce a roster of fields that needed to be populated. Three different types of tools were then developed that catered to each of the three types of fields. Fields with coded value domains use SQL queries to allow donors to select which features get which values from the list. Freeform fields allow donors to set Field Calculator expressions to take values from one or multiple other fields. Fields with additional considerations had tools made specifically for them. This structure gives the post-processing tools flexibility to respond quickly to changes in the NRRD schema. If a new field is added, the most similar tool can be copied and tweaked to produce the new tool. The tools are in ModelBuilder for the most part, meaning that editing is less error-prone than Python.

The tools are also ordered intentionally to help donors work their way through the tools and know which dataset is supposed to be the input to the next. In addition, unlike the preprocessing tools, the post-processing tools do not require every tool to be run, only the ones that the donors actually have data for. This can save donors significant time and headaches.

The post-processing tools do run into the same issues as any data conversion workflow, namely that they either force too much detail onto donor data or that they neglect other data that does not fit into the NRRD schema. This is no fault of the tools themselves though, but rather just puts the impetus on the NRRD schema itself for river reaches, access sites, and campgrounds to include a useful mix of detailed attributes. The tools themselves can quickly respond when this schema changes.

The final append tools in post-processing drop the “DONOR_” fields and only preserve the fields added by the HEM tool or alternative and the fields added during post-processing. One gap right now in this last part of the workflow is a metadata tool which would populate fields like “Source_Originator”, the name of the donor organization. This final tool would help the NRRD team at NPS and RMS keep track of which datasets belong to which organization.

5.2 Justification for Decisions

This section covers justifications for several decisions made throughout the workflow that have not already been discussed. These decisions include working within the ArcGIS framework instead of open-source and providing a downloadable, local toolbox to users as opposed to an online service or standalone script, and requiring linear referencing (in most cases).

Developing and deploying the NRRD data submission tools solely in ESRI’s proprietary ArcGIS framework was done with awareness that this would immediately neglect donor organizations using an open source solution. This decision was made for many reasons. During the first meeting with project sponsors, it became clear that NPS was pursuing an ESRI solution exclusively for the NRRD for the time being. The version of the NRRD that was shown there (completion of NRRD phase I) was also running as an ESRI web map on ArcGIS Online. More importantly than this, however, was that the USGS’s Hydrography Event Management (HEM) Tool, used for linear referencing to the National Hydrography Dataset (NHD), requires ArcGIS. While it was revisited throughout the project, aligning donor features to the NHD was considered a central step in the NRRD submission process, partly due to strengthening the partnership with the USGS’s Hydrography team. As a whole, this was taken as enough justification to focus efforts on an ArcGIS solution, especially given a nine-week timeframe for development. Efforts were instead put into best realizing the benefits of an ESRI solution. Several of these benefits are described below.

One of the greatest benefits of an ESRI solution is the highly functional and well-documented environment for developing custom tools similar to those found in ArcToolbox. Tools can be developed in either the graphical interface of ModelBuilder or in Python code and still take advantage of ArcGIS’s automatic validation of inputs and outputs. Tools themselves can be quickly copied and modified to create similar tools, sometimes within minutes, and a robust metadata interface is provided for documenting each tool in FGDC or ISO geospatial metadata standards. The performance gains of working in this environment enabled the creation of more tools and functionality than would have otherwise been possible, and this benefit will extend to any future tools needed, as will be described in the Business Case and Implementation Plan section.

An ArcGIS toolbox format also has several advantages for users of the tools (donor organizations) compared to standalone scripts and other solutions for performing geoprocessing tasks. Zandbergen provides a complete list of these benefits, but some are restated here more specifically (Zandbergen, 2013). Users of the NRRD data submission tools in ArcGIS will not have to install special software or interact with any code directly. Instead, the tools will be immediately compatible with all versions of ArcGIS 10.1 and higher (10.3.1 is the highest as of this writing). Data donors will be able to see the output of each step of the workflow in ArcMap, allowing for quick visual QA/QC of each step, or more robust analysis if needed. Users can quickly make use of any other geoprocessing tools available in ESRI's ArcToolbox if they decide to add extra steps, making the workflow more responsive to different data donors. Finally, the tools in the NRRD data submission toolboxes can themselves be added to ESRI's ModelBuilder or called from Python scripts. This enables users to automate the NRRD data submission workflow once they have run through it once, yielding significant time savings for subsequent "update" submissions.
Figure 49 shows the first three preprocessing tools after being chained within ModelBuilder.

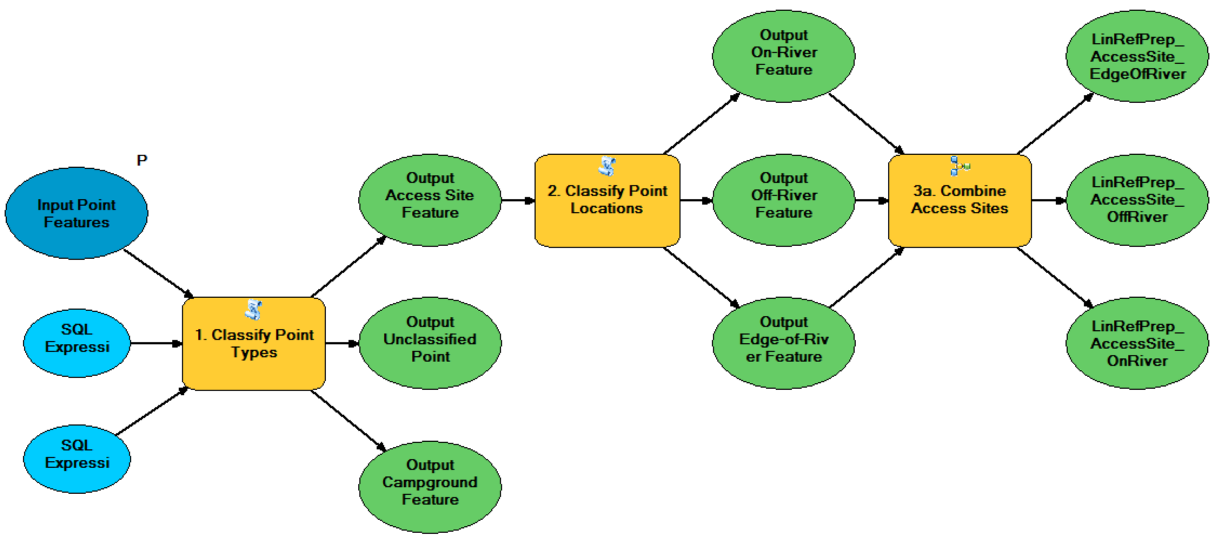


Figure 49: The preprocessing and post-processing tools can all be added to ModelBuilder themselves. This allows donors to chain and standardize their workflow after they have established it the first time.

GIS professionals working with donor organizations will thus have a built-in way to standardize their submission workflow, leading to more frequent updates and a more justifiable cost-benefit tradeoff for making their initial submission to the NRRD.

The decision to develop a set of local toolboxes for use with ArcGIS for Desktop as opposed to an online geoprocessing service was made in order to provide a simpler, more transparent, and easier to update set of tools to donor organizations. Along with being more familiar to the team during the nine weeks of this project, the tools, especially those made in ModelBuilder, will be easier for donors to review for their own knowledge. This helps build trust with users, as an online geoprocessing service would not offer this ability to "peek under the hood" at the tool execution logic. This also gives users familiar with

Python or ModelBuilder the opportunity to customize the tools for their own organization if they see an opportunity to do so. Deploying the toolbox and associated folders in a simple ZIP file download also avoids the need for hosting a live geoprocessing service on the NPS servers, along with all the associated server maintenance required. Finally, a set of local tools will be easier to update as the additional task of optimizing for server performance is eliminated. However, users will have to re-download the tools after any updates. A final reason to keep the toolboxes local is that the data submission itself from a donor organization will just be a file geodatabase with the required feature classes. Rather than passing data to and from NPS servers, it makes more sense to keep data preparation local with donor organizations and then just submit the final geodatabase.

Requiring linear referencing to the NHD was a decision that was revisited throughtout the project and that changed as new situations were encountered. Initially, all rivers and points were to be referenced to the National Hydrography Dataset. This would help support spatial queries in the NRRD by ensuring that all points (e.g. access sites) were on top of river reaches. However, some points did not make sense to reference, especially in cases where donors felt that they had accurate latitude and longitude locations for their points. As a result, the Classify Point Locations tool was introduced to give donors the option of how significantly to linearly reference their points. River reaches were still required though to be referenced, however. This would help with some quality control and ensure that river segments matched across donor jurisdiction boundaries. However, due both to the difficulty of linearly referencing line features not already coincident with the NHD and an opportunity to improve the NHD, an exception was created for donor datasets that were higher in detail than the NHD. Instead of requiring them to be linearly referenced, they would be submitted in their original locations and the USGS and state NHD steward would be notified of an opportunity for improving the NHD using those features. This would offer a more long-term solution to the problem of river reaches that do not match the NHD.

5.3 Challenges Encountered

This section covers specific challenges encountered during the nine weeks developing this toolkit and cookbook tutorial. Focus will be put on challenges that have not been covered in detail yet. These include choosing between ModelBuilder and Python for individual tools, preserving domains through tools, and transitioning between ArcGIS version 10.2 and 10.3 development environments.

Custom ArcGIS tools can be created in both ModelBuilder and Python environments. In general, Python script tools allow greater flexibility and more complex functionality while ModelBuilder provides a graphical user interface and automatic handling of input and output schemas. ModelBuilder was the preferred development environment during the nine weeks for rapid prototyping and greater readability after giving the final tools back to the project sponsors. Almost every tool in the preprocessing and post-processing tools started out as a ModelBuilder tool. However, while ArcGIS provides a conversion function for when switching to the Python environment is required, this function does not capture all of the additional functionality that ModelBuilder automatically handles. This made for a significant investment every time that a tool needed to be switched to Python just to replace this functionality, which happened with almost every preprocessing tool and two of the post-processing tools.

As an example, ModelBuilder automatically deletes intermediate datasets from the scratch geodatabase with a simple option. In Python, this automatic deleting functionality had to be coded manually and changed slightly with each new tool. Figure 50 shows the difference between this functionality in ModelBuilder and Python.

Two particular situations required this conversion, which were encountered for almost every preprocessing tool. The first was running an operation on a set of attribute fields that is not known exactly until processing time (e.g. “calculate all fields that start with ‘DONOR_’”). As handling donor fields was something all preprocessing tools did, this function required almost every one be converted to Python. The second situation involved creating output datasets that were both dynamically named (original feature class as a suffix), but always in the same output location. This was the case with some of the preprocessing tools; ModelBuilder could not perform this function.

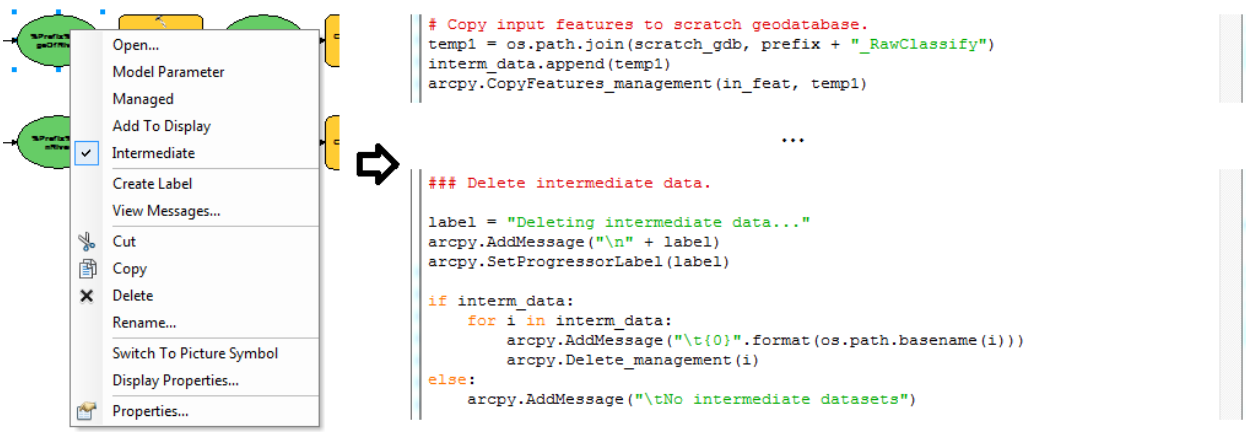


Figure 50: A significant time investment was required whenever a tool needed to be converted from ModelBuilder to Python. The Python code at right illustrate what was needed to replicate the automatic deleting of intermediate datasets in ModelBuilder.

A second challenge involved trying to preserve domains in fields after passing through many tools. In some of the post-processing tools, the newly added NRRD fields did not carry the domains with them, which could cause some confusion since only the codes (1, 2, 3, etc.) could be seen in the output of that tool. This was solved in the final append tools, but could still cause confusion during the workflow itself about whether the tool functioned correctly. A solution to this was found by copying the template feature class to the scratch geodatabase before joining the new field; joining a field from the same geodatabase preserves the domain while joining fields across geodatabases does not. This was not known until late in the nine weeks, so was not implemented in every tool. A second case of domain preservation issues was with original donor fields. Sometimes, they would not have their domains carried with them after linear referencing, leading to more difficult SQL queries in post-processing with only being able to see the code (1, 2, etc.) instead of the description (“Boat launch”, “Emergency access”, etc.). This issue extended to subtype fields as well. Figure 51 shows an example.

DESCRIPTION	DONOR_DESCRIPTION
River Access - No Ramp	10
Emergency Access	3
Boat Ramp	1

Figure 51: Preprocessing tools are not currently able to carry over the descriptions of subtypes in the "DONOR_" fields; only codes are carried over. This can cause difficulties for the donor later if they try to use these fields in SQL queries as the descriptions they are used to may not be available to them.

A third challenge involved compatibility of tools between ArcGIS version 10.2 and 10.3. Tools built in ModelBuilder in version 10.3 do not even show up if the same toolbox is opened in ArcGIS 10.2. Saving using the "10.2 Compatible" option also was not successful as, while the tools would show up, they would crash version 10.2 upon opening. To make matters more complicated, the HEM tool does not work in ArcGIS 10.3 at all. As a result, one of the two team members kept version 10.3 as it was required for their employment while the other used version 10.2. In addition, the team member with version 10.3 purchased a laptop to install version 10.2 to run the HEM tool as well. This meant that the full workflow could only be tested by the team member with both version 10.3 and 10.2 leading to longer times between tests. Development was geared towards supporting all ArcGIS versions 10.1 - 10.3.

5.4 Next Steps

This section covers the most pressing or immediate improvements that could be made to the NRRD data submission toolkit after returning it to the project sponsors at the National Park Service and River Management Society. These next steps mainly respond to the issues raised in the assessments of each phase of the workflow and the challenges presented in the previous section. The following section, Business Case and Implementation Plan, describes an ideal environment to make these improvements so that the toolkit and cookbook remain up-to-date along with the NRRD itself.

The main issue remaining with preprocessing tools is the treatment of coded value domain and subtype fields during the creation of "DONOR_" copy fields. The best approach may be to separate fields with domains into a code field and a description field (text) so that both are preserved later in the workflow. The Domain To Table tool in ArcToolbox can be used to accomplish this. This will help donors write SQL queries later. Subtype fields will require more work as they have to be "un-subtyped" to be deleted, but the end result should be similar to domain fields, with both a code and description output "DONOR_" field. However, subtypes enable users to set different domains on the same field (for different subtypes). This may compound the problem of expanding domain fields into two output fields, as each unique domain would have to be expanded separately. This is a fringe case, however, and likely will not affect many donors; and those it does affect will probably be experienced enough in ArcGIS to handle their fields without assistance. A longer-term task with the preprocessing tools is to formalize the transfer of attributes from points to lines (as with the Join Rapids to River Reaches tool) and vice versa. The easiest way to accomplish this may just be to instruct donors on how to use the Spatial Join tool along with perhaps a post-processing tool. This would enable them to join any set of attributes from one feature class to another and convert to the NRRD schema. Another short term task is to split the river reaches preprocessing tools into a sequence of three rather than two options of a single tool. Prep River Reaches should be first, followed by Join Rapids. This would enable more tools to be added to the river reaches workflow without having to copy functionality.

The most serious gap in the workflow currently is the lack of a HEM alternative update tool for line and point features. If any of the point or line features need to be moved manually by the donor after linear referencing, which is very likely with the HEM alternative tools, there is no way to automatically update the “Reachcode” and “Measure” attributes. This is half of the usefulness of linear referencing, so making these additional tools should be a priority if the HEM alternative tools are to be a legitimate option in the workflow. Luckily, there are large sections of the HEM alternative import tools that essentially already perform this function that can be copied over to separate tools, so this should not take much time. A second short task is to add functionality to the HEM Alternative (Point Import) tool to keep points in the output even if they fall outside of the reference tolerance. This will help donors not have to rerun the tool to catch all points. A much larger, long-term task is improving the HEM Alternative (Line Import) tool to more accurately linearly reference lines. This may never be done, however, so be cautious about how much time is spent developing these alternative tools beyond basic functionality versus training donors on how to use the HEM tool itself.

The post-processing tools are in the best shape out of the three phases. Most effort should be put into ensuring that they remain up-to-date with the NRRD schema, adding additional tools for additional fields. The field “roster” should be maintained with changes to the NRRD schema as well. One aspect that could be improved is the handling of freeform text fields. Minnesota access sites had several “DONOR_” fields that were longer than the destination NRRD fields, such as “Site_Name” and “Direction”. This required the creation of longer Field Calculator expression on the part of the donor to truncate after a certain number of characters. Using Calculate Value in ModelBuilder, however, would enable adding a section of Python code to each Field Calculator expression to do this automatically, saving the donor the extra effort and confusion. A small improvement could be made by using Copy Features on the Template NRRD Feature Class in each model to copy it to the scratch geodatabase before the Join Field. This would allow the transfer of domains for those fields also with the field itself, yielding more aesthetic output in the stages geodatabase. Finally, it will be essential to add a metadata tool to populate the fields “Source_Originator” (donor name), “Source_DataDesc” (notes), “Source_FeatureID” (unique identifier), and an additional field called “Source_Date” in each feature class in the *NRRDgdb_Final.gdb* geodatabase. This will be essential in keeping track of donor datasets in the live NRRD to know which features to replace when updates are submitted.

These short-term, mid-term, and long-term next steps are summarized in Table 5. Along with these is the on-going task that new tools may need to be added or existing tools updated with any change to the NRRD schema. The Business Case and Implementation Plan section offers an outline of how this might best be done.

Phase	Improvements
Preprocessing Tools	1. Split tool #4 into a sequence of 3 tools for river reaches to accommodate more tools ***(short-term)*** 2. Improve copying of original domain fields to both a code and value "DONOR_" field ***(short-term)*** 3. Improve copying of original subtype fields by demoting the old subtype field and treating like a domain field ***(mid-term)*** 4. Formalize transferring of attributes between point and line features classes ***(long-term)***
Linear Referencing	5. Add HEM Alternative (Point Update) and (Line Update) tools ***(short-term, URGENT)*** 6. Improve HEM Alternative (Point Import) tool to copy all points to output regardless of finding match ***(mid-term, URGENT)*** 7. Improve all-around functionality of HEM Alternative (Line Import) tool ***(long-term)***
Post-processing Tools	8. Add Metadata tool ***(short-term URGENT)*** 9. Update freeform text models to truncate input to field length (or increase field lengths) ***(short-term, URGENT)*** 10. Update models to copy template feature class to scratch geodatabase before running Join Field ***(mid-term)*** 11. Continue updating field roster and tools to keep up with changes to NRRD schema ***(long term)***

Table 5: Improvements to the NRRD tools. This table summarizes next steps that the project sponsors can take with the NRRD toolkit, along with a general timeframe for each task and indicator if the task is urgent for the data donation process to be smooth.

6. Business Case and Implementation Plan

This project fits within the second phase of deployment of the National River Recreation Database (NRRD), where live datasets from donor organizations replace internal test data. In the original scoping document, the National Park Service (NPS) and River Management Society (RMS) state that "[completion] of this project will move the NRRD forward from the seminal phase of establishing a data framework...[to seeking] external datasets with confidence" (NPS & RMS 2015, 4). The realization of this opening of the NRRD to external data donors is "reliant on a data integration process in which we have confidence," which is the need this NRRD data submission framework seeks to fulfill. The justification for this project is thus wrapped into the goals of the NRRD itself, to accommodate recreational data on rivers on a national level, requiring a large community of contributors to "populate [it] and maintain its utility and vibrant relevance." While the NPS and RMS have already planned this process well in requesting a comprehensive data donation toolkit, this section will reiterate the business case for one, including justifications for particular decisions. This section will also cover recommendations for its long-term implementation with the goal of providing as much value as possible to the National Park Service, River Management Society, and all organizations that will be associated with the NRRD in the future.

6.1 Business Case

With the NRRD's schema created during phase one of development, prior to this project, data donors could conceivably add their features to copies of the NRRD feature classes and convert their attributes into the NRRD attribute fields manually. This would appear to create the same final geodatabase for submission as the workflow in this project does. However, data donors would be on their own in interpreting the meaning of each feature class and field. Inconsistencies between donor interpretations would lead to a range of inputs in each field, meaning that the live NRRD would not be able to rely on certain information. For example, fields from the HEM tool, such as ReachCode and Measure, would not be clear as to their intended values: connections to the National Hydrography Dataset (NHD). More critically, donors would be unaware of additional considerations for each feature class that are not reflected in the schema such as linear referencing. This could lead to a wide range of data quality; some datasets would be well-aligned with the NHD while others might diverge significantly, not from having higher quality data, but from being unaware of the option to reference to the NHD. Access sites may or may not be snapped to the river reaches, so the NRRD would not be able to rely on spatial queries for finding sites.

Finally, each donor GIS specialist would be forced to spend the majority of their time developing something similar to the NRRD data submission framework. The toolkit would essentially be developed anyway by one or more donor organizations, but without the direct support and knowledge of the NRRD that development by NPS and RMS would have. It would likely take a minimum of several weeks to develop even by a skilled GIS practitioner due to the time it would take to understand the NRRD, find relevant tools (e.g. the HEM tool), write additional tools, and document the workflow. The resulting third-party toolkit would not be as responsive to changes in the NRRD schema over time as a toolkit directly maintained by NPS and RMS. The third-party toolkit would likely still be widely shared, leading to outdated schema in submitted datasets and a proliferation of any other mistakes in the tools. On the other hand, donors who do not use a toolkit will be at a severe disadvantage if needing to update their submission in the future, as it is imagined every donor will need to periodically.

Overall, the lack of a data submission toolkit would present a more costly proposition to every potential donor thinking of making a submission to the NRRD, leading to less submissions and a less successful NRRD. A toolkit and workflow together provide an introduction to the NRRD and signify support from NPS and RMS in the donor organization's submission effort. A toolkit means that the donor GIS specialist can spend most of their time understanding their own data and how it compares to the NRRD and less time on developing the actual steps to perform that conversion. The tools are documented and verified to meet the requirements of the NRRD and to support a variety of potential inputs. The tools are also updated as soon as changes are made to the NRRD schema, ensuring submissions correspond to the latest schema. On the other hand, the cookbook tutorial places the tools in a standardized workflow to address data conversion issues at the most effective time. It walks data donors through each tool and brings up any issues that may be a concern at a given step. The cookbook tutorial also answers common questions quickly and directs more complicated questions to the NRRD support team. If more than 2 or 3 organizations use this toolkit, it will likely have saved the time that it took to develop during the nine-week project and will only continue providing benefits from that point on. It will also help to build more goodwill and long-term partnerships with donor organizations.

Several decisions were made that seek to further increase the value offered by the NRRD data submission framework. These decisions include developing the toolkit in ArcGIS, making the toolkit local rather than an online service, and requiring linear referencing to the NHD (in most cases). Justifications for these decisions are given in more detail in the Discussion section, but are reiterated here in terms of their benefits to the business case. Tools developed in ArcGIS benefit from the program's automatic user-interface (tool dialog box), error handling, and XML documentation structure. This helps with usability for donors as the tools will work like familiar ArcToolbox tools and they can visually assess the output in ArcMap at each step. ModelBuilder and Python tools can all be viewed and edited by users who are comfortable doing so to tailor the workflow to their data. Finally, all ArcGIS custom tools can be added to ModelBuilder themselves or called from Python, so the tools can be sequenced by the donor using their unique inputs to potentially produce a tool that executes their entire workflow. This repeatable workflow means that donors will be able to submit updates to their data much quicker than the initial submission. Deploying the tools in a zipped local toolbox enables faster development time and lets donors see the source code. Linear referencing helps quality control incoming data and standardizes spatial relationships for points on lines, enabling spatial queries within the NRRD. The exception to the linear referencing requirement is when donors have greater detail in their river reaches than the NHD. In these cases, the data submission process is also an opportunity to request updates for improvement of the NHD, strengthening the mutual benefit of a partnership with USGS. USGS can contribute to the NRRD project by helping train donors in use of the HEM tool to perform linear referencing as this was very helpful during this project.

As NPS and RMS have both recognized, a data submission toolkit and tutorial is practically required for an aggregated database like the NRRD. This is evidenced by similar efforts like ESRI's Community Maps Program, which aggregates authoritative data into a set of basemaps for all ArcGIS users (ESRI, 2015). This program has a set of data preparation tools similar to the one created in this project (ESRI, 2015).

6.2 Implementation Plan

The current set of data submission tools for the National River Recreation Database was designed with an implementation and update plan in mind. As authoritative data provided by donor organizations is essential to the NRRD's vision as a truly *national* one-stop shop for certain information regarding recreation on rivers, the most essential part of the implementation plan is how to get these tools to potential donors and how to get their prepared data back to NPS. The actions of donor organizations are also relevant to the implementation plan, as more involved donors will likely submit more and better data to the NRRD. Another aspect has to do with how tools will be added and updated as the NRRD schema changes throughout time. This is important early in the NRRD's lifespan as gaps in useful data are noticed and filled – several changes to the NRRD schema were made during this project – but also as donor agencies and the public point out other data they would like to submit or see. A final aspect of the implementation and update plan covers the context in which the tools will need to be supported by NPS and RMS in order to be functional and fully meet their goal of making the NRRD "poised to seek external data sets with confidence" (Rosebrough, 2015).

It is recommended that the full set of toolboxes and accompanying folders, including the "cookbook" tutorial, be provided for download on ESRI's ArcGIS Online platform. This places it similarly to ESRI's Community Maps Data Prep Tools at http://www.arcgis.com/home/item.html?id=9238b9012beb4f30adbdd6e709412b6c (ESRI, 2015). ArcGIS Online is already being utilized by NPS and RMS to test the NRRD schema itself for online performance, so using it to deploy the data submission tools will not be a new environment. In addition, while there are many benefits to starting a paid ArcGIS Online organization account, just posting the tools on the ArcGIS Online gallery is something any free account can do, meaning that no additional burden will be placed on NPS or RMS in the form of a premium account or hosting on their own website. The tools will also receive their own page in the gallery where any updates will be recorded in a Last Modified date and some documentation can also be provided, including links, tips, and a changelog. Finally, users of the tools will be able to leave public comments (e.g. feedback), which can be responded to and help other users of the tools as well, all without NPS or RMS having to host anything. This helps fulfill one of the secondary objectives of this project, providing feedback functionality for data donors and the public.

For returning prepared data, it is also recommended to follow the example of ESRI's Community Maps Program in having donors submit just a zipped file geodatabase. This keeps an organization's data consolidated and in the form that it will be in for the actual live NRRD. Feature classes will be named the same and all the same attribute fields will be present, meaning that a simple Append will add the donor's data to the NRRD (though later we will also recommend some internal QA/QC first). This also avoids donors' having to use more complicated web forms to submit their data. Periodic updates to donors' data can also be done through submitting a complete geodatabase just like the first submission. This may seem like overkill if only small changes have been made, but tools may have been updated in the meantime and it is easier for NPS and RMS to wholesale replace one organization's data with a new set than to look for any changes. Once donors become more familiar with the tools, they will also know just which tools to run to update their data. For example, they will only have to run the AccessSite tools if updating just access site information, while keeping the RiverReach and Campground feature classes from their original submission. Having the final submission be just a file geodatabase means that donors can pick from unchanged and newly updated feature classes to create a new file geodatabase to submit.

Critical to the implementation of these tools is donors' willingness to use them. This was driven home early in work on the project as perhaps the primary hurdle of making these tools. As such, those responsible for maintaining and developing the tools at NPS and RMS (described in more detail just after this paragraph) should be available to answer questions over email and assist donors as they begin their donation process. The "cookbook" tutorial is a great resource, but cannot cover all situations that a data donor could run into (e.g. "Is {x} really an access site? How should I represent it?"). This is especially important for the most difficult step in the process, linear referencing. The two 2-hour tutorial sessions on the Hydrography Event Management (HEM) Tool given by Michael Tinker of the USGS were invaluable to this project in terms of clarifying what was possible with the tool (and how to do it) and in increasing confidence that it could be done. As alignment with the NHD was decided multiple times as being essential for the NRRD's ability to combine datasets across the country, this knowledge will be essential to data donors as well. NPS and RMS have made an excellent decision in working closely with the USGS on the NRRD and should continue to do so as donor organizations begin to participate. Someone very familiar with the HEM tool should also be part of the NRRD team, described in the next subsection. Submitting data to the NRRD can be oriented as an incentive for learning the HEM tool, which could lead to intangible benefits with other river management organizations' being more comfortable with the USGS's hydrography data model and tools. The alternative tools to the HEM are useful in certain cases, but cannot provide the level of functionality given by the HEM likely required for large donations.

6.3 The NRRD Team

A fully operating National River Recreation Database will likely require a small team of people within either the National Park Service or River Management Society, or both. While a complete assessment of the needs for the online public release of the NRRD is outside the scope of this project, a description of the part of the "NRRD Team" focused on future development of these tools can be given. In addition, any other NPS/RMS responsibilities associated with *preparing data* for inclusion in the NRRD are important to describe to help the tools realize their full potential. It is also worth noting that NRRD team members do not necessarily need to dedicate 100% of their time to the project, but can work throughout their respective organizations and donate their specific talents when needed to the NRRD project. It is most important just to ensure that certain aspects of the NRRD data donation process are covered by someone familiar with the tools and the functionality underlying them.

The NRRD will be a living schema subject to continual updates driven both by scope changes in an internal/administrative capacity and by feedback from donor organizations and the public. Updates to the NRRD attribute fields, as well as recommendations for what content should go where, will be common early in the NRRD's lifespan, but also when major changes are made on ESRI's side with additional ArcGIS functionality. As such, there should be at least one or two GIS developers associated with the NRRD that, if not 100% dedicated to the project, are intimately familiar with the dependencies of all of the tools so as to understand the implications of changes in scope or in ESRI functionality. She/he/they will be able to provide reality checks and costs of any desired changes to the NRRD in terms of changes needed to the preparation tools and "cookbook" tutorial. As data submission is essential to the NRRD's functioning, it is important that this voice is represented in decision-making related to the NRRD. This developer (or developers) should be proficient in both Python and ModelBuilder development to make best use of an ESRI solution to these tools.

An equally important addition to the NRRD development team is a GIS developer or analyst intimately familiar with the USGS Hydrography Event Management (HEM) tool. This can be the same person as the tool updater, or someone different. As it took a total of 4-hours of video training to even gain literacy with this tool, it will require a significant amount of time to help train members of potential donor organizations at large. While tutorial videos and documentation do exist for the HEM tool, a more critical issue is making the GIS specialist at the donor agency feel comfortable enough to make the investment in learning it. An alternative to the HEM tool may be possible for small areas or where the NHD is already being used for river reaches, but any full solution would necessarily be essentially the HEM anyways, so encouraging donor organizations to just make the investment in the HEM is advised. As a result, training donor organizations will require more time than could be reasonably given by someone not associated with the NRRD project. The focus of this developer or analyst would be on providing the smoothest HEM experience for donor organizations in specifically the functions required to submit data to the HEM. This is what is currently outlined in the linear referencing section of the "cookbook" tutorial. She/he would likely be meeting with many potential donor organizations to clarify the HEM process and to help answer any questions that they may have. This might even go as far as helping troubleshoot specific situations that donors are running into – several confusing situations were encountered during this project alone with only 2 sample datasets. This is worth it primarily for two reasons: 1) there is usually only one source of authoritative data for certain areas (e.g. each National Park), so it is important that they are supported enough to submit their data, or else the NRRD will have gaps, and 2) donor GIS specialists will become more familiar over time with the tools and be able to submit updated data more independently in the future.

Finally, the tools will not be able to handle all cases encountered by donor organizations. Especially with the linear referencing of river reach data, there can be errors introduced that lead to results unsuitable for posting. ESRI's Community Maps implements a two-phase review process where an automatic review checks the geodatabase for the right feature classes and right attributes fields. Then, the submission waits on a manual check by ESRI employees (ESRI, 2015). The automatic check can likely be skipped due to the small number of features classes currently part of the NRRD (3 + 1 table), but the manual quality assessment is more critical as the data will be used for more than basemap symbology. At least two GIS specialists should be familiar with the NRRD data model enough to QA/QC incoming submissions and updates. They should be in close communication with the developer(s) and HEM specialist to understand common errors that can be introduced and how to advise donors on correcting them. They should also be available to answer questions that donors may have, along with the developer(s) and HEM analyst. *No data should make it into the live NRRD without being checked, to some degree, by someone on the NRRD team.* This becomes even more important as analysis functions are added to the NRRD for answering questions from “both the public and river managers” as per the scoping document (e.g. “How many NPS [Wild and Scenic Rivers] have public river access sites? How many do not have any river access sites?”) (2). Testing some of these spatial queries on incoming datasets will help determine if they are suitable for the NRRD. The GIS specialists can then also perform the actual Appends to the live NRRD once data is deemed suitable.

As such, the data submission workflow provided in the “cookbook” sits inside of a larger workflow for maintenance of the National River Recreation Database and the data submission tools. Figure 52 shows a swimlane diagram for this maintenance and update workflow process, including essential tasks for each entity. The workflow is divided into two main phases (left labels): the (mostly internal) maintenance of the NRRD and tools and the collaboration with donor organizations and the public.

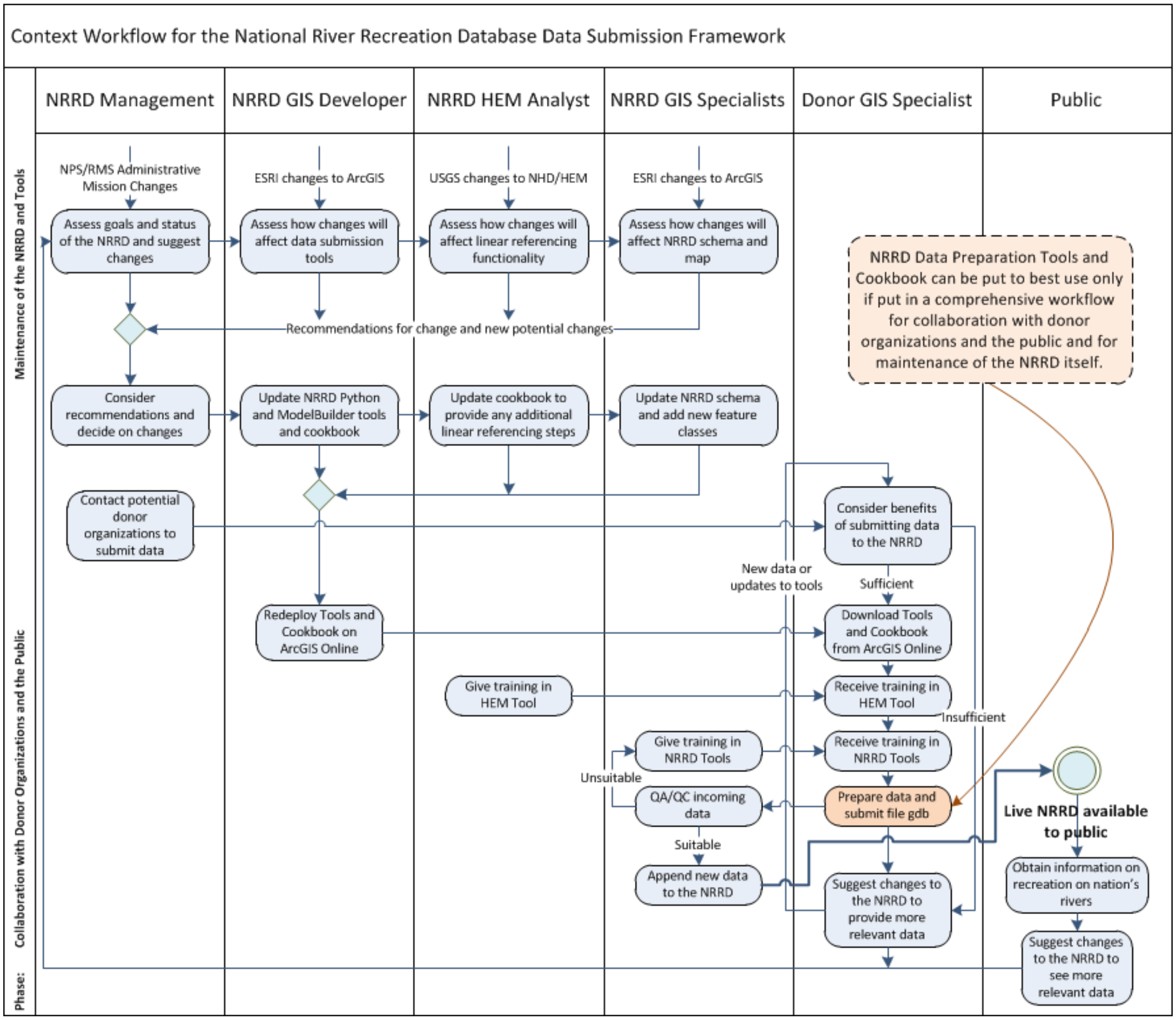


Figure 52: Swimlane Diagram of Support Workflow. The NRRD toolkit and cookbook developed in this project will continue providing long-term benefits only if placed within a comprehensive workflow for maintenance and collaboration. Changes to the NRRD schema should be immediately reflected in new and updated tools. Collaboration with data donors and among all members of the NRRD team is critical to ensuring the tools are used to their maximum potential.

All 4 parties within the development division of the NRRD team (4 left columns) will collaborate in some respect with donor organizations. NRRD management, whether part of NPS, RMS, or another partner entity, will contact potential data donors to introduce them to the project and explain the benefits of contributing. It may be at some point that enough organizations are finding the website and submitting data voluntarily that this may become less of a task. The NRRD GIS developer(s) will update the tools and cookbook themselves and deploy those updates to the ArcGIS Online location. The NRRD HEM analyst will train the GIS specialists of donor organizations in the HEM tool and help troubleshoot. The NRRD GIS specialists will troubleshoot the NRRD-specific tools, QA/QC incoming data (including perhaps running the HEM tool on datasets that were created using the HEM alternative workflow), and append checked data to the live NRRD map service. The public, donor organizations, and the NRRD team will all use the NRRD and suggest changes, which are considered by all 4 parties of the NRRD team to understand the implications of any upgrade. In this way, the toolbox and documentation helping donor organizations to contribute to the NRRD will be part of a comprehensive organizational workflow, responsive to changes both internal – administrative mission and scope – and external – ESRI's updates to ArcGIS, USGS's updates to the HEM tool, and requested functionality from donors and the public. It is our hope that the tools and cookbook created in this project will evolve alongside the National River Recreation Database throughout its lifespan and that a robust framework for support and collaboration will help all parties involved realize the maximum benefits of participating.

Works Cited

American Whitewater. (2015). Retrieved August 11, 2015, from American Whitewater: https://www.americanwhitewater.org/

Burger, J. (2000, April 17). Landscapes, tourism, and conservation. *Science of The Total Environment, 249*(1-3), pp. 39-49.

ESRI. (2015). *Community Maps Data Prep Tools*. Retrieved August 2015, from http://www.arcgis.com/home/item.html?id=9238b9012beb4f30adbdd6e709412b6c.

ESRI. (2015). *The Community Maps contribution process*. Retrieved August 2015, from http://doc.arcgis.com/en/community-maps/contribute/contribution-process.htm

Flowerdew, R. (1991). Spatial Data Integration. In M. F. D. J. Maguire, *Geographical Information Systems: principles and applications* (pp. 375-387). London: Longman.

Forster, L. E. (2005, July). Effect of boat launches on benthic invertebrates in a lentic system. *Journal of Ecological Research*, pp. 8 - 12.

Liddle, M. a. (1980, April). The effects of recreation on freshwater plants and animals: A review. *Biological Conservation*, pp. 183 - 206.

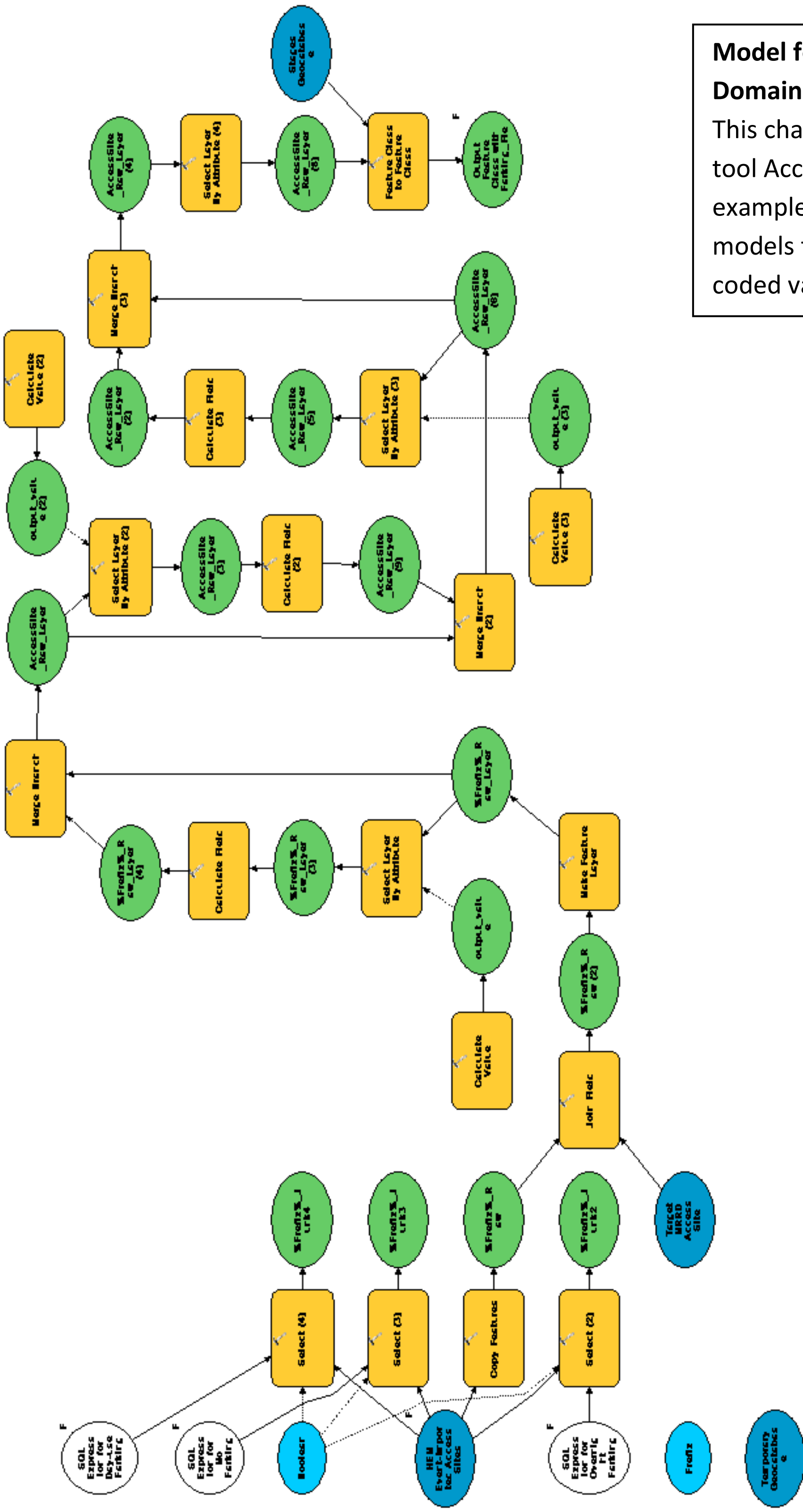
Rosebrough, S. (2015). *Workshop Sponsor Questionnaire.* National Park Service.

USGS. (2015). Retrieved August 2015, from National Hydrography Dataset: http://nhd.usgs.gov/

USGS. (2015). *NHD Tools*. Retrieved August 2015, from National Hydrography Dataset: http://nhd.usgs.gov/tools.html

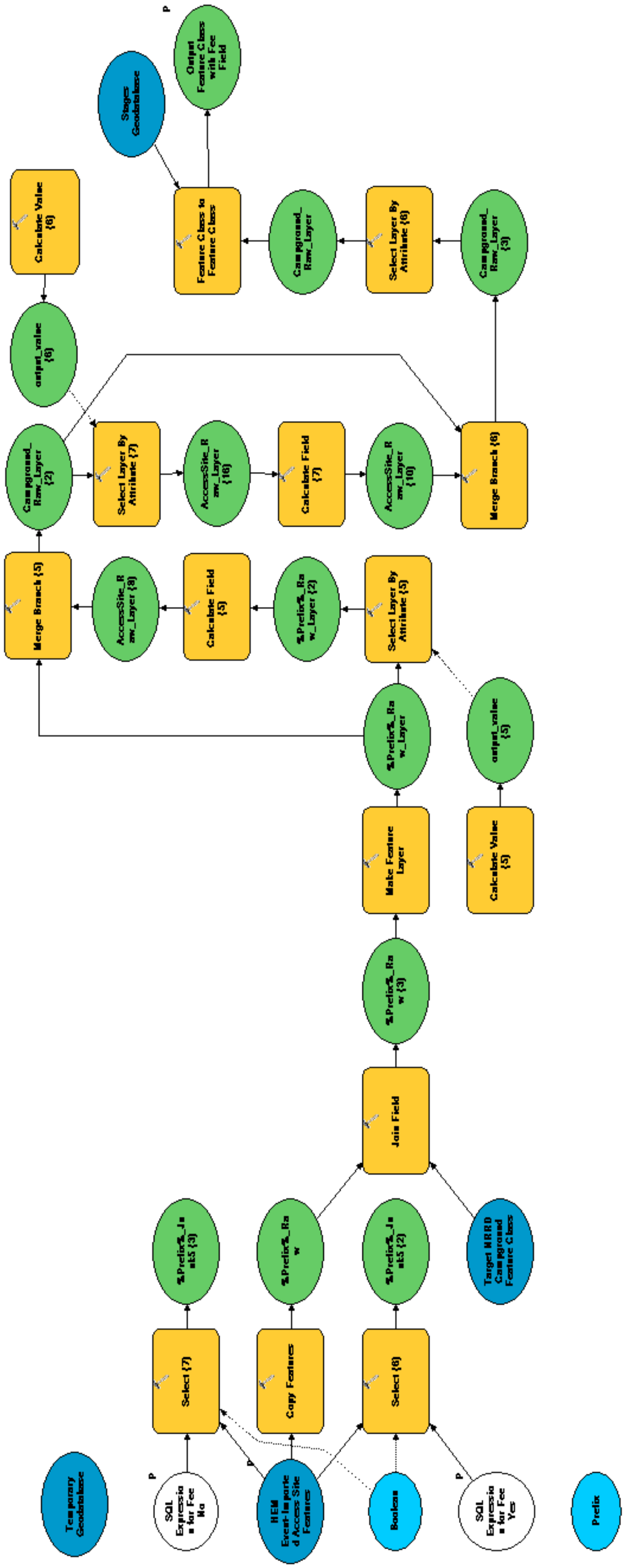
Zandbergen, P. A. (2013). *Python scripting for ArcGIS.* Redlands, CA: ESRI Press.

Appendix A: Examples of Tool Models and Scripts



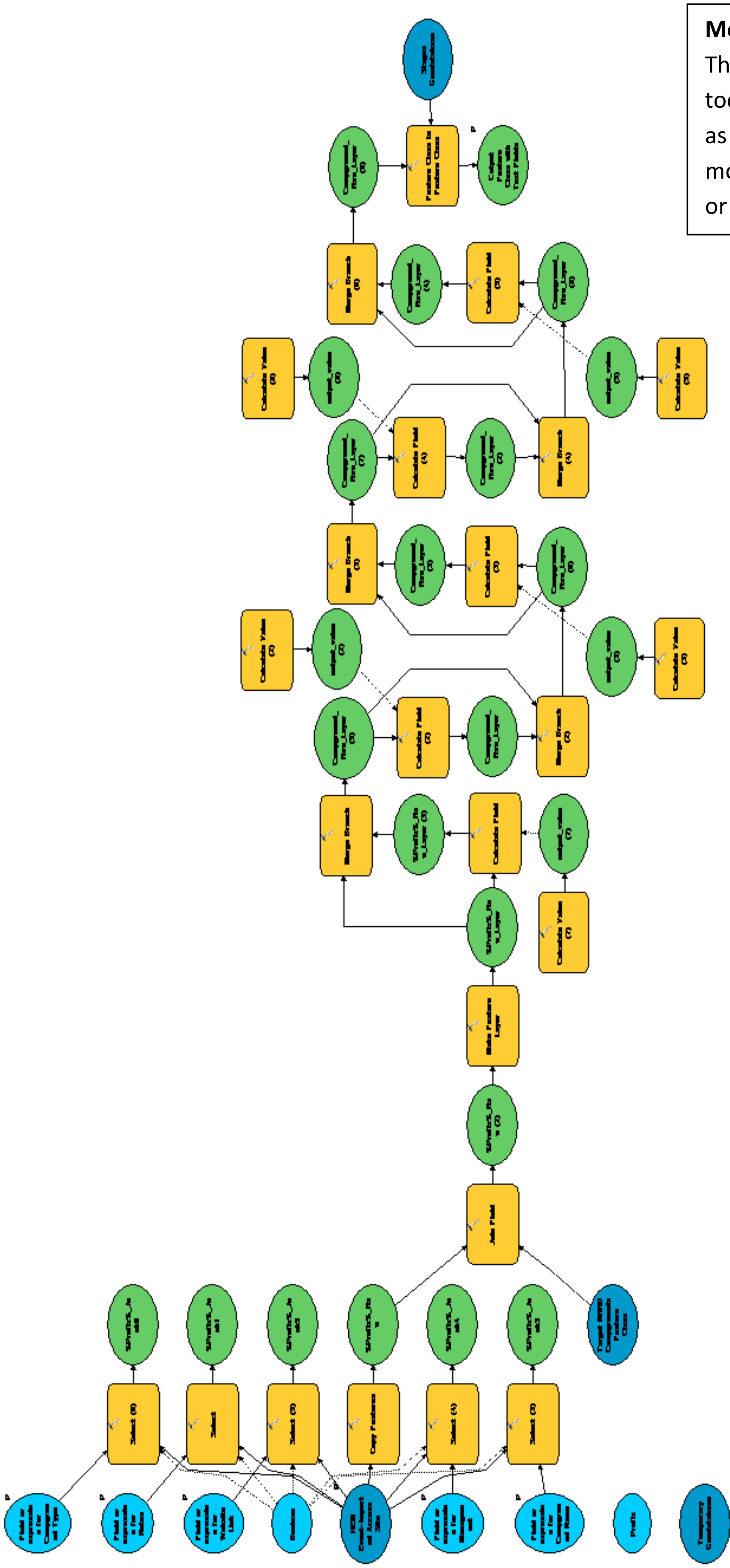
Model for Populating Coded Value Domain Fields

This chart shows the model for the tool Access Site – 2. Parking as an example of the workflow for all models that populate fields with coded value domains.



Model for Populating Yes/No Fields

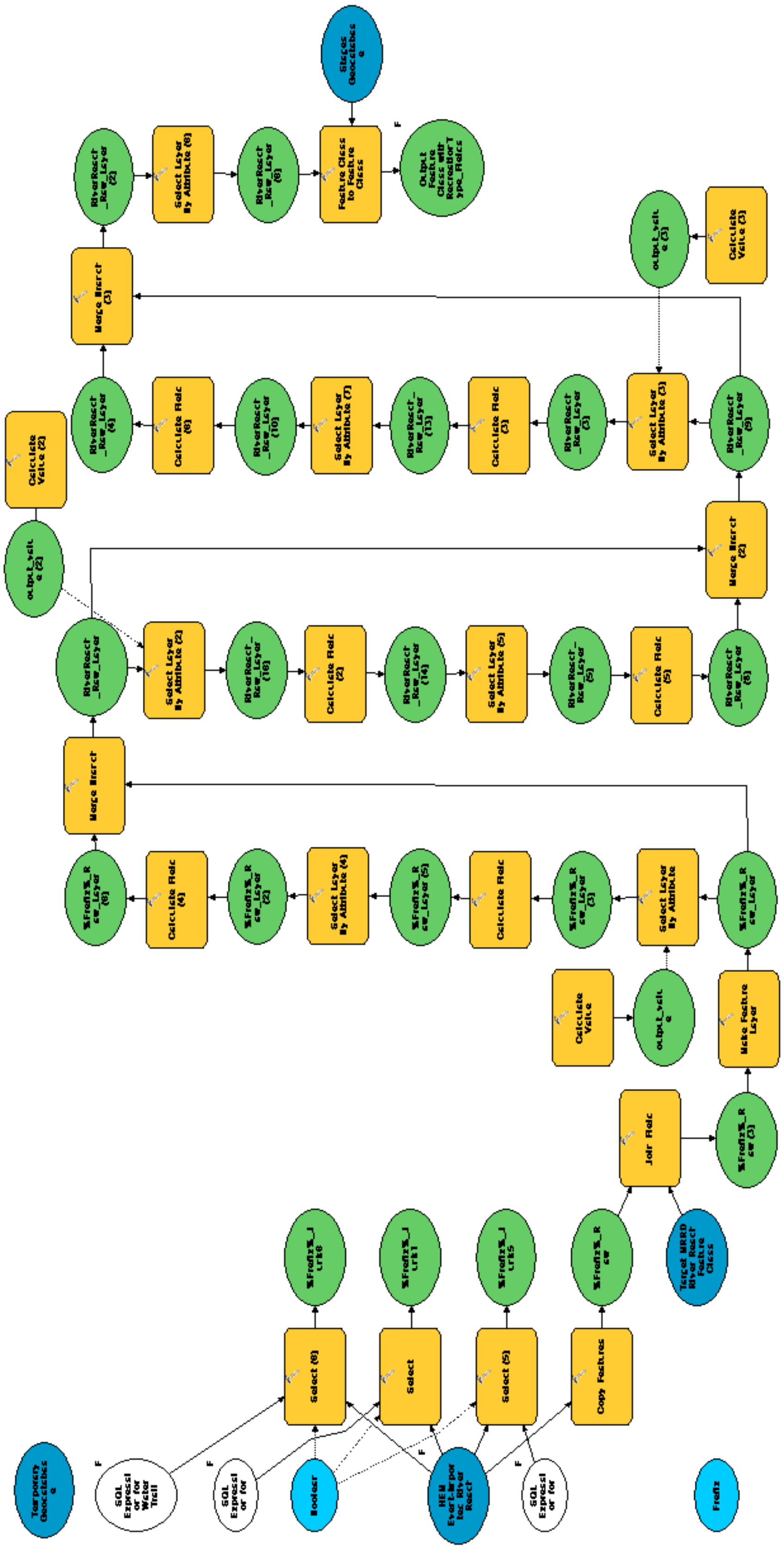
This chart shows the model for the tool "Campgrounds – 1. Fees" as an example of the workflow for all models that populate fields with yes/no values.



Model for Populating Text Fields
This chart shows the model for the tool “Campgrounds – 2. Text Fields” as an example of the workflow for all models that populate fields with text or string values.

Model for Populating Recreation Type Field

This chart shows the model for the tool "River Reach – 08. Recreation Type," which is the only required yes/no field in the database and is therefore treated differently than other yes/no fields.



Script for 1. Classify Point Types
The next 8 pages contain the Python script for the tool “1. Classify Point Types” from the preprocessing toolbox. This tool is designed to split a feature class into 3 output feature classes: access sites, campgrounds, and unclassified based on two SQL statements. The tool also removes all original fields and replaces them with identical copies prefixed with “DONOR_” to avoid field name conflicts later in the workflow. Finally, the script contains sections at the beginning and end that are common to all scripts in this project, including deleting of intermediate datasets after the tool successfully completes. This represents a kind of “template” for creating additional NRRD tools. Check the comments (after #) for more information about each section.

```
### Import relevant modules.

import arcpy

import sys
import os


### Start the script.

arcpy.AddMessage("\n{0:*^33}".format(" Script initiated. "))


### Define input parameters.

label = "Defining input parameters..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Import parameters from tool dialog box.
in_feat        = arcpy.GetParameterAsText(0)
sql_access_site = arcpy.GetParameterAsText(1)
sql_campground  = arcpy.GetParameterAsText(2)


# Define additional parameters.

addl_params = []

# Script Location
script_path = __file__
addl_params.append(("Script Location", script_path))

# Locate the current working directory assuming intended file structure.
cwd = os.path.dirname(os.path.dirname(__file__))
addl_params.append(("Current Working Directory", cwd))

# Scratch Geodatabase
scratch_gdb = os.path.join(cwd, r"Scratch\NRRDgdb_Scratch.gdb")
addl_params.append(("Scratch Geodatabase", scratch_gdb))
```

```
# Name of the input features.
in_desc = arcpy.Describe(in_feat)
in_name = in_desc.name
addl_params.append(("Input Feature Class Name", in_name))

# Locations of the derived output datasets.
out_access_site = os.path.join(
    cwd,
    r"Preprocessing_NAD83.gdb\Raw_AccessSites",
    "AccessSite_NRRD_Raw_" + in_name
    )
out_campground   = os.path.join(
    cwd,
    r"Preprocessing_NAD83.gdb\Raw_Campgrounds",
    "Campground_NRRD_Raw_" + in_name
    )
out_unclassified = os.path.join(
    cwd,
    r"Preprocessing_NAD83.gdb\Raw_Unclassified",
    "Unclassified_NRRD_Raw_" + in_name
    )

deriv_types = {"Feature Class", "Table"}
deriv_params = [param.displayName for param in arcpy.GetParameterInfo() if
                param.parameterType == "Derived" and
                param.datatype in deriv_types]
deriv_paths = [out_access_site, out_campground, out_unclassified]
addl_params.extend(zip(deriv_params, deriv_paths))

# Today's Date and Time
today = datetime.datetime.now()
addl_params.append(("Today's Date and Time", today))

# Intermediate Data Prefix (to avoid name conflicts in scratch gdb).
prefix = "ClassifyTypes"
addl_params.append(("Intermediate Data Prefix", prefix))


# Restate inputs.

lbl = "Tool Dialog Box"
arcpy.AddMessage("\n{1}\n{0}\n{1}".format(lbl, "-" * (len(lbl)+1)))
```

```
for param in arcpy.GetParameterInfo():
    if param.parameterType != "Derived":
        arcpy.AddMessage("\n{0}:\n{1}".format(param.displayName, param.value))

lbl = "Automatic"
arcpy.AddMessage("\n{1}\n{0}\n{1}".format(lbl, "-" * (len(lbl)+1)))

for name, value in addl_params:
    arcpy.AddMessage("\n{0}:\n{1}".format(name, value))


### Main Processing.

lbl = "MAIN PROCESSING"
arcpy.AddMessage("\n\n{1}\n{0}\n{1}\n".format(lbl, "*" * (len(lbl)+1)))

# Create list to hold paths to intermediate datasets to delete at end.
interm_data = []


### Add new fields.

label = "Adding new fields..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Copy input features to scratch geodatabase.
temp1 = os.path.join(scratch_gdb, prefix + "_RawClassify")
interm_data.append(temp1)
arcpy.CopyFeatures_management(in_feat, temp1)


# Add fields.

# DONOR_ copies of original fields.
skip_types = {"OID", "Blob", "Guid", "Geometry", "Raster"}
donor_fields = [field for field in in_desc.fields if
                field.type not in skip_types]

for field in donor_fields:
    arcpy.AddMessage("\tDONOR_" + field.name)
```

```
    arcpy.AddField_management(
        temp1,
        "DONOR_" + field.name,
        ("TEXT" if field.type == "String" else
         "SHORT" if field.type == "SmallInteger" else
         "FLOAT" if field.type == "Single" else
         "LONG" if field.type in {"Integer", "OID"} else field.type.upper()),
        field.precision,
        field.scale,
        field.length,
        field_domain=field.domain
        )

# New NRRD_ preprocessing fields.
arcpy.AddMessage("\tNRRD_OrigFC")
arcpy.AddField_management(temp1, "NRRD_OrigFC", "TEXT", field_length=100)

arcpy.AddMessage("\tNRRD_PointType")
arcpy.AddField_management(temp1, "NRRD_PointType", "TEXT", field_length=25)


### Calculate new fields.

label = "Calculating new fields..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Make feature layer to run selection on one table.
temp1_layer = "temp1_layer_" + arcpy.Describe(temp1).name
arcpy.MakeFeatureLayer_management(temp1, temp1_layer)


# Calculate fields.

field = "NRRD_OrigFC"
arcpy.AddMessage("\t" + field)
arcpy.SetProgressorLabel(label + "\n\t" + field)
arcpy.AddMessage("\t\t{0}".format(in_name))
arcpy.CalculateField_management(temp1_layer, field,
                                repr(in_name), "PYTHON_9.3")
```

```
# Fields with SQL expressions for categories.

# Define field.
field = "NRRD_PointType"
arcpy.AddMessage("\t" + field)
arcpy.SetProgressorLabel(label + "\n\t" + field)

# Match SQL expressions and values.
cats_and_sqls = [
  ("Access Site", sql_access_site),
  ("Campground", sql_campground)
  ]

# Run calculations for SQL expressions.
for cat, sql_exp in cats_and_sqls:
  arcpy.AddMessage("\t\t{0}".format(cat))

  if sql_exp:
    arcpy.AddMessage("SQL Expression:\n{0}\n".format(sql_exp))

    # Build Field Calculator expression. Do not overwrite non-Null
    # values.
    expression = "{0} if not !{1}! else !{1}!".format(repr(cat), field)

    # Run selection and make calculation.
    arcpy.SelectLayerByAttribute_management(temp1_layer,
                                            "NEW_SELECTION", sql_exp)
    arcpy.CalculateField_management(temp1_layer, field,
                                    expression, "PYTHON_9.3")

# Run the "remainder" calculation.
cat = "Unclassified"
arcpy.AddMessage("\t\t{0}".format(cat))
expression = "{0} if not !{1}! else !{1}!".format(repr(cat), field)

# Clear selection and make calculation.
arcpy.SelectLayerByAttribute_management(temp1_layer, "CLEAR_SELECTION")
arcpy.CalculateField_management(temp1_layer, field,
                                expression, "PYTHON_9.3")


# Calculate DONOR_ copies of original fields and remove the originals.
```

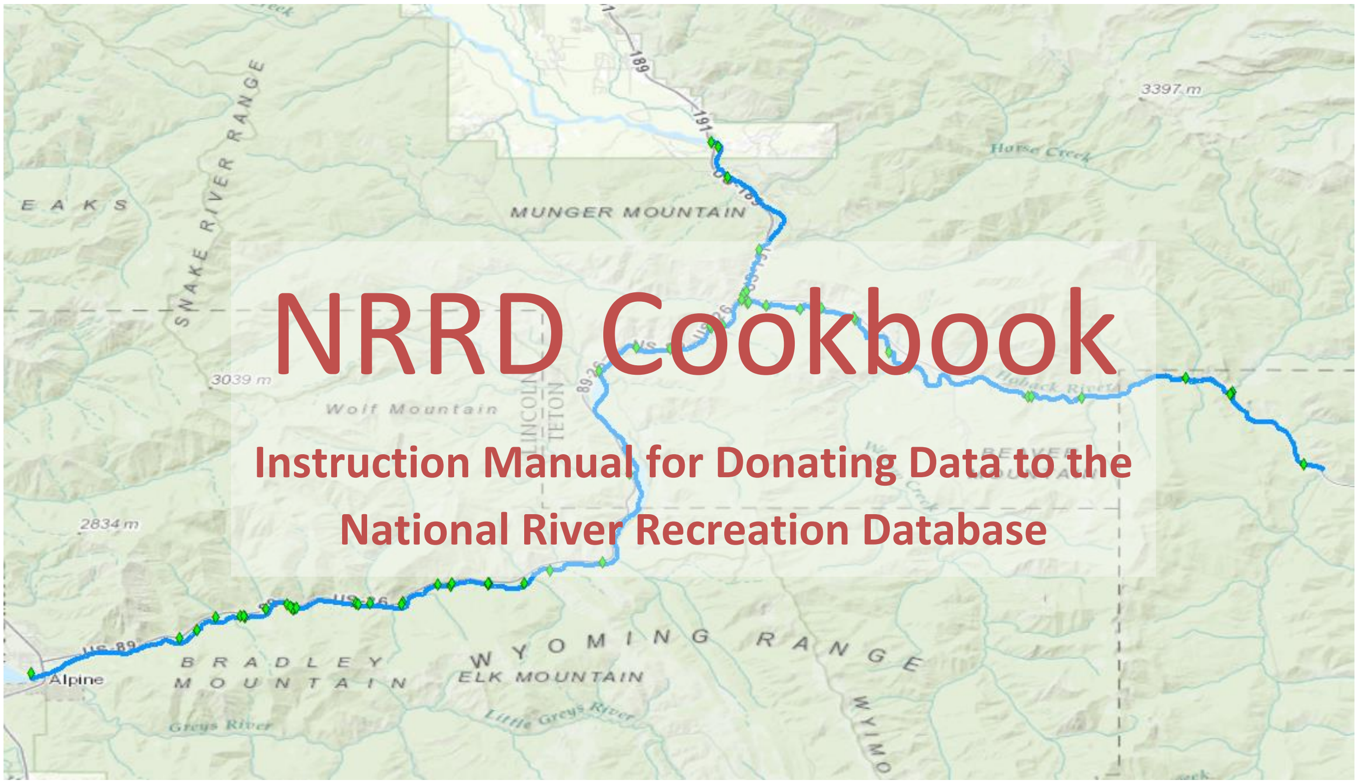
```
arcpy.AddMessage("\tDONOR_ copies of original fields...")
arcpy.SetProgressorLabel(label + "\n\tDONOR_ copies of original fields...")
for field in donor_fields:
    arcpy.CalculateField_management(
        temp1_layer,
        "DONOR_" + field.name,
        "!" + field.name + "!",
        "PYTHON_9.3"
        )

arcpy.AddMessage("\t\tRemoving original fields...")
for field in donor_fields:
    if not field.required:
        try:
            arcpy.DeleteField_management(temp1_layer, field.name)
        except Exception as e:
            arcpy.AddWarning("Could not delete field {0}".format(field.name))
            arcpy.AddWarning(e)


### Split feature class.

label = "Splitting feature class..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Define split parameters.
field = "NRRD_PointType"
temp_as = os.path.join(scratch_gdb, prefix + "_AccessSite")
temp_cg = os.path.join(scratch_gdb, prefix + "_Campground")
temp_uc = os.path.join(scratch_gdb, prefix + "_Unclassified")
interm_data.extend([temp_as, temp_cg, temp_uc])

cats_and_fcs = [
    ("Access Site", temp_as),
    ("Campground", temp_cg),
    ("Unclassified", temp_uc)
    ]

# Run Select to split feature class.
for cat, fc in cats_and_fcs:
```

```
    arcpy.AddMessage("\t{0}".format(cat))

    # Build SQL expression.
    sql_exp = "{0} = {1}".format(field, repr(cat))

    # Run selection.
    arcpy.Select_analysis(temp1_layer, fc, sql_exp)


### Produce output datasets.

label = "Producing output datasets..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

# Define output dataset parameters.
dss_and_outs = [
    ("Feature Class", temp_as, out_access_site),
    ("Feature Class", temp_cg, out_campground),
    ("Feature Class", temp_uc, out_unclassified)
    ]

# Get the indices and parameter types of output parameters that have the
# above data types.
types_here = {i[0] for i in dss_and_outs}
out_inds = []
out_paramtypes = []
for ind, param in enumerate(arcpy.GetParameterInfo()):
    if param.direction == "Output":
        if param.datatype in types_here:
            out_inds.append(ind)
            out_paramtypes.append(param.parameterType)

# Produce output datasets according to data type.
for param_ind, param_type, (data_type, temp_ds, out_ds) in\
        zip(out_inds, out_paramtypes, dss_and_outs):
    arcpy.AddMessage("\t{0}".format(os.path.basename(out_ds)))

    # Throw a warning with the number of records.
    count = int(arcpy.GetCount_management(temp_ds).getOutput(0))
    arcpy.AddWarning("\t\tOutput table contains {0} records.".format(count))
```

```
    # Run an output method based on the data type.
    if data_type == "Feature Class":
        arcpy.CopyFeatures_management(temp_ds, out_ds)
    elif data_type == "Table":
        arcpy.CopyRows_management(temp_ds, out_ds)
    else:
        err_msg = ("output method not defined for data_type {0}".
                   format(data_type))
        raise ValueError(err_msg)

    # Set output parameter if derived.
    if param_type == "Derived":
        arcpy.SetParameterAsText(param_ind, out_ds)


### Delete intermediate data.

label = "Deleting intermediate data..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

for i in interm_data:
    arcpy.AddMessage("\t{0}".format(os.path.basename(i)))
    arcpy.Delete_management(i)


### Restate results.

label = "Restating results..."
arcpy.AddMessage("\n" + label)
arcpy.SetProgressorLabel(label)

for param in arcpy.GetParameterInfo():
    if param.direction == "Output":
        arcpy.AddMessage("\n{0}:\n{1}".format(param.displayName, param.value))


### End the script.

arcpy.AddMessage("\n{0:*^33}\n".format(" Script completed. "))
```

Appendix B: Cookbook

The following pages hold the instruction manual created to assist users in navigating the NRRD tool kit.

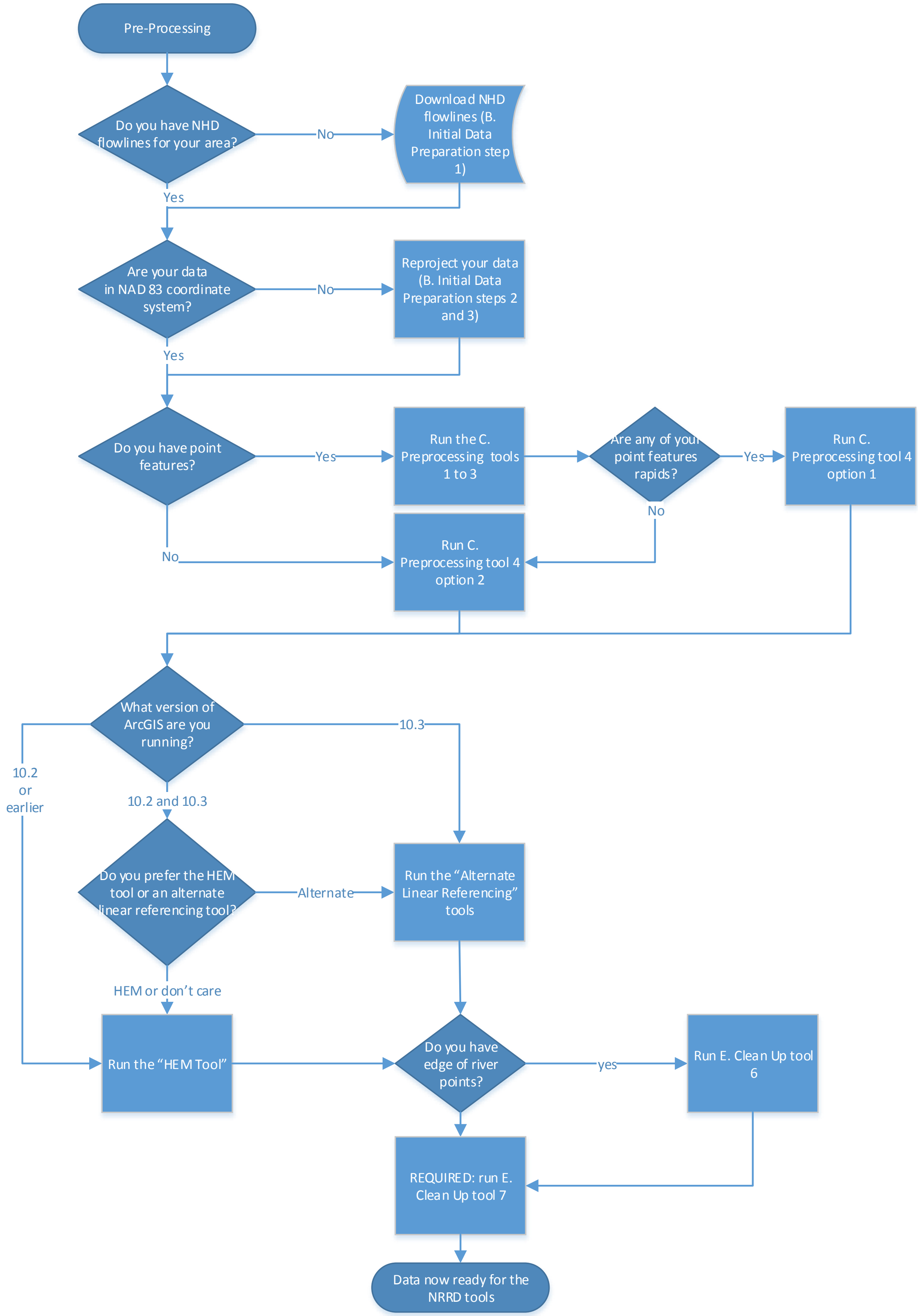


NRRD Cookbook

Instruction Manual for Donating Data to the National River Recreation Database

Contents

Figure 53: Workflow Diagram for NRRD Processing



A. Introduction

This handbook will guide you through the steps necessary to transform your river recreation data so that it can be incorporated into the National River Recreation Database (NRRD). Minimum skills and requirements include:

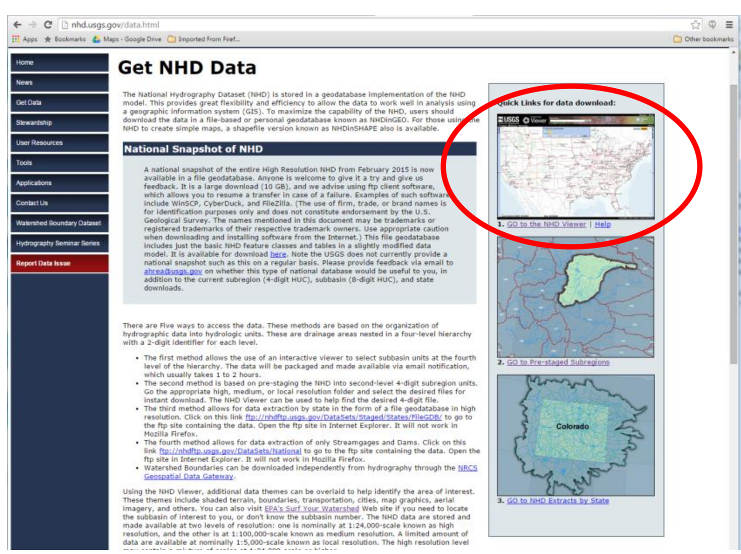
7. Access to ArcGIS version 10.2 and/or 10.3
8. Ability to navigate to folders and geodatabases in ArcCatalog
9. Ability to use tools in ArcToolbox
10. Ability to edit line and point features in an ArcMap editing session
11. Ability to use the field calculator to set the content of fields in attribute tables
12. Ability to compose SQL expressions in a selection dialog box

There are five basic sections, each with a suite of steps or tools to follow and use in order. These include:

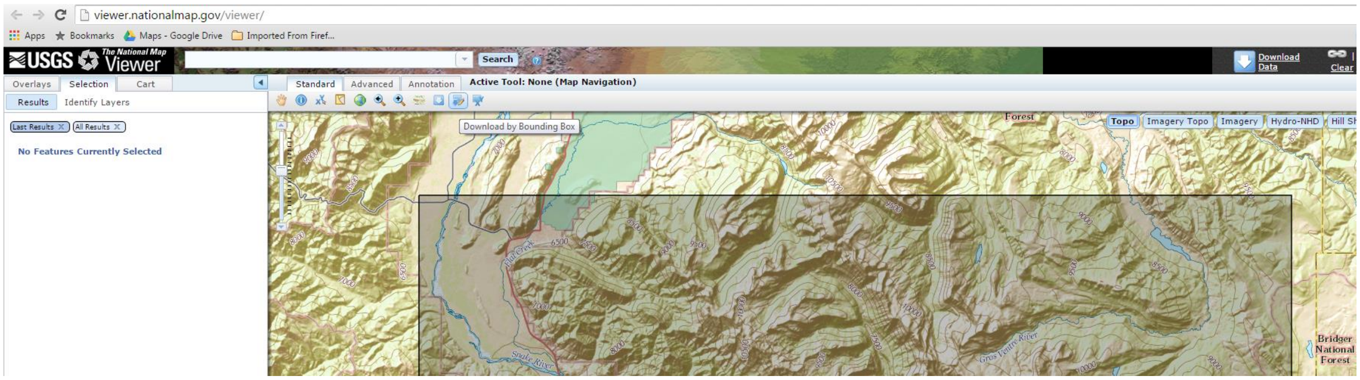
1. **Initial data preparation** to ensure that you have all the necessary feature classes and they are all in the correct projections;
2. **Preprocessing point and line data**, which will set point and line data up for linear referencing;
3. **Linear referencing**, where you will choose one of two options for relating your point and line data to the National Hydrography Dataset (NHD);
4. **Post linear referencing clean up tools,** where you will restore your original geometry and combine your output into three feature classes as inputs for the NRRD tools; and
5. **NRRD tools**, which will walk you through a series of tools that populate the fields in the NRRD database with comparable data from your databases.

B. Initial Data Preparation

1. If you do not already have the National Hydrography Dataset (NHD) for your area, download the data. *Tip: for faster processing, choose the smallest NHD area that still encompasses your entire region*.
   a. Go to http://nhd.usgs.gov/data.html. Download data using any of the options, or:
   b. Go to the NHD viewer (right side of webpage) and navigate to your project area on the map.

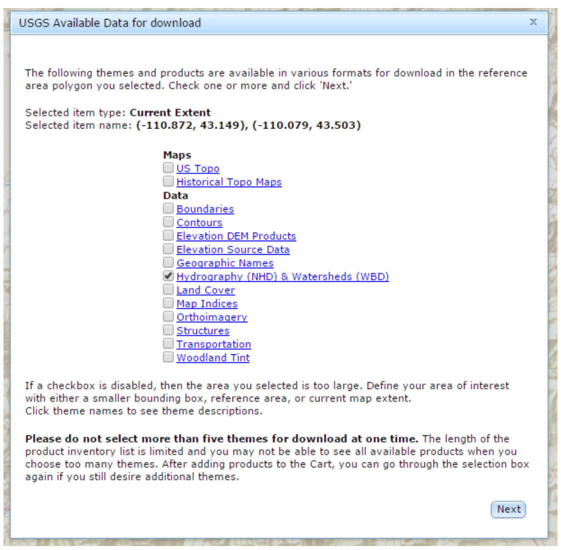


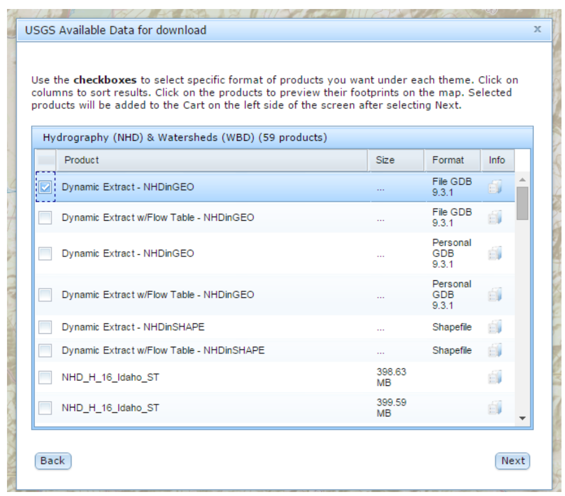
c. From the standard toolset above the map, choose the “download by bounding box” tool and draw a rectangle that encompasses your entire area.



d. In the menu that pops up, choose “Hydrography (NHD) & Watersheds (WBD)” and click “next.”

e. In the next dialog box (“USGS Available Data for Download”), choose the first option, “Dynamic Extract – NHDinGEO,” and click “next.”

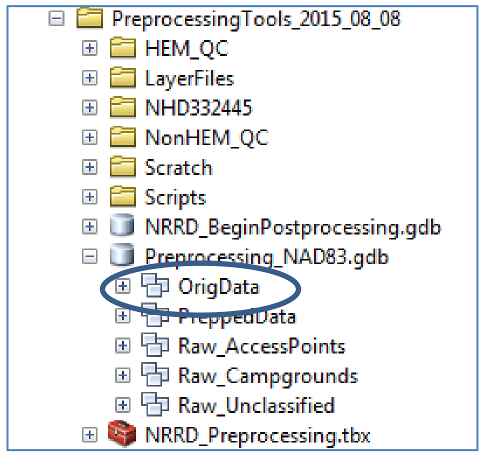




f. On the left side of your screen, the tab “Cart” comes to the front with your extraction highlighted. Click “Checkout,” provide your e-mail address, and click “place order.” It may take some time for your files to appear in your in-box.

2. Open a blank map in ArcGIS and add your NHD flowline feature class to the map. *Tip: if you load other layers into the map first, be sure that your data frame is in the same coordinate system as the NHD, or the linear referencing tools will not function properly*.

3. Add your source data to the map and, if it is not already in the same coordinate system as the NHD, reproject the data.
   a. From ArcToolbox, go to Data Management Tools →Projections and Transformations → Project.
   b. In "Input Dataset or Feature Class," choose your data from the drop-down list.
   c. In "Output Dataset or Feature Class," Navigate to the "OrigData" feature dataset in the Preprocessing_NAD83.gdb" geodatabase in your Preprocessing Tools folder.
   d. For the output coordinate system, go to Geographic Coordinate Systems → North America → NAD 1983 and click Okay.



NOTE: All of your data needs to be in the "OrigData" feature dataset described in step 3.c. above. If your data does not need to be reprojected, you need to copy each feature class to the "OrigData" feature dataset.

C. Preprocessing Point and Line Data

These tools will prepare your point and line data for processing using the HEM tool or the Alternate Linear Referencing tool. If you do not have point data, you can skip to tool No. 4. They allow you to specify which points are in the river (snapped to the centerline), which points are on the edge of the river (river right or river left), and which points are off the river (e.g. campgrounds); which points are access sites or campgrounds; and which points are rapids. Point features that are not access sites, campgrounds, or hazards (e.g. river miles) will not be incorporated into the NRRD. All line features must also be run through one of the two line preprocessing tools.

Tools in the NRRD_preprocessing toolbox

1. Classify Point Types tool REQUIRED if you have point data
2. Classify Point Locations tool REQUIRED if you have point data
3a. Combine Access Points tool REQUIRED if you have access point data
3b. Combine Campgrounds tool REQUIRED if you have campground point data
4. Line feature class prep REQUIRED if you have line data: OPTION 1: Join rapids to river reaches tool OPTION 2: Prep River Reaches Tool
5. (Linear Referencing tool of your choice)
6. Restore Edge Of River Points tool REQUIRED if you have edge of river points
7. Combine Linear Referenced Feature Classes REQUIRED for all data

Figure 54: NRRD preprocessing toolkit catalog tree

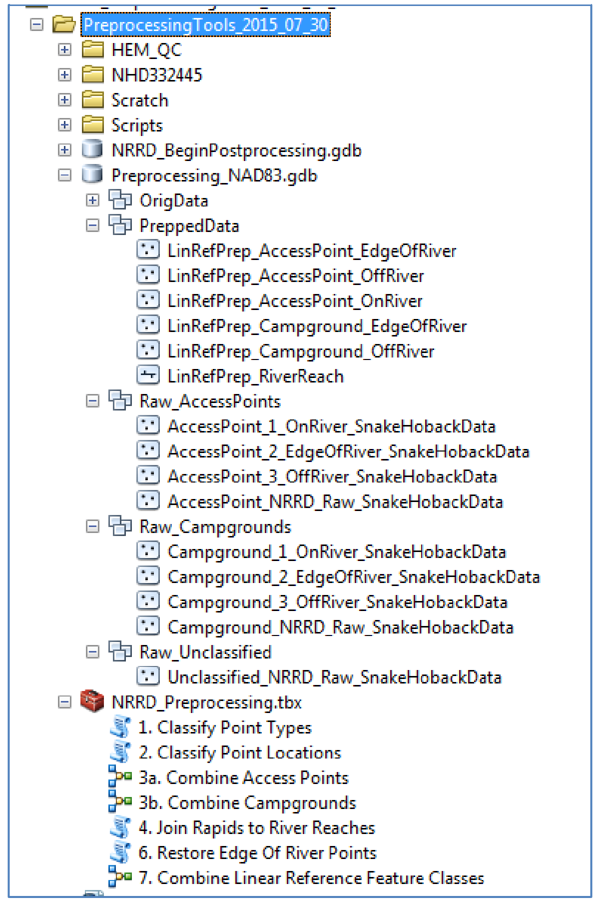


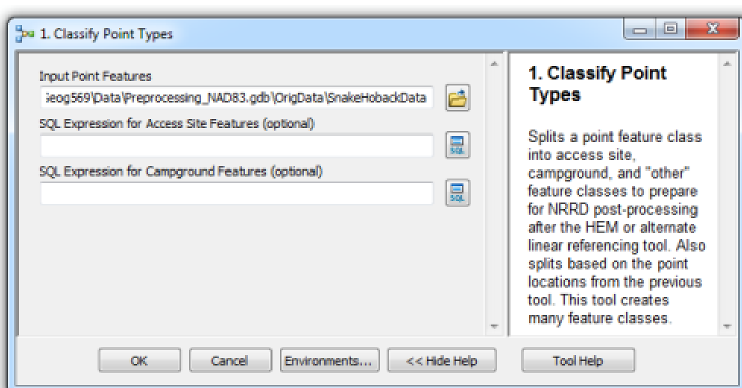
Figure 2 shows the catalog tree you will find when you open the "Preprocessing Tools" folder. To begin with, each of the feature datasets in the "Preprocessing_NAD83.gdb" will be empty. They will be populated by intermediate feature classes as you run the tools, and you will find the required input feature classes for the later tools in these feature datasets. Refer back to figure x to help locate the inputs you need to run the following tools.

IF YOU HAVE MULTIPLE POINT FEATURE CLASSES CONTAINING DATA ON ACCESS SITES AND CAMPGROUNDS, YOU SHOULD RUN TOOLS 1 AND 2 ON EACH OF THOSE FEATURE CLASSES.

1. *Classify Point Types tool*

This tool separates your point features into three categories: access sites, campgrounds and "other." The NRRD only makes use of access site and campground data, so other types of point data (such as river miles) will not be retained. Note that rapids will be dealt with in tool 4, "Join Rapids to River Reaches." This tool adds a field identifying each feature as one of the three point types. THIS TOOL ALSO ADDS FIELDS TO YOUR POINT FEATURE CLASS THAT ARE NECESSARY FOR THE FUNCTIONING OF FUTURE TOOLS, SO YOU NEED TO RUN THIS TOOL EVEN IF YOU HAVE ONLY ONE TYPE OF POINT IN YOUR DATASET (E.G. ONLY ACCESS SITES).

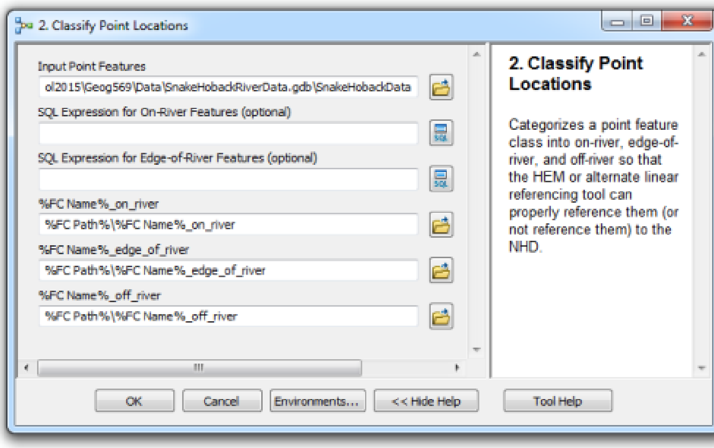
1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "1. Classify Point Types."
2. For "Input Point Features," choose your point feature class from the dropdown menu or navigate to its location. *Tip: make sure your feature class is in the "orig" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase.*
3. In the next two boxes, click the SQL icon to the right of each box and create an expression that indicates which point features correspond to each of the NRRD feature types. If you have no features of a particular type, leave that box blank. *Tip: if you have only one type of point feature, use the IS NOT NULL buttons in the SQL dialog box to select all records (e.g. Access_Site_Name IS NOT NULL).*



4. The output of this tool is three feature classes in the "Preprocessing_NAD83.gdb" geodatabase. One, prefixed "AccessPoint_NRRD_Raw…," is in the "Raw_AccessPoints" feature dataset. Another, prefixed "Campground_NRRD_Raw…," is in the "Raw_Campgrounds" feature dataset. The third, prefixed "Unclassified_NRRD_Raw" is in the "Raw_Unclassified" dataset. You will use these as inputs to the tools that follow. The tool should also add these feature classes to your map document, so you should be able to access them from the dropdown list.

2. Classify Point Locations tool

This tool separates your point features into three categories: on the river, adjacent to the river, and off the river. Each of these point types will be treated differently in the HEM or alternate linear referencing tools workflow. You need to run this tool up to TWO TIMES, once for access points (if you have them), and once for campgrounds (if you have them). This tool adds a field identifying each feature as having one of the three location types. AGAIN, YOU MUST RUN THIS TOOL EVEN IF ALL YOUR POINT FEATURES ARE IN THE SAME LOCATION (E.G. ALL ARE ACCESS SITES ON THE EDGE OF THE RIVER).

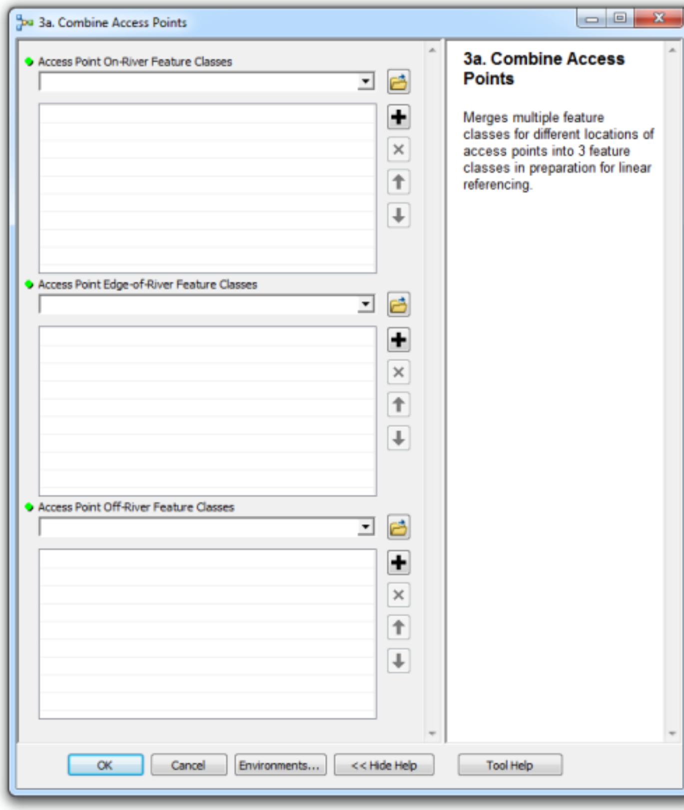


1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "2. Classify Point Locations."
2. For "Input Point Features," choose either the feature class prefixed "AccessPoint_NRRD_Raw…" or that prefixed "Campground_NRRD_Raw…" from the dropdown menu or navigate to its location.
3. In the next two boxes, click the SQL icon to the right of the box and create an expression that indicates which of your point features are on the river or at the edge of the river (note: all other points will be assigned to the "off-river" cagetory). If you do not have point features that are either on or at the edge of the river, leave that box blank. *Tip: if you have only one type of point location, use the IS NOT NULL buttons in the SQL dialog box to select all records (e.g. Access_Site_Name IS NOT NULL).*

3a. Combine Access Points tool

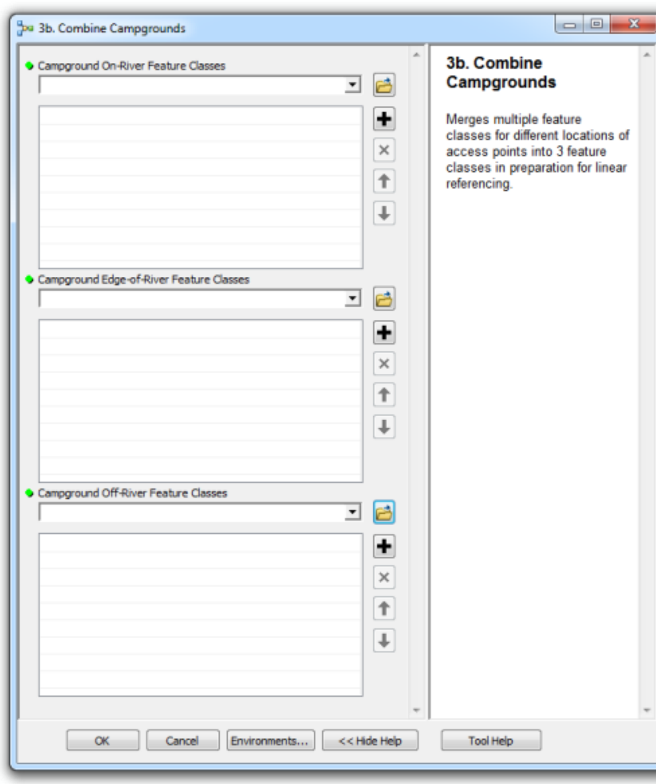
If you have entered data from multiple original feature classes, this tool will combine them all into one feature class for inputting into your linear referencing tool of choice. Note that even if you have only one original feature class from which you entered data, you still need to run this tool and specify the feature class that corresponds to each point location to set the stage for the linear referencing. Note also that each field is required. Even if you had only one location type, the previous tool ("classify point locations") created feature classes for each location type, and you need to input these (empty) feature classes into the dialog box.

1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "3a. Combine Access Points."
2. In the "Access Point On-River Feature Classes" dropdown box, enter ALL of the feature classes prefixed by "AccessPoint_1_OnRiver_..." one at a time. If you have only one, enter that. You can also find these feature classes by navigating to the "Raw_AccessPoints" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase.
3. Repeat this process in the "Access Point Edge-of-River Feature Classes" box for all feature classes prefixed by "AccessPoint_2_EdgeOfRiver_...".
4. Repeat this process in the "Access Point Off-River Feature Classes" box for all feature classes prefixed by "AccessPoint_3_OffRiver_...".
5. The output will be three feature classes located in the "PreppedData" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase, named "LinRefPrep_AccessPoint_OnRiver," "LinRefPrep_AccessPoint_OffRiver," and "LinRefPrep_AccessPoint_EdgeOfRiver."



3b. Combine Campgrounds tool

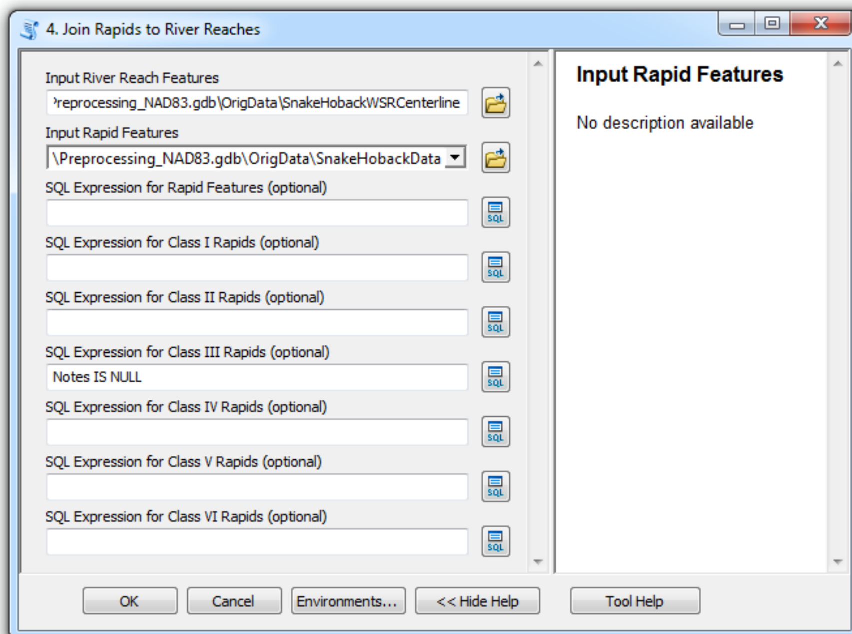
If you have entered data from multiple original feature classes, this tool will combine them all into one feature class for inputting into your linear referencing tool of choice. Note that even if you have only one original feature class from which you entered data, you still need to run this tool and specify the feature class that corresponds to each point location to set the stage for the linear referencing. Note also that each field is required. Even if you had only one location type, the previous tool ("classify point locations") created feature classes for each location type, and you need to input these (empty) feature classes into the dialog box.



1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "3b. Combine Campgrounds."
2. In the "Campground On-River Feature Classes" dropdown box, enter ALL of the feature classes prefixed by "Campground_1_OnRiver_…" one at a time. If you have only one, enter that. You can also find these feature classes by navigating to the "Raw_Campgrounds" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase.
3. Repeat this process in the "Campground Edge-of-River Feature Classes" box for all feature classes prefixed by "Campground_2_EdgeOfRiver_…".
4. Repeat this process in the "Campground Off-River Feature Classes" box for all feature classes prefixed by "Campground_3_OffRiver_…".
5. The output will be up to three feature classes located in the "PreppedData" feature dataset in the "Preprocessing_NAD83.gdb" geodatabase, named "LinRefPrep_Campground_OnRiver," "LinRefPrep_Campground_OffRiver," and "LinRefPrep_Campground_EdgeOfRiver."

4. Line feature class prep: OPTION 1: Join rapids to river reaches tool

If you have point data identifying rapids, use this tool to classify those rapids and join them to your river reach linear data. This is necessary to prepare your data so that rapids may be added to the NRRD database as part of your linear features. NOTE: If you do not have rapids as point data, choose OPTION 2 on the next page



1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "4. Join Rapids to River Reaches."
2. For "Input River Reach Features," choose your original line feature class from the dropdown menu or navigate to its location.
3. For "Input Rapid Features," choose your *original* (not processed by tools 1-3) point feature class that contains your rapids data.
4. In each of the next six boxes, click on the SQL icon to the right of the box and enter an expression that points to rapids of the class indicated. If you do not have any rapids in given class, leave that box blank.

OPTION 2: Prep River Reaches Tool

If you do NOT have point features that represent rapids, use this tool to prepare your line features for linear referencing. Simply open the tool and input your line feature class into the dialog box.

5. The output from either option will be a feature class located in the “PreppedData” feature dataset in the “Preprocessing_NAD83.gdb” geodatabase, named “LinRefPrep_RiverReach.”

D. Linear Referencing

At this point all of your data is prepped and ready to run through the linear referencing tool of your choice.

For line feature classes:

1. If you already use NHD flowlines for your river reach data, or if your river reach geometry is extremely close to the NHD flowline geometry, or if you do not have access to ArcGIS 10.2, you may opt to use the Alternate Linear Referencing tool, which is a pared-down version of the HEM tool and generally runs faster. However, you may have to update some fields in your data by hand. If your line data does not closely match the NHD and you use the Alternate tool, you may have many features to edit by hand, so budget time for that.
2. If your river reach geometry does not correspond closely to the NHD flowline geometry, and if you have access to ArcGIS 10.2, we recommend that you choose to use the Hydrography Event Management (HEM) tool, which is more thorough and has a built in process for updating all fields in the data.
3. Run your chosen linear referencing tool on “LinRefPrep_RiverReach,” located in the “PreppedData” feature dataset.

For point feature classes:

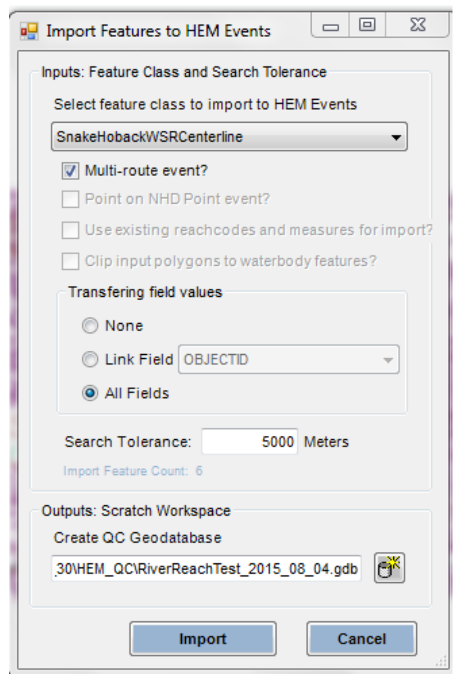
You will need to run a linear referencing tool on the following feature classes (if you have them) from the “PreppedData” feature dataset in the “Preprocessing_NAD83.gdb” geodatabase:

1. “LinRefPrep_AccessPoint_OnRiver” and “LinRefPrep_AccessPoint_EdgeOfRiver.”
2. “LinRefPrep_Campground_OnRiver” and “LinRefPrep_Campground_EdgeOfRiver.”
3. NOTE: you do NOT need to run it on the feature classes with the suffix “_OffRiver.”

	HEM tool	Alternate tool
Where to get the tool	download from USGS website	included with NRRD toolkit
ArcGIS version	10.1 or 10.2	10.3
Tool type	ArcGIS toolbar	tool in NRRD toolbox
Tool speed	slow to very slow	fast to slow
Complexity	complex	fairly simple
QA/QC	built-in	manual, time consuming
Output	identical	

HEM Tool

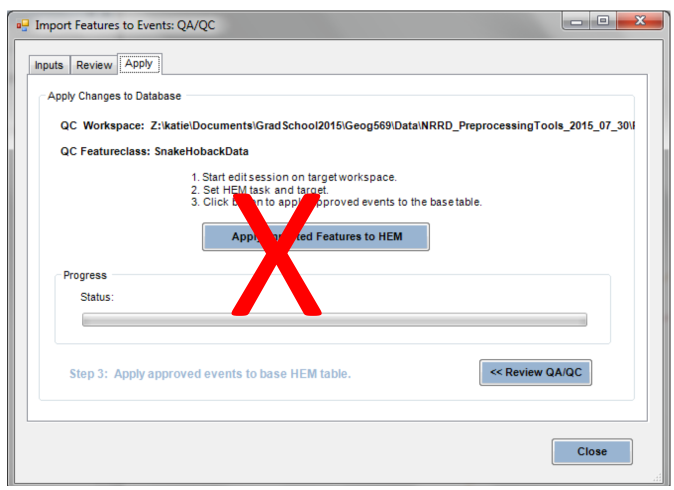
1. Download the Hydrography Event Management (HEM) Tool from USGS. *Tip: make sure you have ArcGIS version 10.2 or earlier available, or the HEM tool will not function properly*.
    a. Go to http://nhd.usgs.gov/tools.html and click on the link to the Hydrography Event Management Tool.
    b. Under "Tool Downloads," click on the link for "HEM 2.7 for ArcGIS 10.2.2."
    c. The tool will be downloaded in a zip file. Extract the files and run "setup.exe."
    d. Open a new map document in ArcMap. On the "Customize" tab at the top, choose "Toolbars" and check the box to the left of "Hydro Event Management Tools."
    e. The HEM toolbar should now be in your ArcGIS window.
2. On the HEM toolbar on the right side, click the "Edit Tools" button and choose "Import to Events."
3. IF YOU HAVE LINE DATA:
    a. Choose your feature class called "LinRefPrep_RiverReach" from the dropdown menu at the top.
    b. Check the box next to "Multi-route event?"
    c. Choose "All Fields" from the "Transferring field values" list
    d. Give the search tolerance a high value, we suggest 5,000.
    e. Navigate to the folder HEM_QC and name your new geodatabase "HEM_QC_1." Click "Import."



4. IF YOU HAVE POINT DATA:
    a. Choose your feature point feature class prefixed "LinRefPrep_" from the dropdown menu at the top.
    b. Do NOT check the box next to "Point on NHD event?"
    c. Choose "All Fields" from the "Transferring field values" list
    d. Give the search tolerance a high value, we suggest 5,000.
    e. Navigate to the folder HEM_QC and name your new geodatabase "HEM_QC_2." (If run multiple feature classes through the tool, make a new geodatabase for each one and number them sequentially.)
    f. Click "Import."

TIP: be patient, if you have many features this could take a very long time, even hours. Consider letting it run overnight.

5. You should now have a results window called "HEM: Import Features to HEM Events," and with luck it will tell you that all of your features were successfully imported and none were unsuccessful. This should be the case if you put a sufficiently high value in the "Search Tolerance" box above. Close the window. Note that if not all of your features were successfully imported, you will have the opportunity to import them during QA/QC.
6. Next, you will QA/QC your data. Back on the HEM toolbar, on the right side click the "Edit Tools" dropdown and click on "Import to Events QA/QC."
7. On the Inputs tab, navigate to the QC geodatabase you created above.
8. In the "Spatial Thresholds" box, enter another high value (such as 5,000) in the Distance (Meters) box. If you have a line feature class, enter "100%" in the "Length % Change" box. These high values will ensure that your features are auto-approved. If you'd prefer to manually approve your features, enter smaller values.
9. Click the "Review" button in the lower right-hand corner.
10. A box will pop up asking you if you want to start an edit session. Click "Yes."
11. The next pop up box explains what the thresholds you input mean. Click "OK."
12. On the "Review" tab, check the box next to "Show approved features" on the left side below the gray box. It should now list all of your features. If you wish to manually check your features, use the "flash" and "zoom" buttons on the left to locate each feature and check that it is associated with the correct river reach.
13. When you are satisfied, click the "Apply to Base Data" button in the lower right corner. It will ask you if you want to save edits and stop editing, click "Yes."
14. IMPORTANT: On the "Apply" tab, DO NOT click the "Apply Imported Features to HEM" button. This step is unnecessary for the NRRD workflow. Instead, just click "Close."



15. Open your new layer's attribute table. You will note that the following fields have been added to your table:

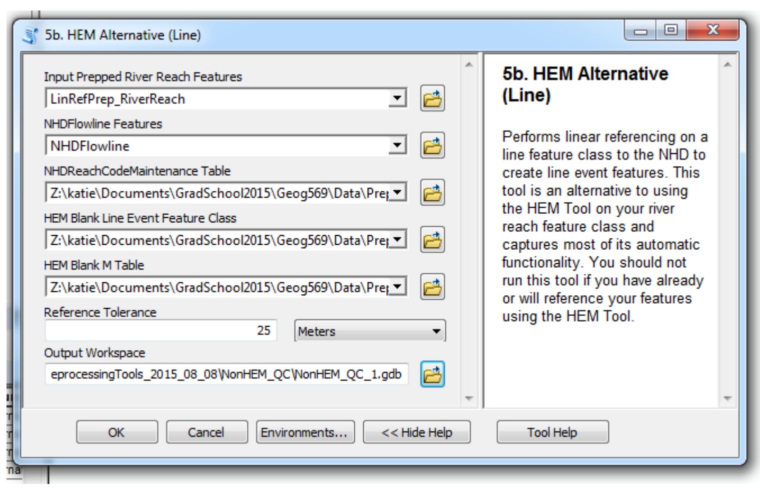
FieldName
Permanent_Identifier
EventDate
FReachCode
TReachCode
ReachResolution
Feature_Permanent_Identifier
FeatureClassRef
Source_Originator
Source_DataDesc
Source_FeatureID
FeatureDetailURL
FMeasure
TMeasure
EventOffset
EventType

16. Check your new table to be sure the fields are populated. Of critical importance are the Permanent_Identifier field and the ReachCode fields.
17. Your data is now in the necessary format to enable you to run the National River Recreation Database import toolset.

Alternate Linear Referencing Tool

Because of the complexity of the HEM tool, some people prefer not to use it. For those people, we provide a simple alternative. The trade-off is that some of the functionality of the HEM tool, particularly the QA/QC portion, is sacrificed, and therefore needs to be done manually. If you are not comfortable conducting QA/QC on your data (checking to make sure the features snapped to the correct NHD flowline, making necessary changes and updating the attribute table to reflect those changes), we recommend that you use the HEM tool instead.

1. Make sure your pre-processed feature classes (prefixed by “LinRefPrep_”) are in your map document.
2. Set up your output geodatabase:
    a. Navigate to the folder “HEMAlt_QC”in the “Preprocessing Tools” folder.
    b. There is one geodatabase, called “NonHEM_QC.gdb” in the folder. Right click on it and choose “copy.”
    c. Right click on the parent folder “HEMAlt_QC”and choose “paste.” You should now have a second geodatabase called “HEMAlt _QC_1.gdb.”
    d. YOU WILL NEED TO CREATE A SEPARATE GEODATABASE FOR EACH FEATURE CLASS YOU RUN THROUGH THE ALTERNATE LINEAR REFERENCING TOOL. They will be numbered sequentially (e.g. “HEMAlt _QC_1.gdb,” “HEMAlt _QC_2.gdb.”)
3. Navigate to the “NRRD_Preprocessing” toolbox.
4. IF YOU HAVE LINE DATA



    a. Open the tool named “5b. HEM Alternative (line).”
    b. In the “Input Prepped River Reach Features” box, choose “LinRefPrep_RiverReach” from the dropdown menu or navigate to its location.
    c. In the “NHDFlowline Features, choose “NHDFlowline” from the dropdown menu or navigate to the NHD flowline feature class for your region.
    d. Leave the next three boxes with their default values: these are pointing to intermediate locations for the tool.
    e. In the Reference Tolerance box, input the SMALLEST value practical for your data. This is the distance perpendicular to the NHD flowline that the tool will search for your input line

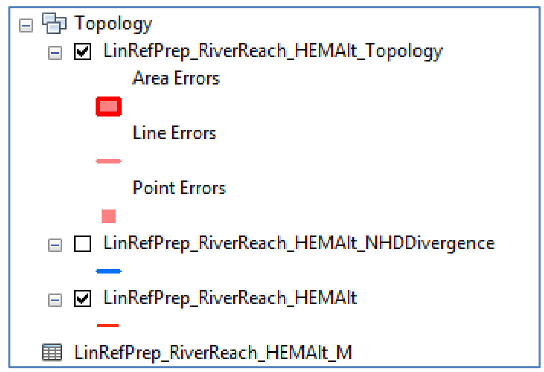
data. If you make the tolerance too small, the tool will not find parts of your river reach lines and you will have to edit them in manually. If you make the tolerance too large, the tool will get off track and start including extraneous tributaries from the NHD feature class. The latter problem takes longer to fix than the former, so we recommend the smallest tolerance that will still capture most of your data.

f. In the Output Workspace box, navigate to the copy of the NonHEM_QC_#.gdb you set up in step 2 above. Remember that this needs to be a new, clean geodatabase (so that the topology will function properly). Therefore, if for any reason you run the tool multiple times on the same data, you must either delete the NonHEM_QC_#.gdb from previous tool runs OR create a new, sequentially numbered NonHEM_QC_#.gdb file.

g. Click "OK." This tool is faster than the HEM tool, but it can still take many minutes to run, so be patient.

5. QA/QC FOR RIVER REACH LINES

a. When the tool is done running, you should have a few things added to your map document table of contents:



i. *LinRefPrep_RiverReach_HEMAlt_Topology* shows you places where the output feature class differs from the NHD flowline. The Point Errors show you all the line endpoints. You should expect these to be at the ends of your rivers. However, if there are any in the middle of river reaches, they will have to be addressed

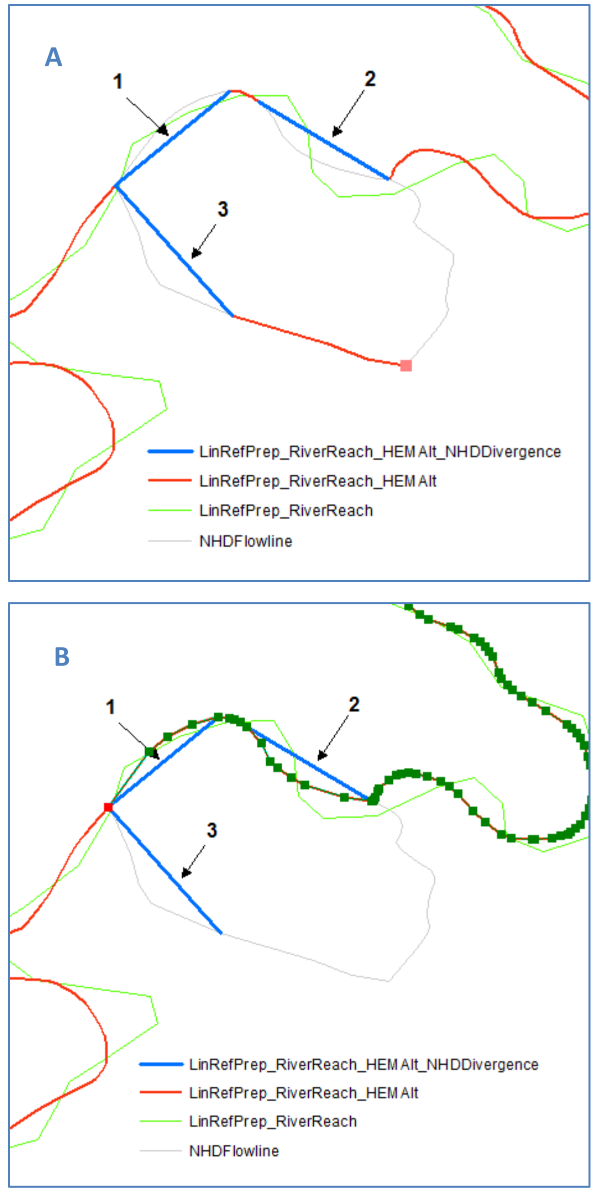
ii. *LinRefPrep_RiverReach_HEMAlt_NHDDivergence* pinpoints the locations where your lines diverge from the NHD flowline. You will need to check these areas to see if the divergence is of a tolerable margin or if you wish to edit them. These are often associated with tributary confluences.

iii. *LinRefPrep_RiverReach_HEMAlt* is your output feature class. This is the feature class you will edit.

b. Open an edit session on your output feature class (*LinRefPrep_RiverReach_HEMAlt*).

c. You will now inspect and/or edit each one of the instances of divergence. Open the attribute table of *LinRefPrep_RiverReach_HEMAlt_NHDDivergence*. Highlight the first record so it is selected, then zoom to it. You will go through the entire attribute table, checking one instance at a time. Alternatively, if you only have a few instances, you can zoom to each one directly.

Figure 55: Example of divergence editing



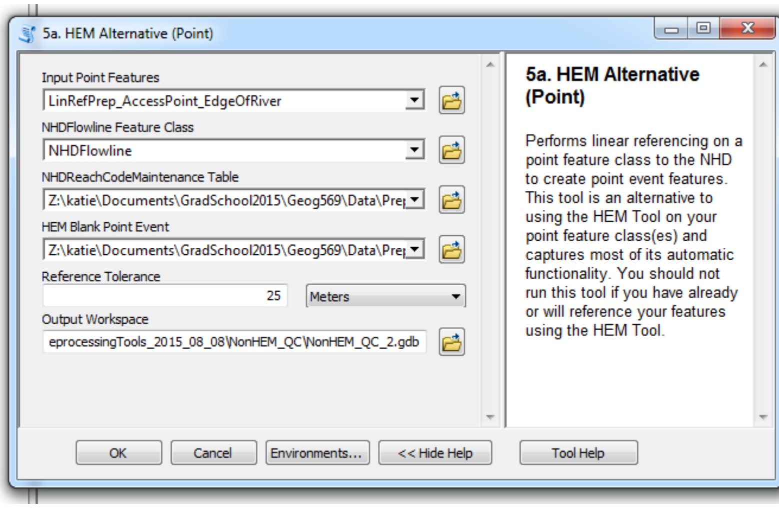
d. In figure 3A, divergence lines (in blue) 1 and 2 are instances where the input line (in green) strayed too far from the NHD flowline (in gray), so the output line (in red) snapped straight between two points. In these instances, you should add vertices to the output line and manually move it to match the NHD flowline. Line 3 is an instance where the tool added a tributary (or in this case a side channel) to the output. You should delete these lines. Notice the red square at the end of the line stub – once you delete this stub, the error square will disappear. Figure 3B shows the same three instances after editing.

e. If you need to move a line from one NHD flowline to another (if your data snapped to the wrong flowline), you will have to manually update the reach code in your table.
   i. Highlight the correct NHD flowline reach.
   ii. Open the NHD attribute table and copy the value in the "ReachCode" field of the highlighted reach.
   iii. Highlight your corrected reach in your *LinRefPrep_RiverReach_HEMAlt* feature class.
   iv. Open its attribute table (one reach should be highlighted), right click on the "ReachCode" field, choose "field calculator," and paste the copied reach code into the dialog box.

6. IF YOU HAVE POINT DATA
   a. Open the tool named "5a. HEM Alternative (point)."
   b. In the "Input Prepped River Reach Features" box, choose your point feature prefixed by "LinRefPrep_" from the dropdown menu or navigate to its location. REMEMBER TO RUN THE TOOL ONCE FOR EACH ON-RIVER AND EDGE-OF-RIVER POINT FEATURE CLASS YOU HAVE.
   c. In the "NHDFlowline Features, choose "NHDFlowline" from the dropdown menu or navigate to the NHD flowline feature class for your region.

d. Leave the next two boxes with their default values: these are pointing to intermediate locations for the tool.

e. In the Reference Tolerance box, input a value large enough to capture all of your on-river or edge-of-river points. Remember that you do not need to run off-river points through this tool.



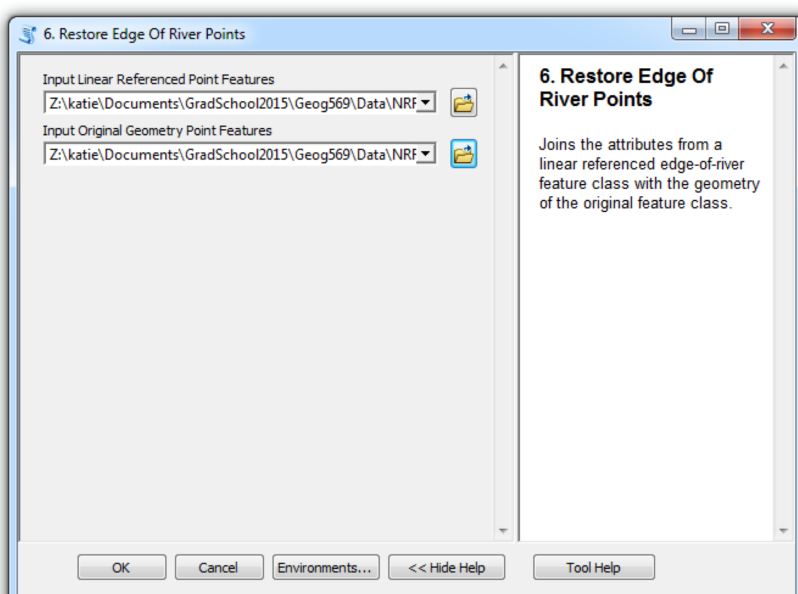
f. In the Output Workspace box, navigate to the copy of the NonHEM_QC_#.gdb you set up in step 2 above. Remember that this needs to be a new, clean geodatabase (so that the topology will function properly). Each new point feature class needs its own new NonHEM_QC_#.gdb file. If for any reason you run the tool multiple times on the same data, you must either delete the NonHEM_QC_#.gdb from previous tool runs OR create a new, sequentially numbered NonHEM_QC_#.gdb file.

g. Click "OK." This tool is faster than the HEM tool, but it can still take many minutes to run, so be patient.

h. QA/QC your point data to be sure the points snapped to the correct NHD flowline, and if there are errors, move the points to the correct line and update the associated ReachCode values.

E. Post Linear Referencing Clean Up Tools

Now that you have referenced your point and line data with the National Hydrography Dataset, you are almost ready to join your data to the NRRD. First you need to run two more tools to finalize the preparations.



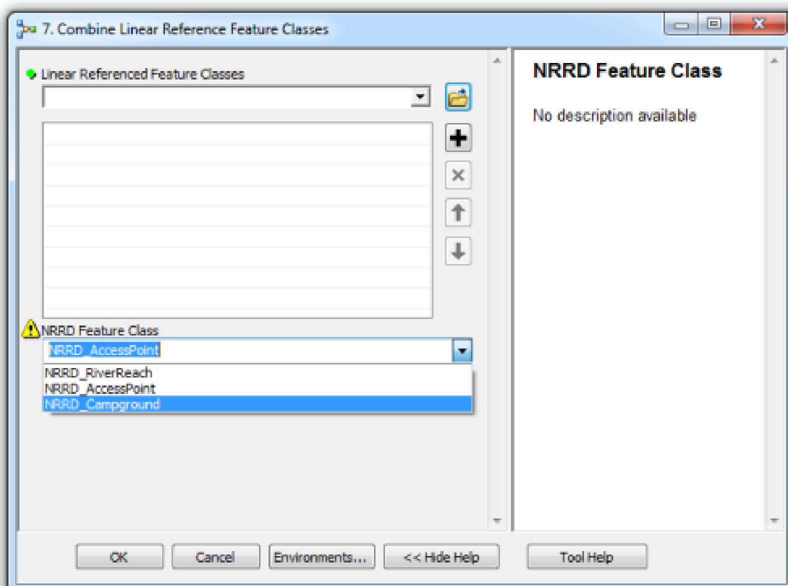
6. Restore Edge Of River Points tool

During the linear referencing, your features were all snapped to the centerline of the NHD flowline feature class. This is fine for river reaches and on-river points, but in order to preserve the geometry of the edge-of-river points, you need to join the linear referenced points back to their original locations.

1. Navigate to the "NRRD_Preprocessing" toolbox in ArcCatalog. Double click on the tool labeled "6. Restore Edge Of River Points."
2. In the box labeled "Input Linear Referenced Point Features," choose from the dropdown box or navigate to the output feature class you got from running your "LinRefPrep_AccessPoint_EdgeOfRiver" and/or "LinRefPrep_Campground_EdgeOfRiver" feature classes through the linear referencing tool (e.g. for the Alternative Linear Reference tool, the output feature class may be named "LinRefPrep_AccessSite_EdgeOfRiver_HEMAlt").
3. In the box labeled "Input Original Geometry Point Features," choose from the dropdown box or navigate to the pre-linear referencing "LinRefPrep_AccessPoint_EdgeOfRiver" and/or "LinRefPrep_Campground_EdgeOfRiver" feature classes, located in the "PreppedData" feature dataset.
4. The output will be written to the geodatabase "NRRD_BeginPostprocessing.gdb" and given the same names prefixed by "LinRefPrep_...".

*7. Combine Linear Referenced Feature Classes: **REQUIRED***

This tool collects all of your output data into three feature classes that correspond to the three feature classes in the NRRD. You will therefore need to run it up to three times, once each for river reach data, access site data, and campground data.



1. For "Linear Referenced Feature Classes," choose from the dropdown box or navigate to all the feature classes that are associated with a given data type.
2. For point data:
   a. Be sure to include the output from tool No. 6, "Restore Edge of River Points," instead of using the Edge of River point feature class from the output of the linear reference tool.
   b. On-River point data should be entered from the output of the linear referencing tool.
   c. Off-River point data was not run through the linear referencing tool, and therefore you should use the feature classes named "LinRefPrep_AccessPoint_OffRiver" or "LinRefPrep_Campground_OffRiver" located in the "PreppedData" feature dataset.
3. For line data:
   a. You likely have only one line feature class for river reaches, but you should still use this tool to transfer it to the next step.

b. Note that you will be required to enter an "M table." This is the "Reach Code Maintenance" table that is associated with the NHD and should be located in your HEM_QC or HEMAlt_QC geodatabase. You can also find it associated with your original NHD data. *Tip: if you for some reason are NOT running a linear referencing tool on your line data, use the Reach Code Maintenance table located in your original NHD geodatabase. Otherwise, use the one associated with your QC data.*

4. The output of this tool will be up to three feature classes located in the "NRRD_BeginPostprocessing.gdb" geodatabase: "NRRD_AccessPoint," "NRRD_Campground," and "NRRD_RiverReach."

YOUR DATA ARE NOW READY TO BE JOINED TO THE NRRD.

F. NRRD Data Import Toolset

The NRRD Data Import Toolset contains one toolbox for each feature class in the NRRD. Each contains a series of model tools that sort your data into the attribute fields for that feature class, with generally one tool per attribute field. You can choose the attributes for which you have comparable data and run only those tools. These instructions walk you through each of the tools in each toolbox.

Tip: Before using any of these tools, open ArcGIS and add your feature classes (the output of the HEM/alternate linear referencing tool as detailed in the "Pre-tool steps" section) to a blank map.

Floatable River Reaches Toolbox

Table 1 lists all of the fields in the feature class RiverReaches. This is a line feature class. You will transform your data to mesh with the fields in this feature class. The following sections provide step by step instructions on how to run the tools associated with each of these fields to transform your data for the NRRD. If you do not have data for one of the fields and cannot easily create it, you may skip the associated tool. Note that the output of one tool is used as the input for the next, so it makes most sense to run through the tools in order. Each tool adds its output to your map document, so you will be able to choose the correct input feature class for each tool from a dropdown list. Note also that the last tool described is required and will write your data into a new feature class that is ready for submission to the NRRD.

Tools in the River Reaches toolbox

River Reach - 01. Stream Difficulty tool
River Reach - 02. Website tool
River Reach - 03. Difficulty Outlier tool
River Reach - 04. Trip Length tool
River Reach - 05. State
River Reach - 06. Camping, Fishing and Permits
River Reach - 07. Text Fields tool
River Reach - 08. Recreation Type tool
River Reach - 09. Management Type tool
River Reach - 10. Final Append **REQUIRED**

Table 6: Fields in the NRRD River Reaches feature class

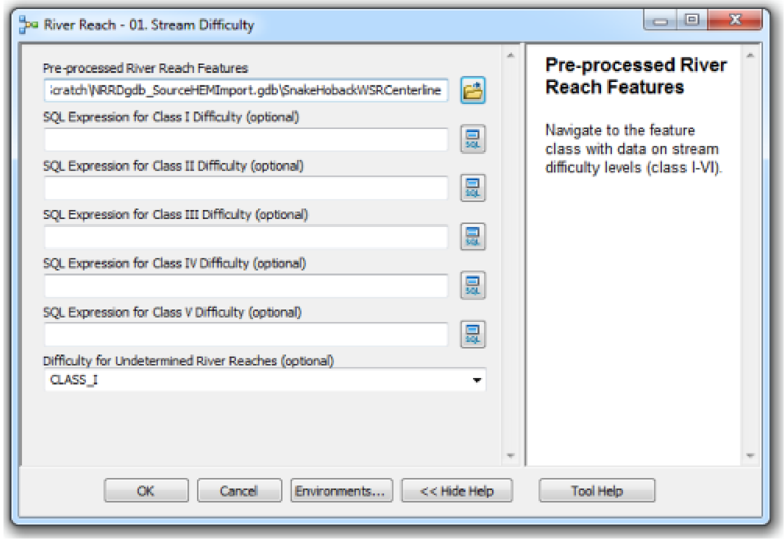
FieldName	Type	Length	Description	Example
Permanent_Identifier	*String*	*40*	*Fields added by the HEM or alternate linear referencing tool*	
EventDate	*Date*	*8*		
FReachCode	*String*	*14*		
TReachCode	*String*	*14*		
ReachResolution	*Integer*	*4*		
Feature_Permanent_Identifier	*String*	*40*		
FeatureClassRef	*Integer*	*4*		
Source_Originator	*String*	*130*		
Source_DataDesc	*String*	*100*		
Source_FeatureID	*String*	*100*		
FeatureDetailURL	*String*	*255*		
FMeasure	*Double*	*8*		
TMeasure	*Double*	*8*		
EventOffset	*Double*	*8*		
EventType	*Integer*	*4*		
GNIS_River_Name	String	35	Name of river or creek. Use the same naming as in the field called GNIS_name on the National Hydrography Data set.	Colorado Escalante Creek
Upstream_Point	String	50	Description of upstream point of the reach, typically the put-in. In the case of water trails, this is the beginning of the water trail.	Horseshoe Bend Put-in
Downstream_Point	String	50	Description of downstream point of the reach, typically the take-out. In the case of water trails, this is the end of the water trail.	Glacier Creek Take-out
State	SmallInteger	2	Primary State associated with the physical location--or beginning -- of the river section. The data is separated by state, and collected in that manner, two capital letters representing state.	CO
Description	String	254	Description of the river reach that includes the information about the area and recreational experience	
SectionID	*String*	*38*	*Automated*	
Section_Name_1	String	200	Name of the river section based on floatable, boatable use (i.e., the socially-used name of the section; e.g., sections in American Whitewater database, sections in printed or electronic boating guides). This name could be a commonly-used name for the section or could be named simply by the put-in to the take-out, if there is no name independent of those (most sections are named as per the latter)". The sections are in geographical order (upstream to downstream). In case of water trails, this would be the water trail name.	

FieldName	Type	Length	Description	Example
Regulations	String	254	Access or management-related notes that managing agency wants to communicate to public including any regulations or restrictions to access, as well as joint management. (note: this field may be difficult to manage; will "test" it and may use just a Weblink).	Private Access Only; Managed Jointly; closed to boating in August
Fishing	SmallInteger	2	Fishing allowed, yes/no	
Camping	SmallInteger	2	Camping allowed, yes/no	
Difficulty	SmallInteger	2	Difficulty as defined by the International Scale of River Difficulty, Class I-VI. Roman numerals I-VI are used, either individually, or in a range, separated by a dash, e.g. I-II.	I,I-II, II, II-III, III, III-IV, IV, IV-V, V, VI
Permit	SmallInteger	2	Is a Special Area Permit (private permit) needed for private boating publics to run river? (Yes/No) Permit types can include: self issue, mail in, lottery, 1st come/1st serve, Boaterpass.	
Permit_Site	String	254	URL to Website that lists information about permit or to more detailed regulations in places where there is no permit	
Website	String	254	Link to local website	
Miles	Double	8	Trip length in miles	
Trip_Length	SmallInteger	2	Trip duration	1=partial day; 2=full day; 3= one night; 4= multiday
Management	String	35	Agency with jurisdiction over this reach, if multiple agencies, list together with a comma in between	National Park Service
Whitewater	SmallInteger	2	Is the difficulty on the International Scale of River Difficulty ≥ II, yes/no	
AW_Link	String	254	American Whitewater website link	
Water_Trail	SmallInteger	2	Water trail, yes/no	
Source_s_	String	300	Data source	
Last_Edited	*Date*	*8*	*Automated*	
State2	SmallInteger	2	Second state in which river is located	
Website2	String	150	Second website for river reach information	
Photo	String	100		
Caption	String	250		
WSR_Section	SmallInteger	2		
Difficulty_Outliers	SmallInteger	2	Maximum difficulty encountered on reach rated lower on the International Scale of River Difficulty (e.g. for a level IV rapid on an otherwise level II reach)	
Address	String	200		
Shape_Leng	Double	8		
Management_2	String	35		
Management_3	String	35		

River Reach - 01. Stream Difficulty tool

This tool translates your data into the format needed to append it to the field "Difficulty" in the NRRD. The level of difficulty is the standard class system, with classes from I to VI.

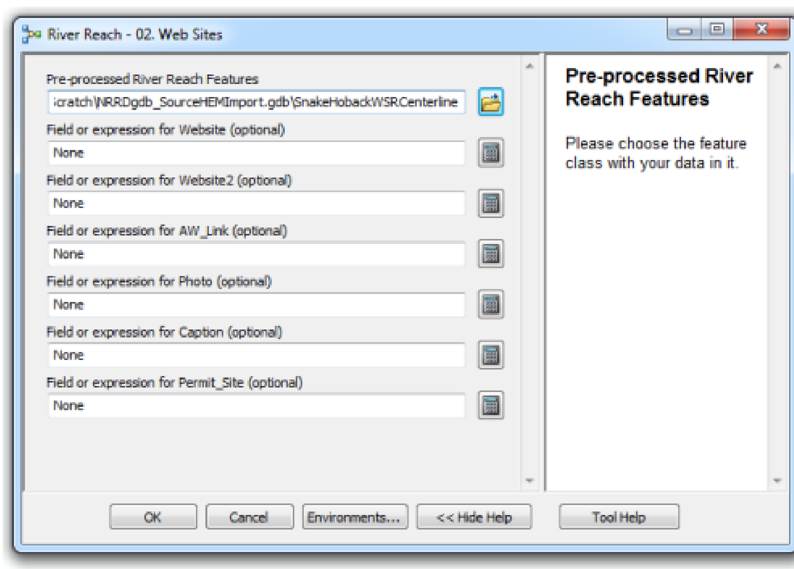
1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 01. Stream Difficulty." Wait for the dialog box to open.
2. In the "Pre-processed River Reach Features box, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the "SQL" button to the right and create an expression that points to the data in your feature class that corresponds to each level of difficulty. For example, if you have an attribute field called "CLASS" that lists rivers as "recreational," "challenging" and "difficult," choose "CLASS = 'recreational'" for Class I difficulty, "CLASS = 'challenging'" for Class III difficulty, and so forth. If you do not have any river reaches in your data that correspond to a given difficulty class, leave that box blank.
4. In the "Difficulty for Undetermined River Reaches" box, choose the class that will be assigned to any river reaches not processed in the previous boxes. The default is Class I.
5. Note that it takes a few seconds after each entry for it to appear in the dialog box.



River Reach - 02. Website tool

This tool allows you to indicate where in your data there are fields with web addresses in them.

1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 02. Web Sites." Wait for the dialog box to open.
2. In the "Pre-processed River Reach Features" box, if you have run tool No. 1, choose the feature class from the drop-down list that begins "RiverReach_1_". If this is the first tool you're running, choose your feature class from the drop



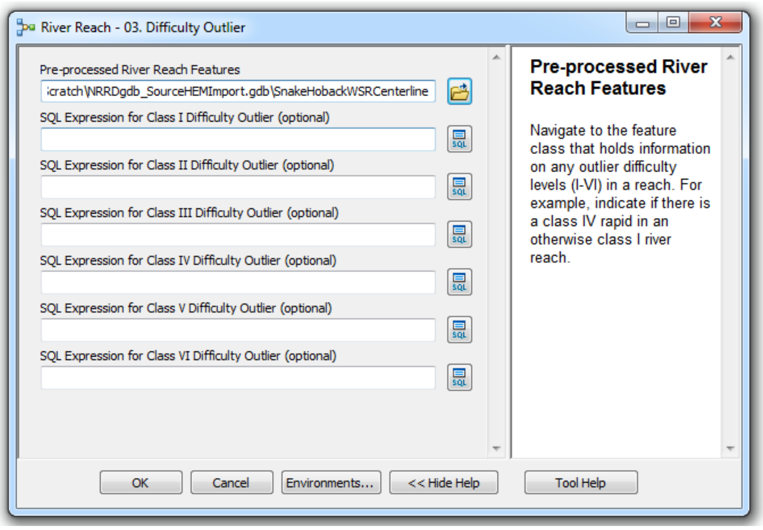
down menu, or navigate to its location.

3. For each of the next boxes, click the field calculator icon to the right of the box and choose the field in your data that corresponds to the type of website or link called for by the NRRD field. If you do not have a field containing website information of a given type, leave the "None" in the box and the NRRD field will be calculated to null.
4. Note that it takes a few seconds after each entry for the dialog box to register your choice.

River Reach - 03. Difficulty Outlier tool

This tool captures data that indicate the greatest level of difficulty encountered on a river reach, for example, class III rapids on a class I reach.

1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 03. Difficulty Outlier." Wait for the dialog box to open.

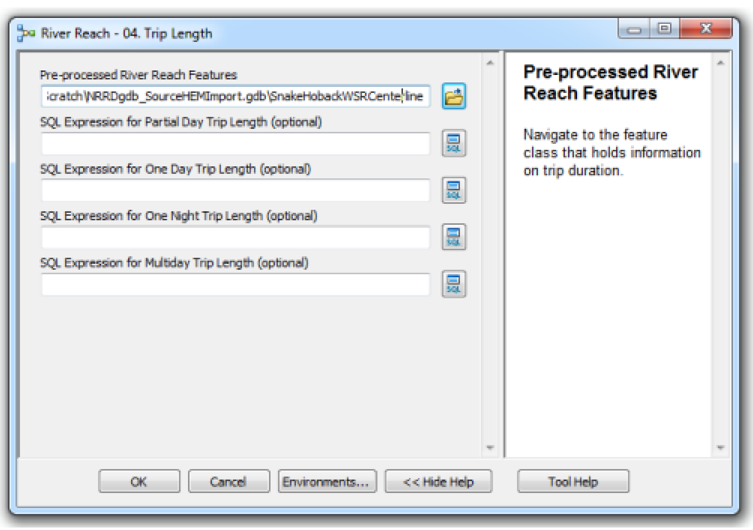


2. In the "Pre-processed River Reach Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "RiverReach_2_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the "SQL" button to the right and create an expression that points to the data in your feature class that corresponds to each level of difficulty. For example, if you have an attribute field called "Rapids" that describes the difficulty of rapids in some river reaches, choose "Rapids = 'III'" for the Class III difficulty and so forth. If you do not have any outliers in your data that correspond to a given difficulty class, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

River Reach - 04. Trip Length tool

This tool defines the duration of a trip on a given river reach as 1. Partial Day, 2. One Day, 3. One Night, or 4. Multiday.

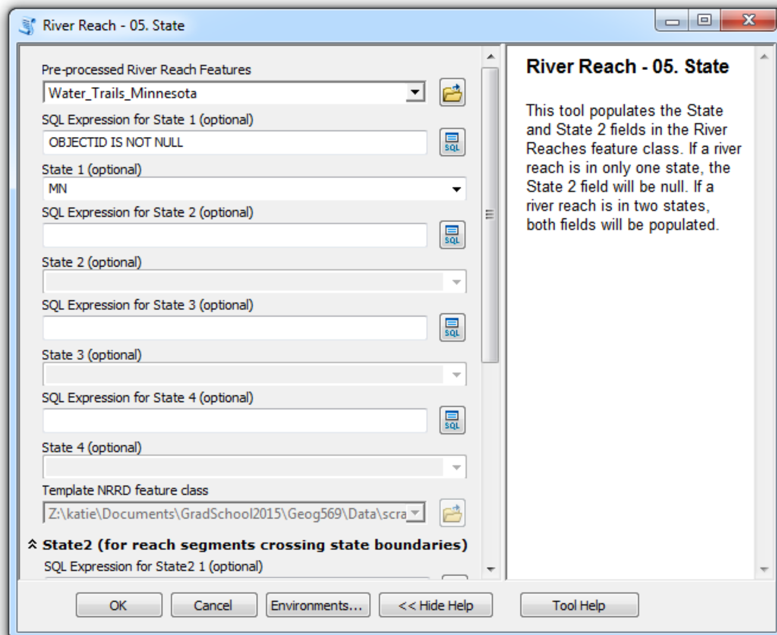
1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 04. Trip Length." Wait for the dialog box to open.



2. In the “Pre-processed River Reach Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “RiverReach_3_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the “SQL” button to the right and create an expression that points to the data in your feature class that corresponds to each trip length. For example, if you have an attribute field called “Duration” that describes the length of trips as “day” or “overnight” , choose “Duration = ‘day’” for One Day Trip Length, and so forth. If you do not have any trips in your data that correspond to a given trip length, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

River Reach - 05. State

This tool identifies the state where your river is located.



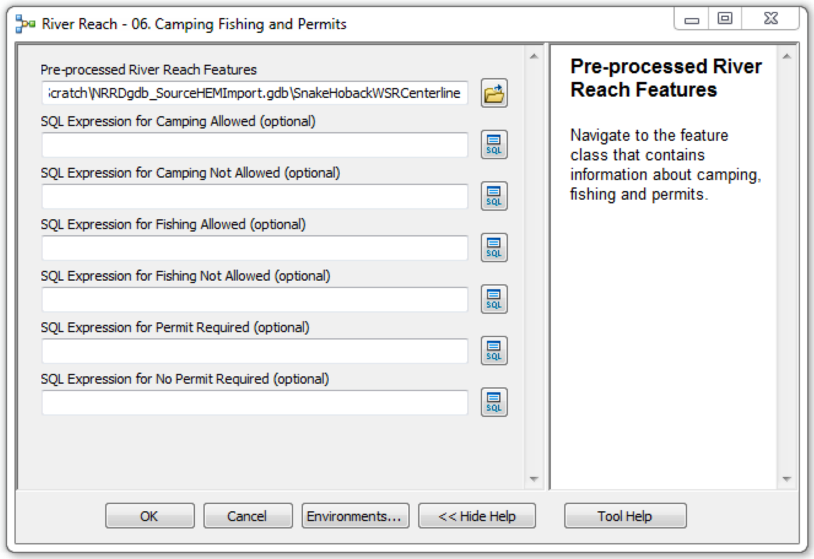
1. Open the “RiverReaches” toolbox. Double click on the tool labelled “River Reach - 05. State.” Wait for the dialog box to open.
2. In the “Pre-processed River Reach Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “RiverReach_4_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your feature class from the drop down menu, or navigate to its location.
3. Once you have input your feature class, you will notice that the box titled “SQL Expression for State 1” is set to “OBJECT ID IS NOT NULL.” This is the default value, and if all of your river reaches are in the same state, you can leave the default value and select your state from the dropdown list titled “State 1.” *Tip: double check to be sure you have a field called “OBJECTID.” If your ID field is called something else, use the “SQL” button on the right to create an expression using the “IS NOT NULL” syntax.*
4. If you have river reaches in multiple states, use the “SQL” button on the right and create an expression that indicates which river reaches are in the first state, then choose that state from the dropdown list. Repeat this with up to 4 states. Note that the State 2 through 4 drop down lists are grayed out until you enter an SQL expression pointing toward that state in your data.
5. At the bottom of the dialog box is an additional set of boxes, accessed by clicking on the double arrow next to “State2.” These boxes populate a second state field in the River Reaches feature

class and are intended to accommodate river reaches that span more than one state. If you have any river reaches that cross state lines, use these boxes to indicate the second state(s).

River Reach - 06. Camping, Fishing and Permits

This tool populates three yes/no fields, indicating whether or not camping or fishing is allowed, and whether or not permits are required.

1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 06. Camping, Fishing and Permits." Wait for the dialog box to open.
2. In the "Pre-processed River Reach Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "RiverReach_5_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the "SQL" button to the right and create an expression that points to the data in your feature class that corresponds to the value indicated. For example, if you have a field that identifies permit types, compose an expression for "Permit Required" that encompasses all permit types.

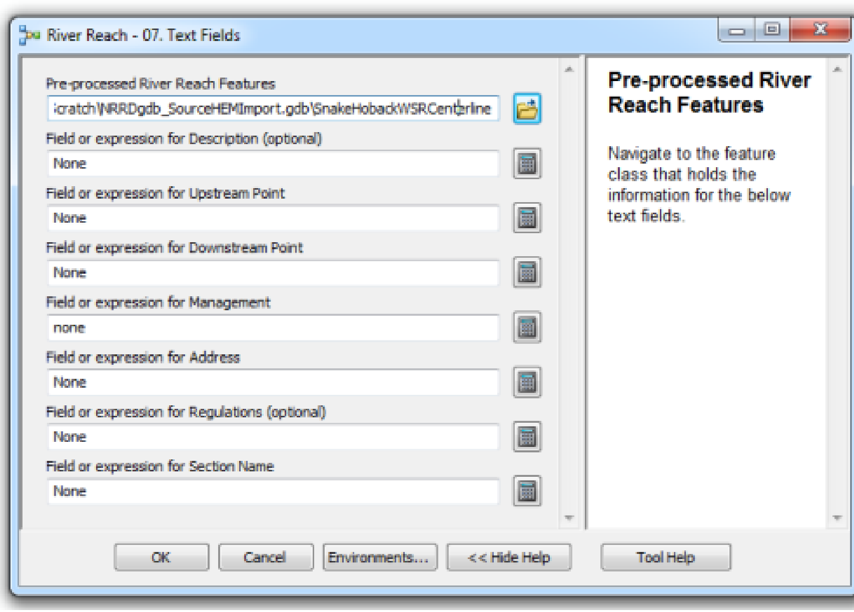


4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

River Reach - 07. Text Fields tool

This tool populates several text fields in the NRRD database with comparable data. Text fields include reach description, upstream point, downstream point, management, address, regulations and section name. For descriptions for each of the NRRD fields indicated, refer to table 1 on pages 20-21.

1. Open the "RiverReaches" toolbox. Double click on the tool labelled "River Reach - 07. Text Fields."

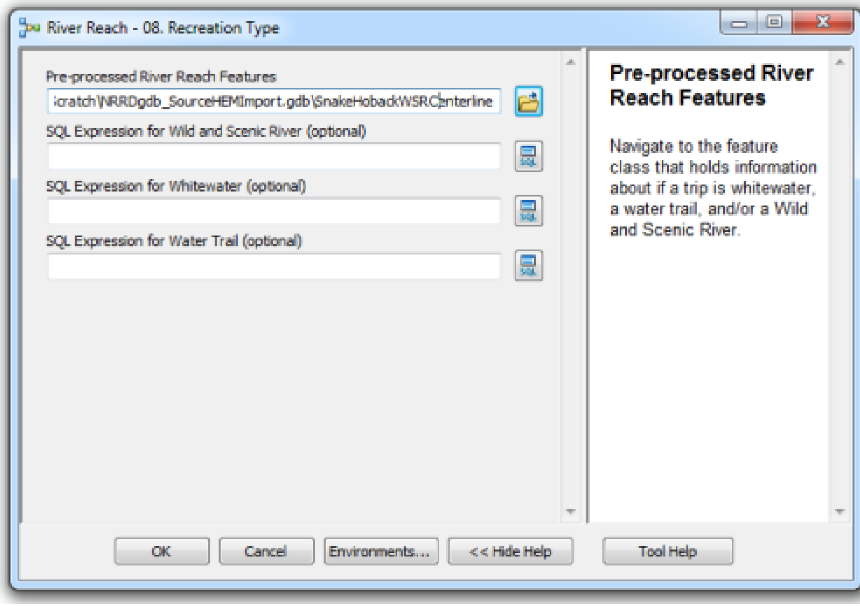


Wait for the dialog box to open.

2. In the “Pre-processed River Reach Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “RiverReach_6_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the field calculator icon to the right of the box and choose the field in your data that corresponds to the text data indicated.
4. Note that it takes a few seconds after each entry for the dialog box to register your choice.

River Reach - 08. Recreation Type tool

This tool populates three yes/no fields, each indicating whether or not a river reach is considered whitewater, part of a water trail, or a Wild and Scenic River.

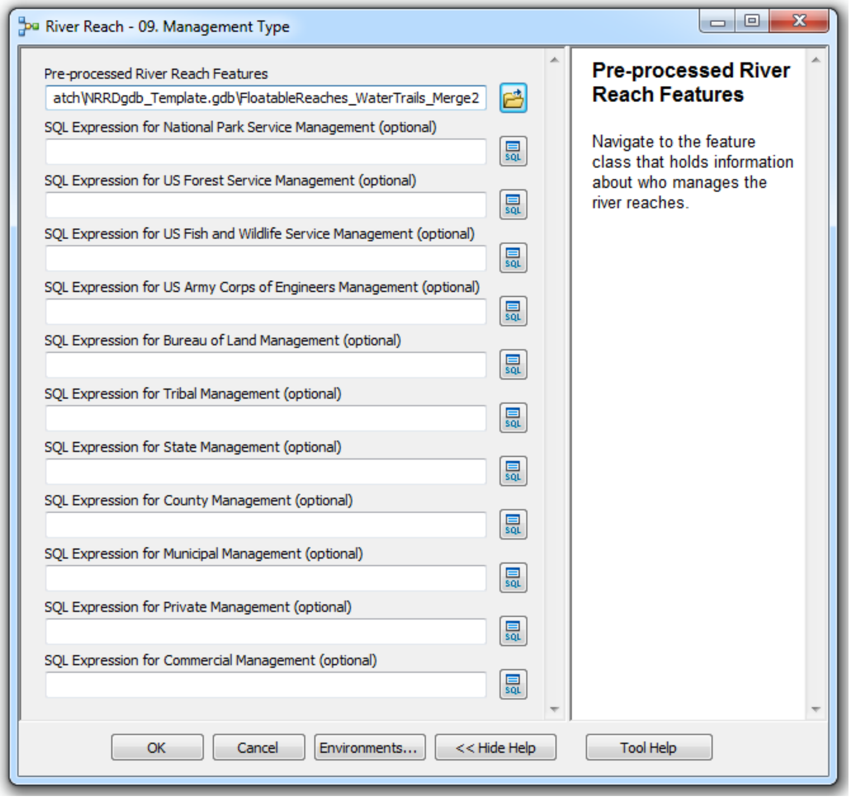


1. Open the “RiverReaches” toolbox. Double click on the tool labelled “River Reach - 08. Recreation Type.” Wait for the dialog box to open.
2. In the “Pre-processed River Reach Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “RiverReach_7_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that indicates which river reaches ARE a given recreation type. All records that are not indicated as being of a certain recreation type will be populated with a “no” value. A single river reach could be one, two or all three recreation types.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

River Reach - 09. Management Type tool

This tool sorts river management into one of eleven pre-set management types: National Park Service, US Forest Service, US Fish and Wildlife Service, US Army Corps of Engineers, Bureau of Land Management, Tribal, State, County, Municipal, Private, or Commercial. Note that there is a text field for entering specific management entities in the “Text Fields” tool (No. 7). So, for example, enter a field that has information such as “Minnesota Department of Natural Resources” or “Bridger-Teton National Forest” in the text fields tool, but “State” and “US Forest Service” here.

1. Open the “RiverReaches” toolbox. Double click on the tool labelled “River Reach - 09. Management Type.” Wait for the dialog box to open.

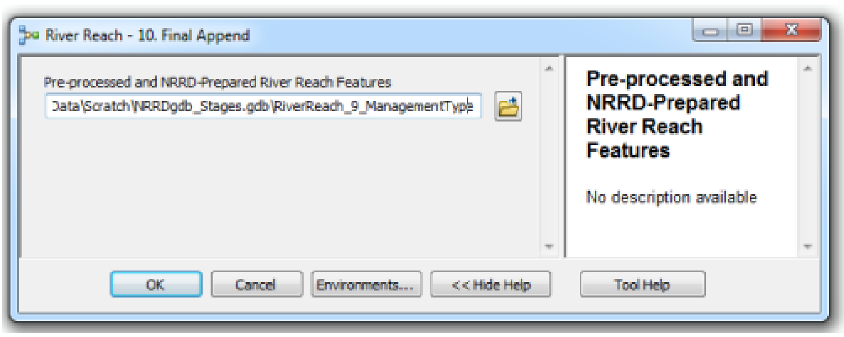


2. In the “Pre-processed River Reach Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “RiverReach_8_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL expression icon to the right of the box and enter an expression that identifies the management type indicated. If there are no reaches under a given management type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

River Reach - 10. Final Append **REQUIRED**

Once you have completed transforming your data with those tools that are applicable, run the “Final Append” tool to write your data to a new feature class that is ready to submit to the NRRD.

1. Open the “RiverReaches” toolbox. Double click on the tool labelled “River Reach - 10. Final Append.” Wait for the dialog box to open.



2. In the “Pre-processed River Reach Features”

box, choose the feature class from the drop-down list that begins “RiverReach_9_” (or the number of the most recent tool you ran).

3. Note that this tool will discard any fields that are not part of the NHD (HEM) or NRRD schema. If you have data you wish to include, it must be already translated into one of those fields using the tools in this toolkit.

AT THIS POINT Y OU HAVE FINISHED PREPARING YOUR RIVER REACH/LINE DATA FOR SUBMISSION TO THE NRRD. NEXT YOU WILL HAVE THE OPPORTUNITY TO TRANSFER INFORMATION ABOUT ACCESS SITES.

Access Sites Toolbox

The following tools populate fields in the NRRD feature class Access_Sites with comparable data from your pre-processed dataset. If you have no data about access sites, you can skip this entire section. If you do not have comparable data for a given field, you can skip that tool. If you use any of the tools in this section, you MUST run the tool "Access Sites 8. – Final Append." Below is a list of the tools available. Each corresponds to a field or series of similar fields in the NRRD feature class. Table X describes these fields.

Tools in the Access Sites toolbox

Access Site - 1. State tool

Access Site - 2. Parking tool

Access Site - 3. River Left or Right tool

Access Site - 4. Management Type tool

Access Site - 5. Access Type tool

Access Site - 6. Yes-No Fields tool

Access Site - 7. Text Fields tool

*Access Site - 8. Final Append **REQUIRED***

Table 7: Fields in the NRRD Access Sites feature class

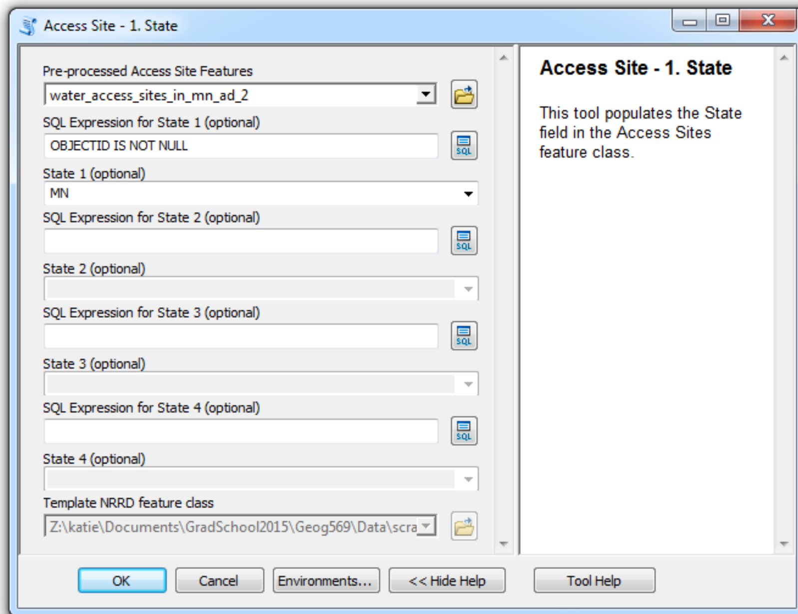
FieldName	Type	Length	Description	Example
Site_Name	String	50	Name of the access site	Horseshoe Bend
Access_Site_Management2	String	50	More detailed name of managing entity	City of Grand Junction
RL_or_RR	SmallInteger	2	Identify if the access site is River Left (RL) or River Right (RR)	River Right, River Left
Access_Type	SmallInteger	2	Yes for day-use parking, No for no parking, and Overnight for overnight parking	Yes, No, Overnight
Parking	String	50	Identify whether the access is one of the following: boat launch, carry-in, boat-in, fly-in	carry-in, boat launch, boat-in, fly-in
River_Mile	String	15	The river mile that the access site is located on (defined by USGS or state)	RM 5
Put_In	SmallInteger	2	Yes/No field to indicate if this is site is used as a put-in	Yes, No
Take_Out	SmallInteger	2	Yes/No field to indicate if this field is used as a take-out	Yes, No
Fee	SmallInteger	2	Yes/No field to indicate if this site has a fee	Yes, No

FieldName	Type	Length	Description	Example
Restrictions	String	250	Other information to share with the public about the access site that doesn't fit into the fields above	Closed from April - May for bird nesting
Website_Link1	String	300	Link to the website for more information/contact information **about this access site** (if not specific to the site, no need to double-up on the web link from section attributes)	
Emergency	SmallInteger	2	Yes/No field answering the question of is this site only available for emergencies? (e.g., to evacuation route)	Yes, No
Observ_Area	SmallInteger	2	Yes/No filed indicated if this access site provides an observation area for people to watch boaters. Would need to have parking, a path, and observation area.	Yes, No
*Source_s_*	*String*	*300*	*Automated*	
River_Name	String	25	Name of river	Rogue
Restrooms	SmallInteger	2	Yes/No field to indicate if this site has restrooms	
Last_Edited	*Date*	*8*	*Automated*	
State	SmallInteger	2	Name of State	Washington
Directions	String	100	Directions to the access site - automated link to google maps or something similar	*AW does this on their directions link*
Management_Type	SmallInteger	2	Type of the managing entity that manages that site	USFS, BLM, USFWS, USACE, State, County, Tribal, Private
Website_Link2	String	300		
Permanent_Identifier	*String*	*40*	*Fields added by the HEM or alternate linear referencing tool*	
EventDate	*Date*	*8*		
ReachCode	*String*	*14*		
ReachSMDate	*Date*	*8*		
ReachResolution	*Integer*	*4*		
Feature_Permanent_Identifier	*String*	*40*		
FeatureClassRef	*Integer*	*4*		
Source_DataDesc	*String*	*100*		
Source_FeatureID	*String*	*100*		
FeatureDetailURL	*String*	*255*		
Measure	*Double*	*8*		
EventOffset	*Double*	*8*		
EventType	*Integer*	*4*		

Access Site 1. – State tool

This tool identifies the state(s) where your access sites are located.

1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 1. State." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, choose your pre-processed feature class from the drop down menu, or navigate to its location.

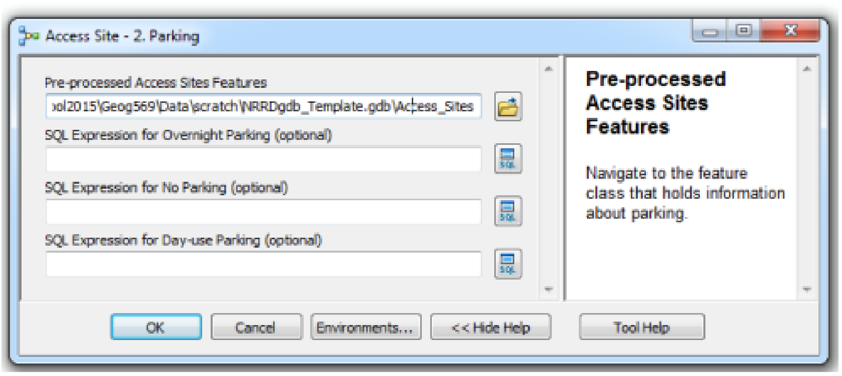


3. Once you have input your feature class, you will notice that the box titled "SQL Expression for State 1" is set to "OBJECT ID IS NOT NULL." This is the default value, and if all of access sites are in the same state, you can leave the default value and select your state from the dropdown list titled "State 1." *Tip: double check to be sure you have a field called "OBJECTID." If your ID field is called something else, use the "SQL" button on the right to create an expression using the "IS NOT NULL" syntax.*
4. If you have access sites in multiple states, use the "SQL" button on the right and create an expression that indicates which access sites are in the first state, then choose that state from the dropdown list. Repeat this with up to 4 states. Note that the State 2 through 4 drop down lists are grayed out until you enter an SQL expression pointing toward that state in your data.

Access Site - 2. Parking tool

Use this tool to designate access site parking as "day-use," "overnight," or "no parking."

1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 2. Parking." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "AccessSite_1_." If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.

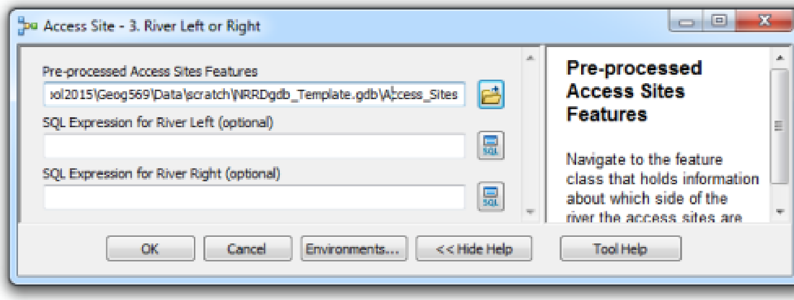


3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the parking type indicated. If you have no sites with a given parking type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Access Site - 3. River Left or Right tool

This tool identifies whether a site is on river left or river right.

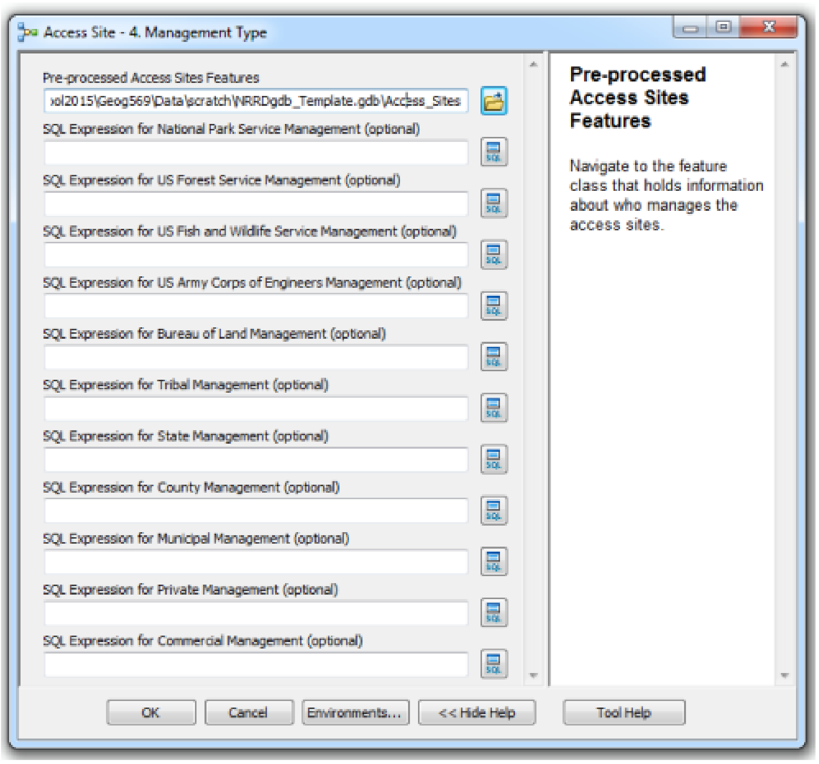
1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 3. River Left or Right." Wait for the dialog box to open.
2. n the "Pre-processed Access Site Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "AccessSite_2_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.



3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the river bank indicated.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Access Site - 4. Management Type tool

This tool sorts access site management into one of eleven pre-set management types: National Park Service, US Forest Service, US Fish and Wildlife Service, US Army Corps of Engineers, Bureau of Land Management, Tribal, State, County, Municipal, Private, or Commercial. Note that there is a text field for entering specific management entities in the "Text Fields" tool (No. 7). So, for example, enter a field that has information such as "Minnesota Department of Natural Resources" or "Bridger-Teton National Forest" in the text fields tool, but "State" and "US Forest Service" here.

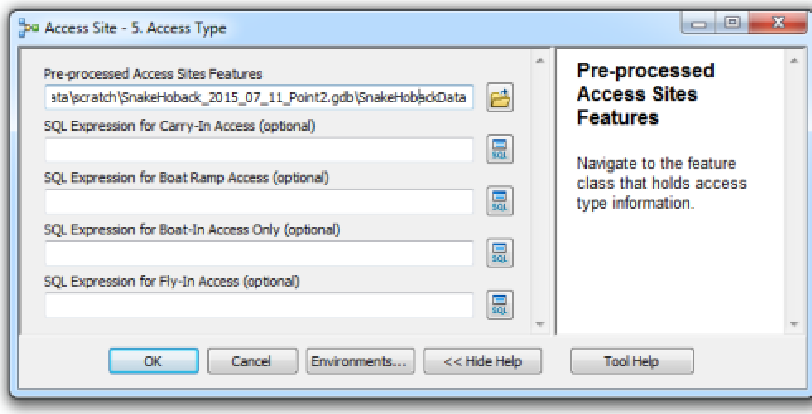


1. Open the “AccessSites” toolbox. Double click on the tool labelled “Access Site - 4. Management Type.” Wait for the dialog box to open.
2. In the “Pre-processed Access Site Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “AccessSite_3_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your pre-processed feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that identifies the management type indicated. If there are no access sites under a given management type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Access Site - 5. Access Type tool

Use this tool to designate access site type as “Carry-In,” “Boat Ramp,” “Boat-In,” and/or “Fly-In.”

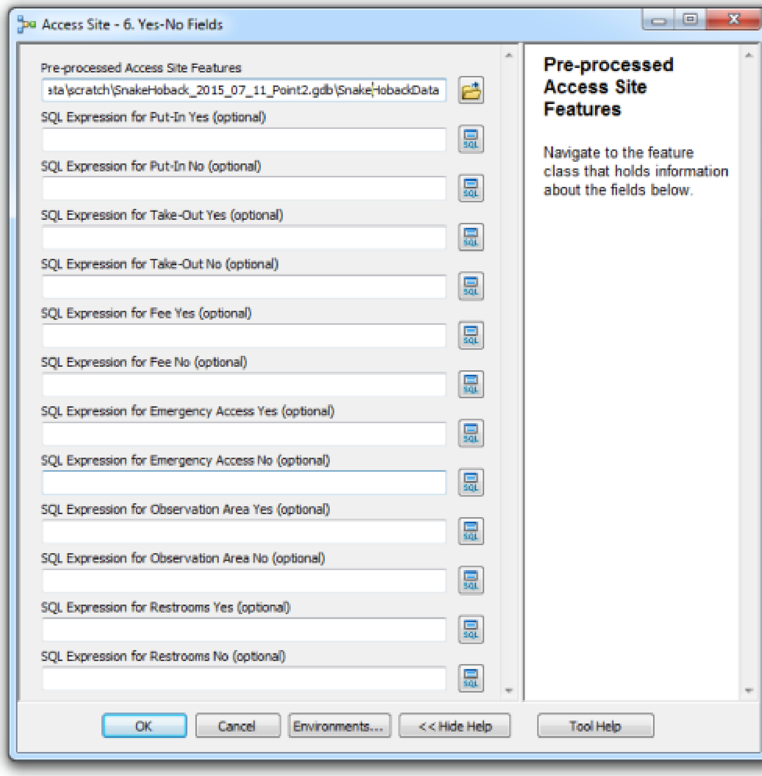
1. Open the “AccessSites” toolbox. Double click on the tool labelled “Access Site - 5. Access Type.” Wait for the dialog box to open.
2. In the “Pre-processed Access Site Features” box, if you have run previous tools, choose the feature class from the drop-down list that begins “AccessSite_4_” (or the number of the most recent tool you ran). If this is the first tool you’re running, choose your pre-processed feature class from the drop down menu, or navigate to its location.



3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the access type indicated. If you have no sites with a given access type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Access Site - 6. Yes-No Fields tool

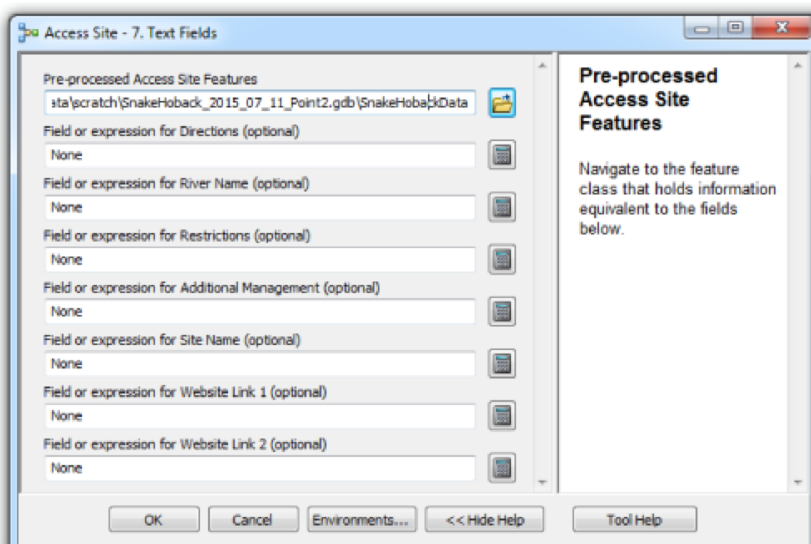
Use this tool to populate yes/no fields for whether your site can be used as a put-in or take-out point, whether or not it requires a fee, whether or not it is an emergency access point, and whether or not it has a designated observation area and/or restrooms.



1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 6. Yes-No Fields." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "AccessSite_5_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the value indicated. If you have no sites with a value, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Access Site - 7. Text Fields tool

This tool populates several text fields in the NRRD Access_Sites feature class with comparable data. Text fields include directions, river name, restrictions (this is a catch-all for other pertinent information), additional management (more detailed version of what was entered in the "Management Type" tool), site name, and two fields for site specific web links. For descriptions for each of the NRRD fields indicated, refer to table 2 on pages 29-30.



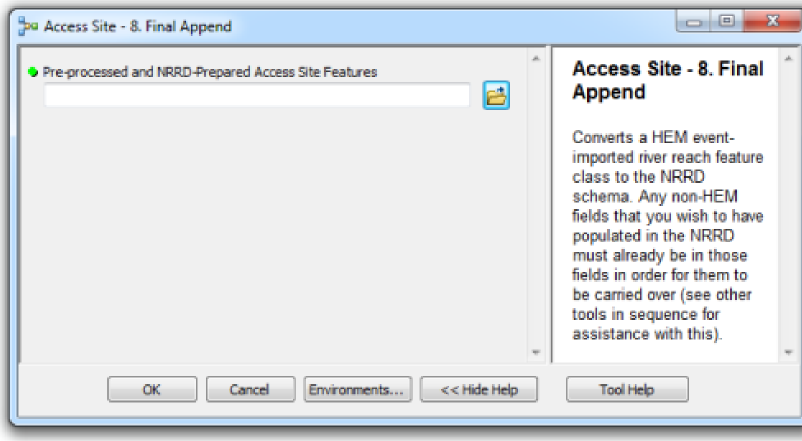
1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 7. Text Fields." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, if you have run previous tools,

choose the feature class from the drop-down list that begins "AccessSite_6_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your feature class from the drop down menu, or navigate to its location.

3. For each of the next boxes, click the field calculator icon to the right of the box and choose the field in your data that corresponds to the text data indicated.
4. Note that it takes a few seconds after each entry for the dialog box to register your choice.

Access Site - 8. Final Append **REQUIRED**

Once you have completed transforming your data with those tools that are applicable, run the "Final Append" tool to write your data to a new feature class that is ready to submit to the NRRD.



1. Open the "AccessSites" toolbox. Double click on the tool labelled "Access Site - 8. Final Append." Wait for the dialog box to open.
4. In the "Pre-processed River Reach Features" box, choose the feature class from the drop-down list that begins "AccessSite_7_" (or the number of the most recent tool you ran).
5. Note that this tool will discard any fields that are not part of the NHD (HEM) or NRRD schema. If you have data you wish to include, it must be already translated into one of those fields using the tools in this toolkit.

AT THIS POINT Y OU HAVE FINISHED PREPARING YOUR ACCESS SITE POINT DATA FOR SUBMISSION TO THE NRRD. NEXT YOU WILL HAVE THE OPPORTUNITY TO TRANSFER INFORMATION ABOUT CAMPGROUNDS.

Campgrounds Toolbox

The following tools populate fields in the NRRD feature class Campgrounds with comparable data from your pre-processed dataset. If you have no data about campgrounds, you can skip this entire section. If you do not have comparable data for a given field, you can skip that tool. If you use any of the tools in this section, you MUST run the tool "Campgrounds 5. – Final Append." Below is a list of the tools available. Each corresponds to a field or series of similar fields in the NRRD feature class. Table X describes these fields.

Tools in the Campgrounds toolbox

Campgrounds - 1. Fees tool
Campgrounds - 2. Text Fields tool
Campgrounds - 3. Management Type tool
Campgrounds - 4. Campground Type tool
*Campgrounds - 5. Final Append **REQUIRED***

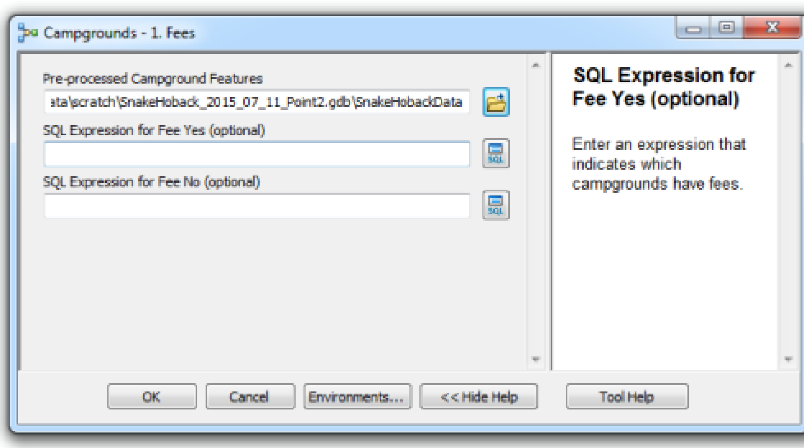
Table 8: Fields in the NRRD Campgrounds feature class

FieldName	Type	Length	Description
Name	String	75	Name of campground
Managing_Entity	String	50	Managing Entity
Website	String	250	Website for campground
Fee	SmallInteger	2	Yes/No
Type	String	100	Developed or Dispersed
Notes	String	300	Notes

Campgrounds - 1. Fees tool

Use this tool to indicate whether or not a campground requires a fee.

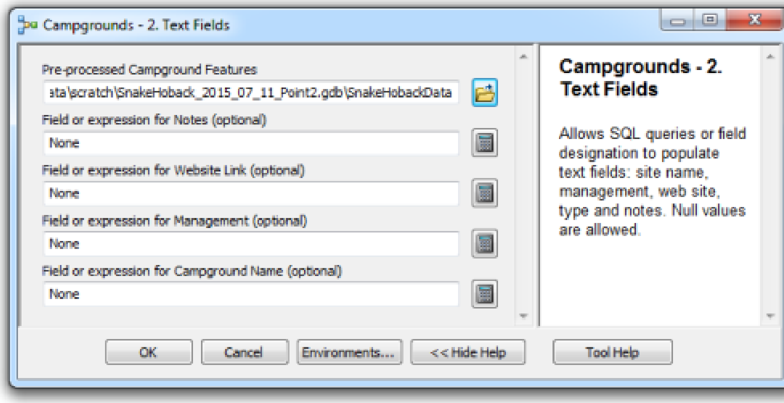
1. Open the "Campgrounds" toolbox. Double click on the tool labelled "Campgrounds - 1. Fees." Wait for the dialog box to open.
2. In the "Pre-processed Campground Features" box, choose your pre-processed feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to where fees are and/or are not required. If you have no sites with a given fee requirement, leave that box blank.



4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Campgrounds - 2. Text Fields tool

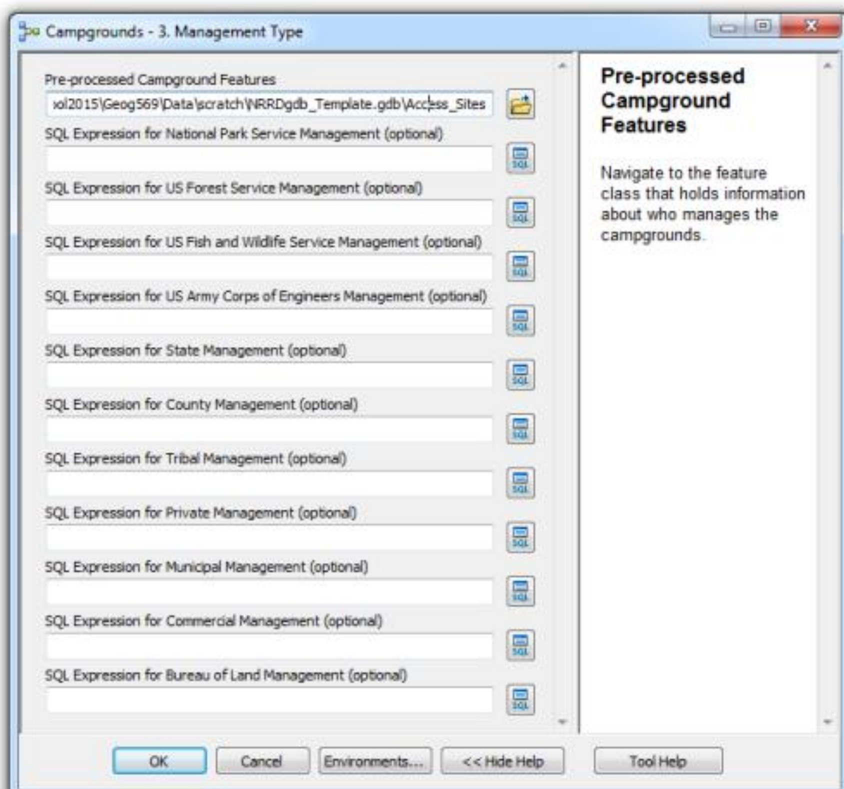
1. This tool populates several text fields in the NRRD Campgrounds feature class with comparable data. Text fields include notes, website link, campground management, and campground name. For descriptions for each of the NRRD fields indicated, refer to table 3 on page 36.
2. Open the "Campgrounds" toolbox. Double click on the tool labelled "Campgrounds - 2. Text Fields." Wait for the dialog box to open.



3. In the "Pre-processed Campground Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "AccessSite_1_." If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.
4. For each of the next boxes, click the field calculator icon to the right of the box and choose the field in your data that corresponds to the text data indicated.
5. Note that it takes a few seconds after each entry for the dialog box to register your choice.

Campgrounds - 3. Management Type tool

This tool sorts access site management into one of eleven pre-set management types: National Park Service, US Forest Service, US Fish and Wildlife Service, US Army Corps of Engineers, Bureau of Land Management, Tribal, State, County, Municipal, Private, or Commercial. Note that there is a text field for entering specific management entities in the "Text Fields" tool (No. 2). So, for example, enter a field that has information such as "Minnesota Department of Natural Resources" or "Bridger-Teton National Forest" in the text fields tool, but "State" and "US Forest Service" here.



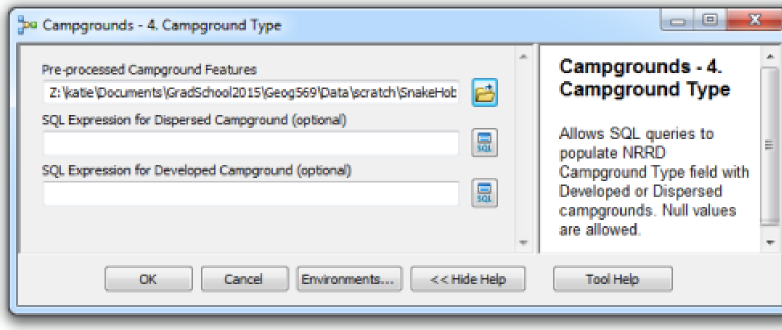
1. Open the "AccessSites" toolbox. Double click on the tool labelled "Campgrounds - 3. Management Type." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features"

box, if you have run previous tools, choose the feature class from the drop-down list that begins "Campgrounds_2_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.

3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that identifies the management type indicated. If there are no access sites under a given management type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Campgrounds - 4. Campground Type tool

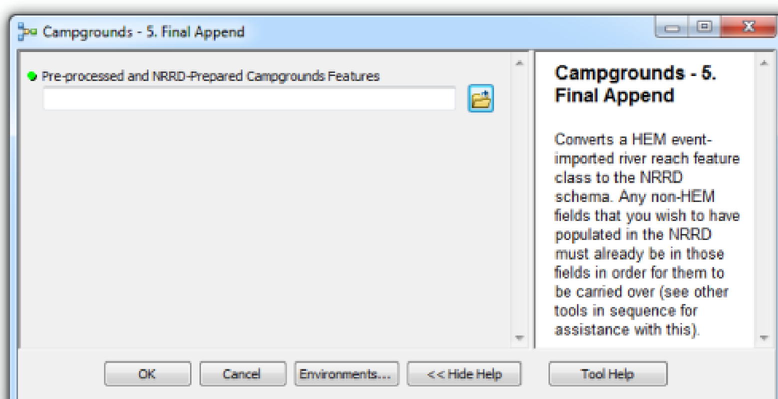
Use this tool to designate access site type as "Developed" or "Dispersed."



1. Open the "Campgrounds" toolbox. Double click on the tool labelled "Campgrounds - 4. Campground Type." Wait for the dialog box to open.
2. In the "Pre-processed Access Site Features" box, if you have run previous tools, choose the feature class from the drop-down list that begins "Campgrounds_3_" (or the number of the most recent tool you ran). If this is the first tool you're running, choose your pre-processed feature class from the drop down menu, or navigate to its location.
3. For each of the next boxes, click the SQL icon to the right of the box and enter an expression that points in your data to the campground type indicated. If you have no sites with a given campground type, leave that box blank.
4. Note that it takes a few seconds after each entry for it to appear in the dialog box.

Campgrounds - 5. Final Append **REQUIRED**

Once you have completed transforming your data with those tools that are applicable, run the "Final Append" tool to write your data to a new feature class that is ready to submit to the NRRD.



2. Open the "Campgrounds" toolbox. Double click on the tool labelled "Campgrounds - 5. Final Append." Wait for the dialog box to open.
6. In the "Pre-processed River Reach Features" box,

choose the feature class from the drop-down list that begins “Campgrounds_4_” (or the number of the most recent tool you ran).

7. Note that this tool will discard any fields that are not part of the NHD (HEM) or NRRD schema. If you have data you wish to include, it must be already translated into one of those fields using the tools in this toolkit.

AT THIS POINT YOU HAVE FINISHED PREPARING YOUR CAMPGROUNDS POINT DATA FOR SUBMISSION TO THE NRRD.