Program TagPro

Acoustic-Tag Data Translation Utility

Version 1.0

Developed by:
Peter Westhagen, John R. Skalski, and James Lady
Columbia Basin Research
School of Aquatic & Fishery Sciences
University of Washington
1325 Fourth Avenue, Suite 1515
Seattle, Washington 98101-2540

Prepared for:
U.S. Army Corps of Engineers, Portland District
Under a Government Order with the U.S. Department of Energy
Contract DE-AC05-76RL01830

Novermber 2021

Acknowledgments

This project was originally funded by the Pacific Northwest National Laboratory for the U.S. Army Corps of Engineers, Portland District under Contract DE-AC05-76RL01830.

Support and upkeep of website and data are provided by Bonneville Power Administration under Contract 76910.

Table of Contents

Chapter 1: Overview

Program TagPro (Acoustic-Tag Data Translation Utility)

Program TagPro is a desktop application that takes valid acoustic-tag events produced by Program FAST or data that has been formatted similarly and creates an output file of capture/detection histories used for survival analysis by Program ATLAS or other third-party software.

Figure 1 shows Program TagPro at startup.

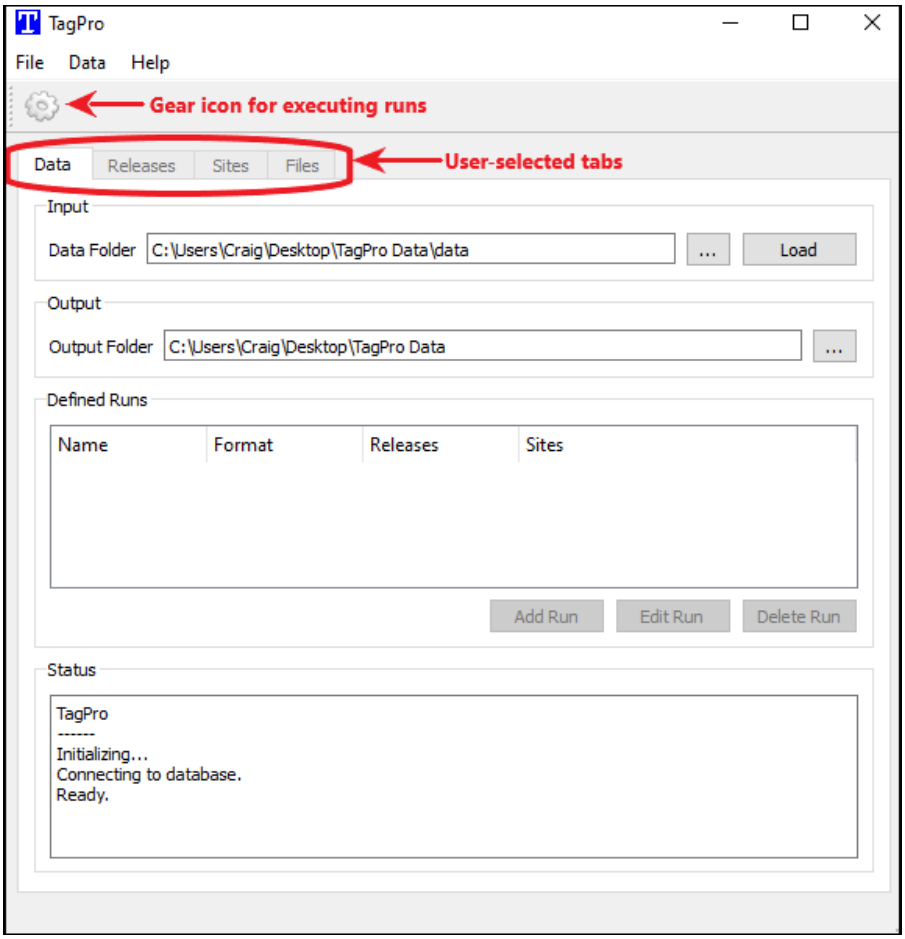


Figure 1. Program TagPro at startup

There are four main tabs—"Data," "Release," "Sites," and "Files" —where the primary work is done. The user will press the gear icon in the upper left to execute the user-defined runs. At startup before any data are loaded, only the "Data" tab is available.

There are also menus across the top. The only functionality in the menus not available elsewhere is the "Clear All" command under the "Data" menu. This will clear all input data, and the TagPro application will revert to its initial startup state. Since this is a destructive command, a warning will be issued first that allows the user to cancel the command.

The steps in Program TagPro to create an output file are as follows:

1. Load the input data
2. Define the release groups
3. Define the sites to be used
4. Define the runs to be executed
5. Execute the runs

Chapter 2: Loading the Input Data

Program TagPro requires three input files:

1. "tags.csv" contains the tagging data
2. "nodes.csv" contains the node definitions
3. "events.csv" contains the valid detection events

Optionally, the input files may include the file "removals.csv," which contains information about censoring events, and the files "tag_attributes.csv" and "tag_life_data.csv. These data are obtained through other sources such as PIT-tag data. The optional files can be uploaded at study creation or any time prior to study verification.

2.1 File Format

Files may contain one or more optional columns that do not need to be present in the uploaded files. They will, however, be included in exported files though their contents will be set to null if not represented in the original imported files. Input and output formats are identical for each of these six types and are given below. The proper column titles for each file type are listed below. The Appendix to this manual further details the types of values that should exist within each column.

Tags File

Unique key: fish_tag_year, tag_code (optional columns highlighted in **blue**)

1. tag_code
2. **project_code**
3. **tag_group**
4. **tagger_name_xlat**
5. **bucket**
6. lot
7. species_code_xlat
8. **length**

9. **weight**
10. pit_code
11. fish_tag_date: format yyyy-mm-dd
12. tag_activate_date: format yyyy-mm-dd hh:mm:ss
13. tag_release_date: format yyyy-mm-dd hh:mm:ss
14. release_location: the main release location (e.g. Roosevelt)
15. release_location_xlat: the sub-release location (e.g. Roosevelt_01)
16. release_location_river_kilometer
17. release_location_latitude
18. release_location_longitude
19. **mort_xlat: 0 if not a mortality**
    - 0: Released Alive
    - 1: 24 hour holding morality release or surgical mortality release .
    - 2: Intentional sacrifice and release.
    - 3:Transport mortality release
20. pri (e.g., *pulse rate interval*)

Nodes File

Unique key: node_code, deploy_year

1. node_code
2. latitude
3. longitude
4. river_kilometer
5. location: this is the array (e.g. CR349.0)
6. location_xlat: this is has an indication of the node (e.g. CR349.0_01)
7. deploy_date: format yyyy-mm-dd hh:mm:ss
8. recovery_date: format yyyy-mm-dd hh:mm:ss
9. elev_or_depth

Events File

1. node_code
2. tag_code
3. first_computed_datetime: format yyyy-mm-dd hh:mm:ss
4. last_computed_datetime: format yyyy-mm-dd hh:mm:ss
5. hits

Tag Life Data File

1. tag_code
2. lot
3. tag_life_days (precision to 100th of a day)

Tag Attributes File

1. tag_code
2. attribute
3. value

Removals File

1. tag_code
2. removal_date: format yyyy-mm-dd hh:mm:ss
3. riverkm
4. removal_type_id
    - 1: PIT- Censoring Indicated by PIT Detection
    - 2: UPRIVER- Upriver behavior
    - 3: BIRD- Bird Predation
    - 4: ATTRANS- Transportation indicated by AT Detections

It is important to note that date format across all input files must be yyyy-mm-dd hh:mm:ss or yyyy-mm-dd (as specified), as dates are used in ordering detections. **WARNING:** If you are using Excel to process or view your data, it will save dates in a different default format from this.

2.2 File Loading

The next step is to specify the input directory for the input files, as shown in Figure 2.

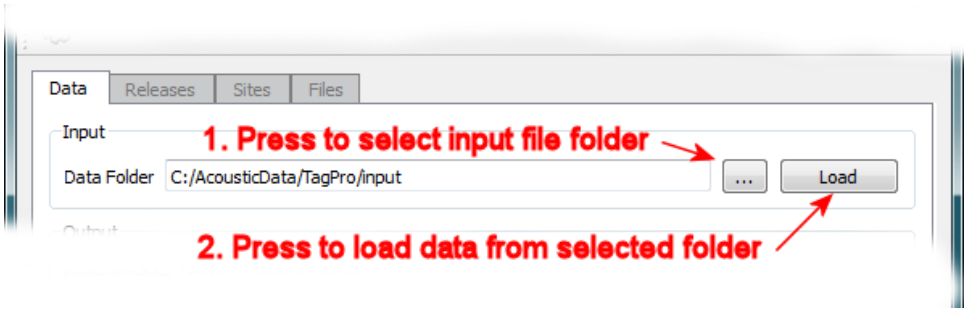


Figure 2. Loading input data in Program TagPro

All of the files mentioned above must be located in the specified input directory. The user then presses the "Load" button to load the data. Large files may take several minutes. When completed, the other tabs—"Release," "Sites," and "Files"—become available.

2.3 Files Tab

After loading the data, the user may select the "Files" tab to show the input files. It will show the required files listed above. It will also display any other files in the input folder that are not used by the TagPro application. This is because some files, such as a tag-life-data file, are important for the subsequent survival analysis, and it is important to keep their association with the input data. An example is shown in Figure 3.

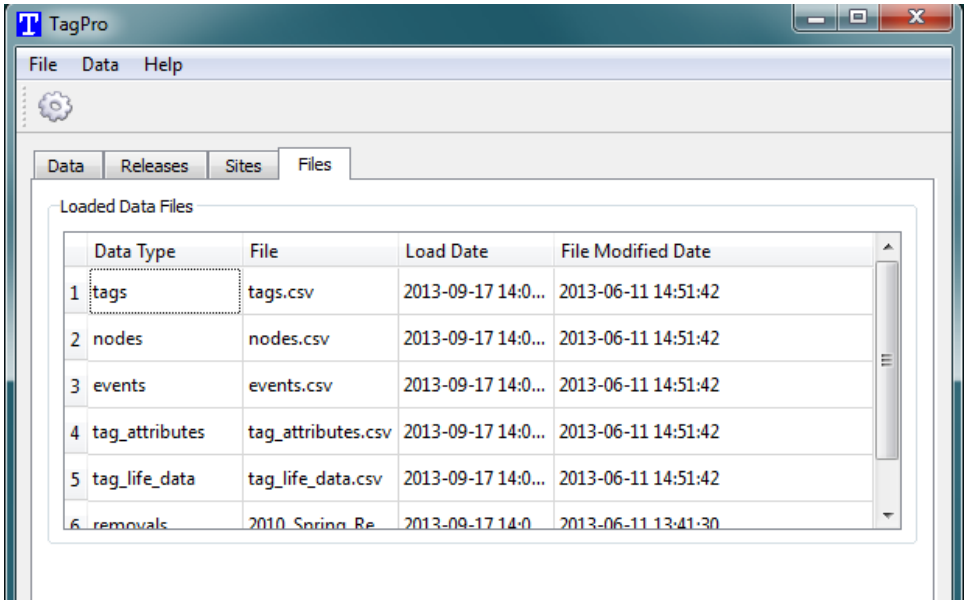


Figure 3. "Files" tab showing input files

Chapter 3: Defining the Release Groups

After loading the data, the user can click on the "Releases" tab to define the release groups. Figure 4 shows the "Releases" tab for a sample data set.

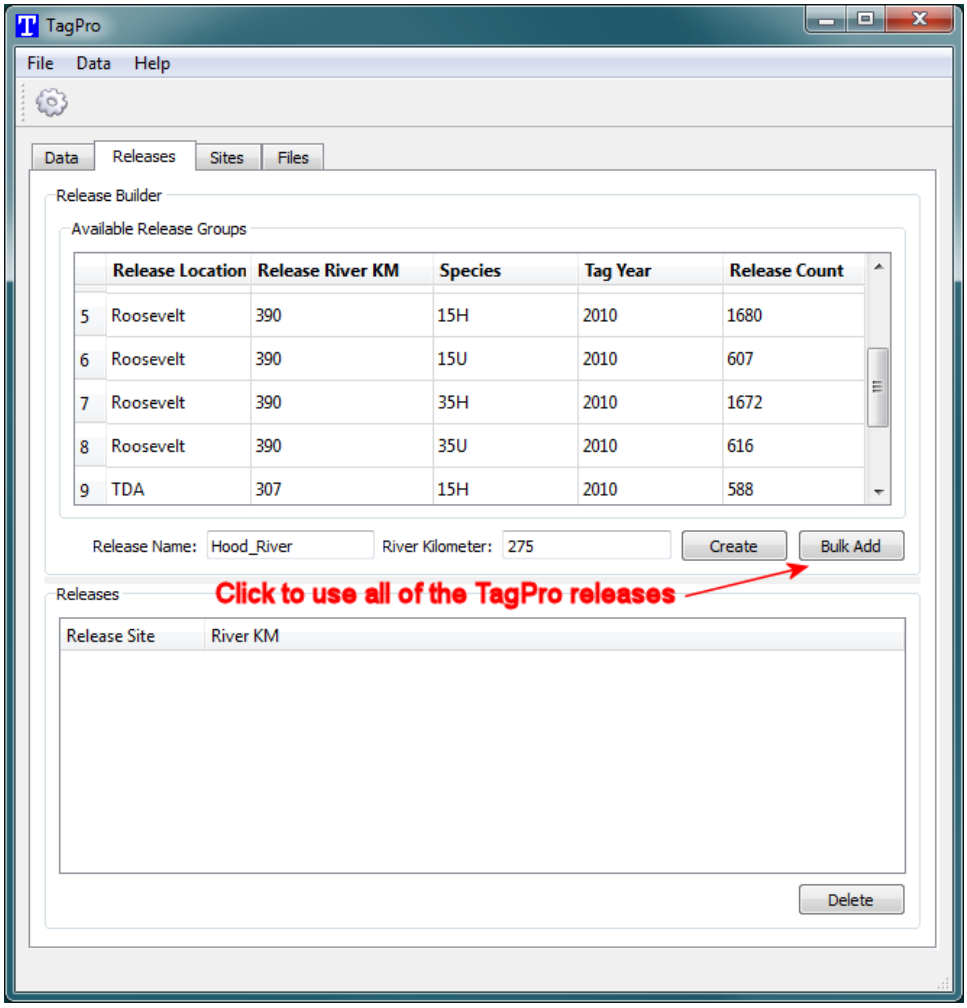


Figure 4. Defining the release groups

It is divided into two parts:

1. The top part is labeled “Available Release Groups.” When TagPro loads the data, it defines these release groups by unique release location and species codes.
2. The bottom part shows the releases that the user defines. It is blank at the start, and the user must create releases to be used for executing the runs.

As shown Figure 4, the user may simply use the *Ctrl* or *Shift* keys to select multiple release groups or a range of release groups and then click the “Bulk Add” button to add the selected groups. Alternatively, the user may click on a single specific release group, change the “Release Name” and “River Kilometer” for that group, and then click “Create” to add that single release group to their releases. The “Bulk Add” function does not allow for changing the release names and river kilometers.

If pooled release group is desired, (for example, individual replicate groups released at the same location which can be considered as a single release), highlight the desired releases, rename the “Release Name,” and use “Create” instead of “Bulk Add.”

In Figure 5, the user has selected the four "Hood_River" releases. The default release name is simply the one used in the tags input file, or in the case of multiple releases selected, the name of the first release followed by a hyphen and then the name of the last release.

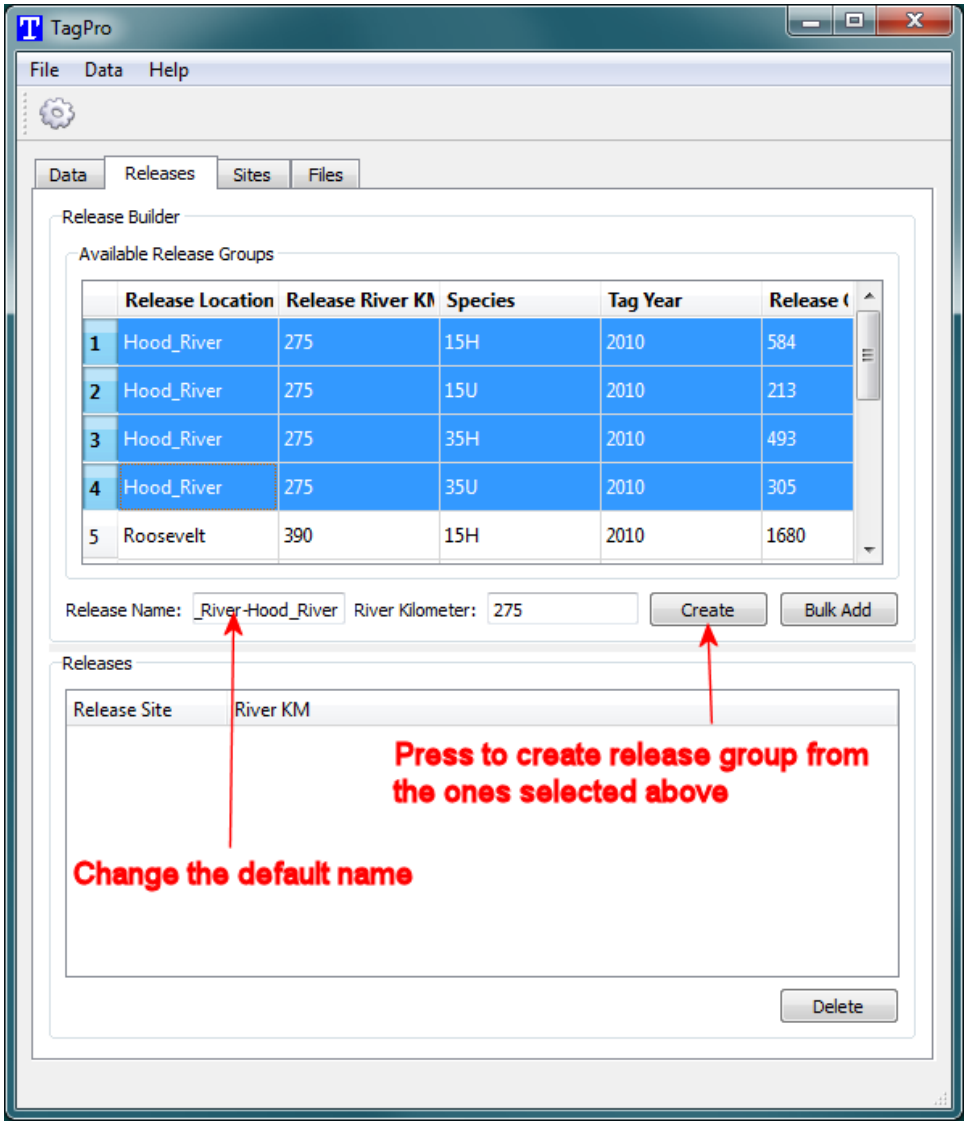


Figure 5. Defining a release group from four TagPro-defined release groups

The user can change the default name by editing the “Release Name” as shown in Figure 5. In this case, if the user changes the name to “All Hood_River,” and clicks “Create,” the result will be as shown in Figure 6. The user can continue to create more user-defined release groups in the same way.

To remove a release group that has previously been defined, select it under the “Releases” table and click the “Delete” button.

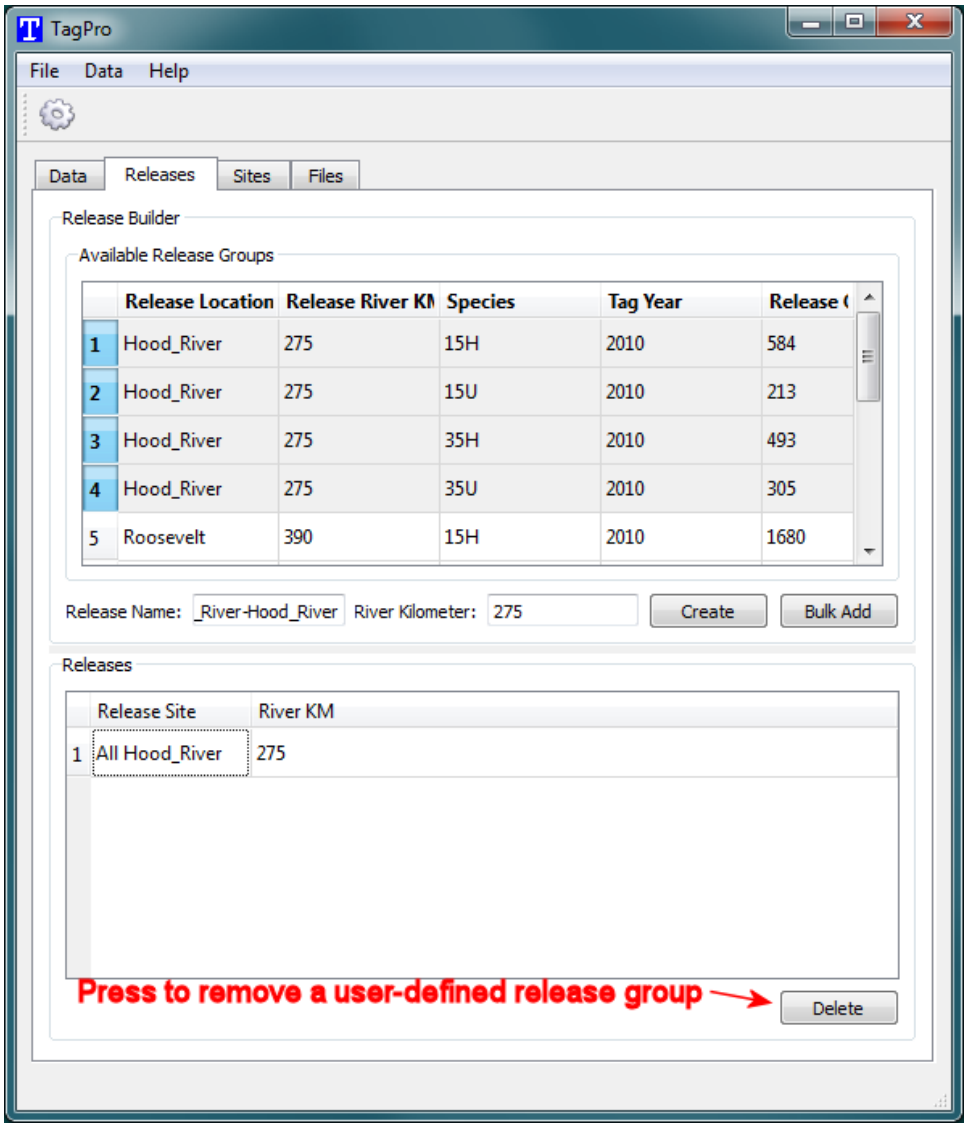


Figure 6. The “Releases” tab after defining a release group

In Figure 7, the user has created three release groups.

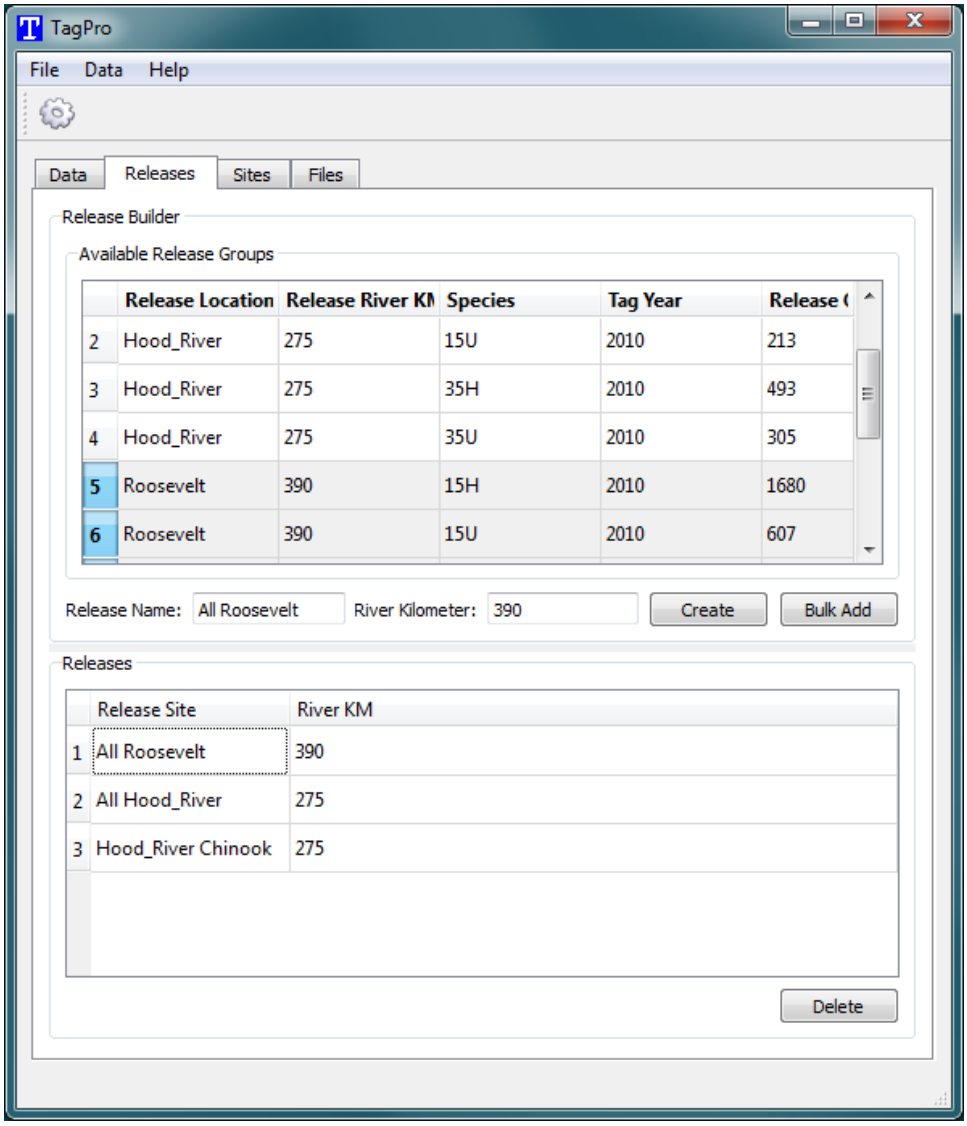


Figure 7. Multiple user-defined releases

Multiple releases in the "Available Releases Groups" are selected by holding the *Shift* key and left-clicking to select a range of releases, or by holding the *Ctrl* key and left-clicking to select the releases one at a time.

Chapter 4: Defining the Sites

The user presses the “Sites” tab to define the sites to be used, as shown in Figure 8.

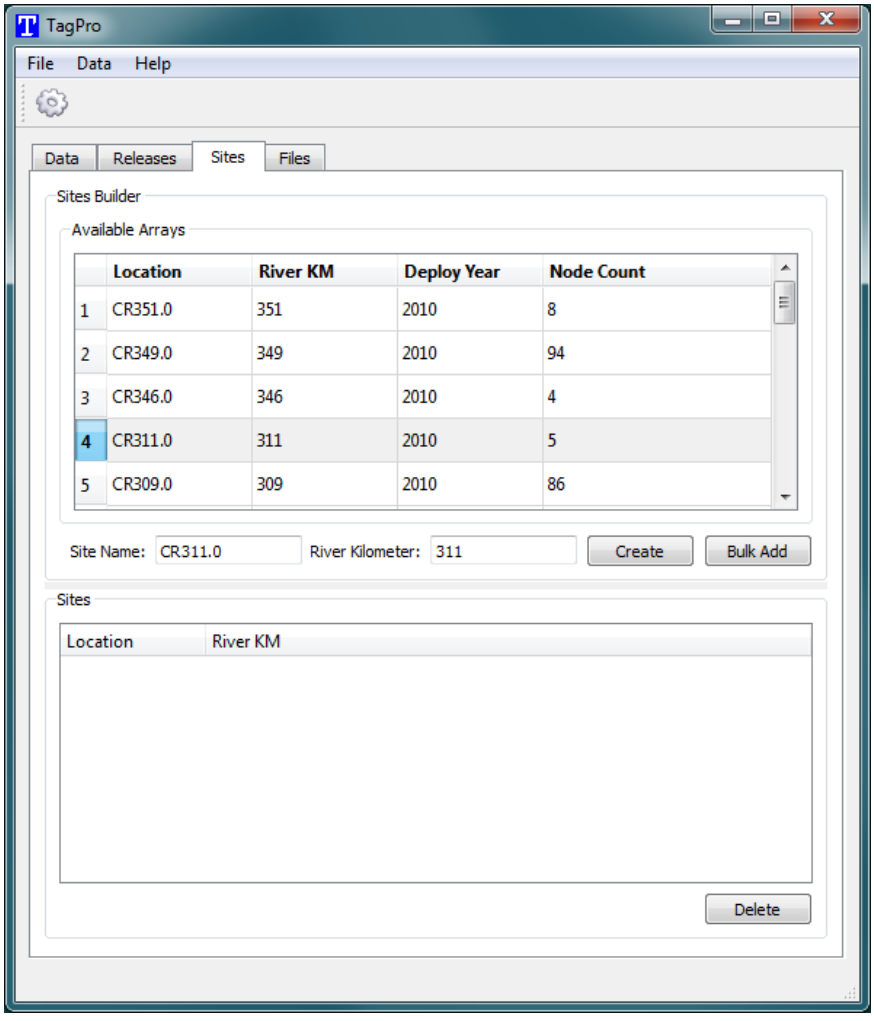


Figure 8. The “Sites” tab for defining sites in Program TagPro

This works in much the same way as defining release groups as described in the previous section. The user can press "Bulk Add" to include an array of selected sites, or select one or more sites, change the default Site Name if desired, and click "Create."

In a hypothetical study, for example, the user might want to look at survival through four upstream sites and may not be interested in survival in the downstream sites beyond that. Thus, all downstream sites could be grouped together into one user-defined site; all detections at any of the downstream sites would then be counted as detections at the user-defined site.

In Figure 9, the user includes the first upstream site, “CR351.0” by selecting it and clicking “Create.”

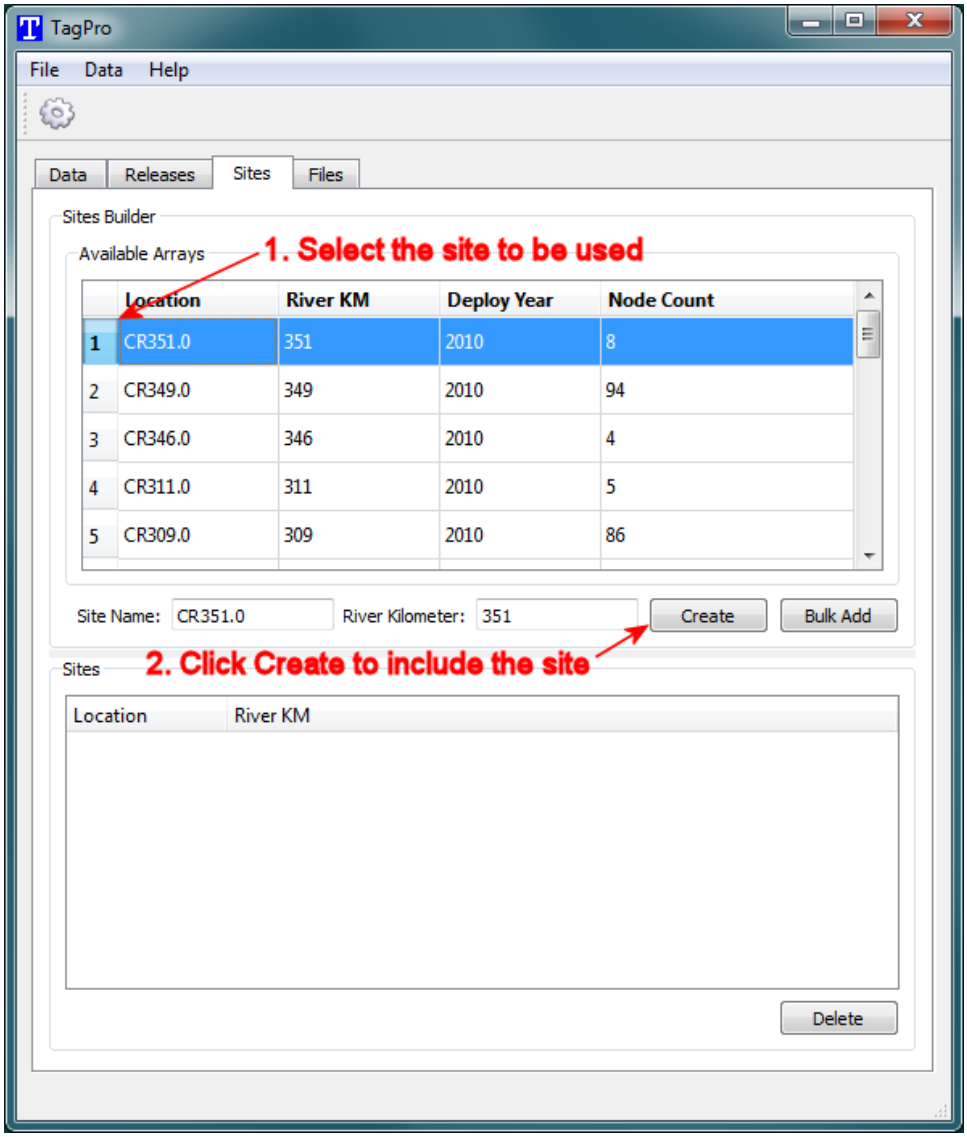


Figure 9. A user includes one available upstream site

The site CR351.0 will then appear in the bottom table labeled "Sites." This process is then repeated for the next three sites with the result shown in Figure 10.

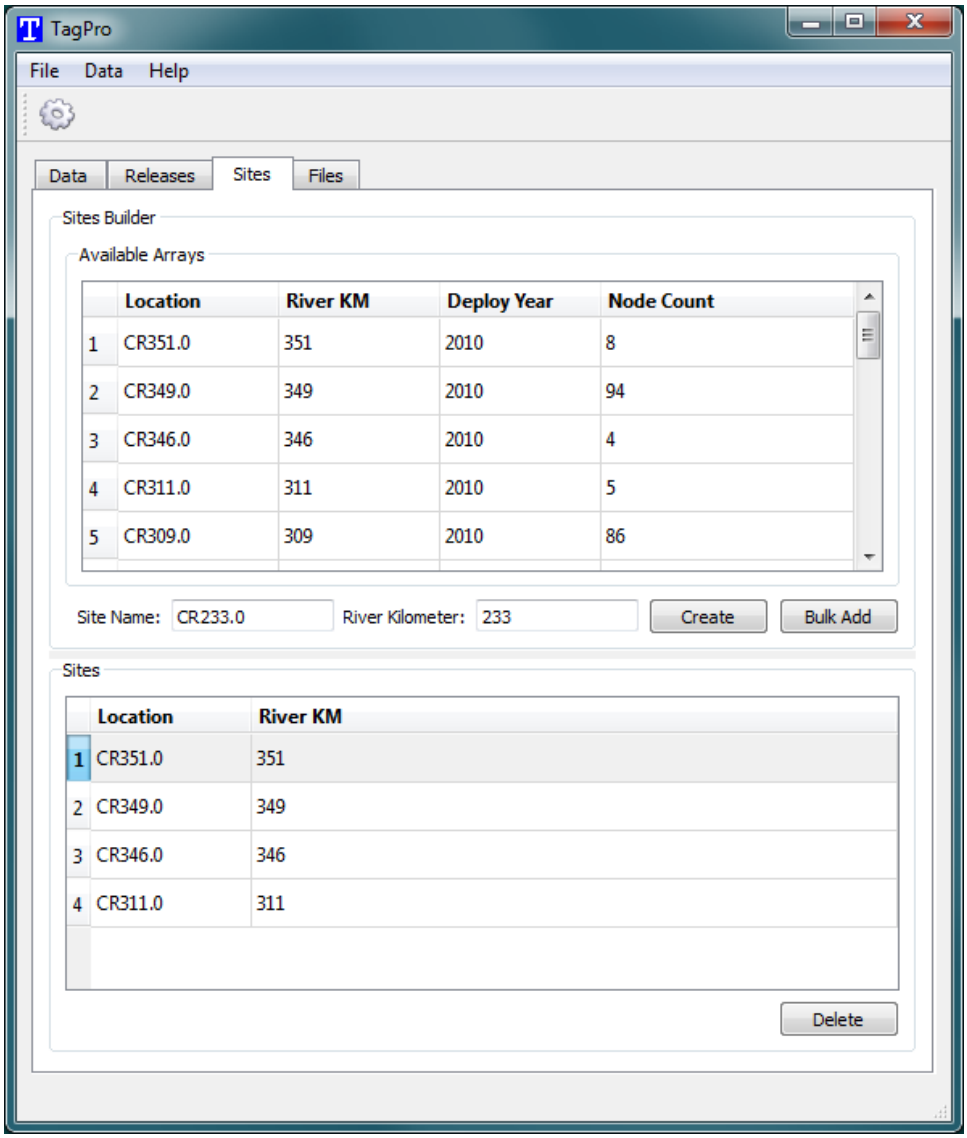


Figure 10. The "Sites" tab after defining four upstream sites

In order to include the remaining downstream sites, the user clicks on the first downstream site (CR309.0), scrolls down to the final site (CR002.8), holds *Shift* and left-clicks on it, and changes the site name to "Downstream, and clicks "Create" as shown in Figure 11. The resulting site called "Downstream" will then appear in the "Sites" table in the lower portion of the Sites tab.

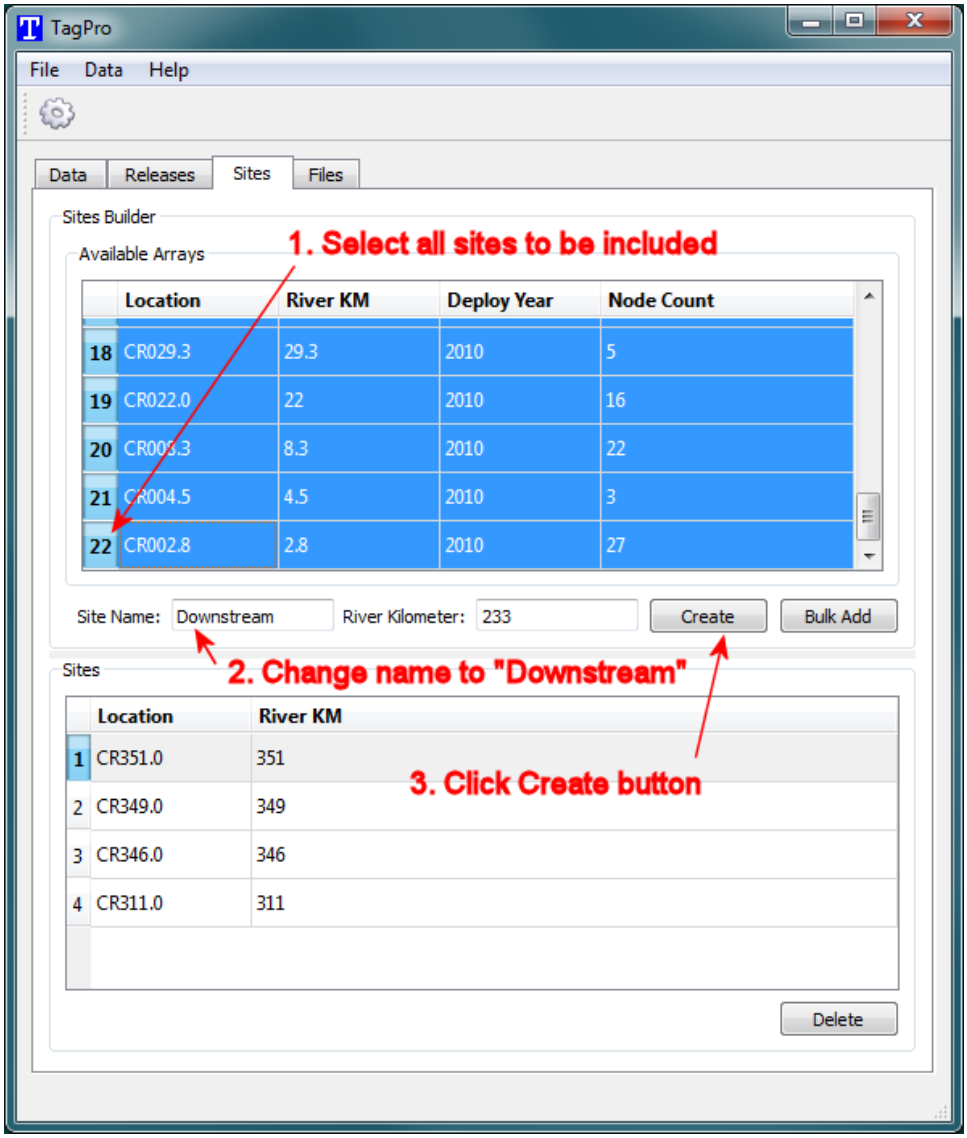


Figure 11. The "Sites" tab when creating one site that includes all downstream sites

Chapter 5: Defining the Runs

Once the release groups and the detection sites have been defined, the user can now set up "runs." A user selects one or more releases, and then selects two or more sites to define a run. When the run is executed, an output file of capture histories will be created that can be used to analyze survival of the chosen releases through the chosen sites.

In order to define a run, click on the "Data" tab and click the "Add Run" button. A dialog will appear as shown in Figure 12. The user must enter a name for the run and select an output file format. There are two formats available: ATLAS format and the Standard format, as explained below in Section 5.1. In the example shown in Figure 12, the user has named the run "Hood River Chinook" and is using the ATLAS format for the output file. The user then presses the "Next" button to continue.

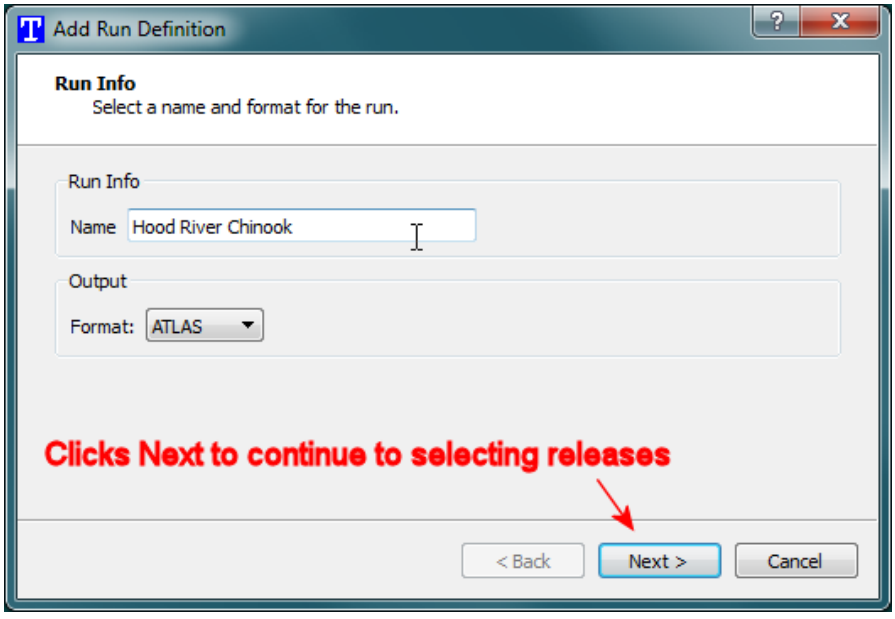


Figure 12. Dialog for creating a run

Figure 13 shows the dialog for selecting the releases for the run. In this example, three releases have been previously defined by the user, and one or more of them must be selected for defining the current run. In this case, the user selects just one release, “Hood_River Chinook” and presses “Next” to continue.

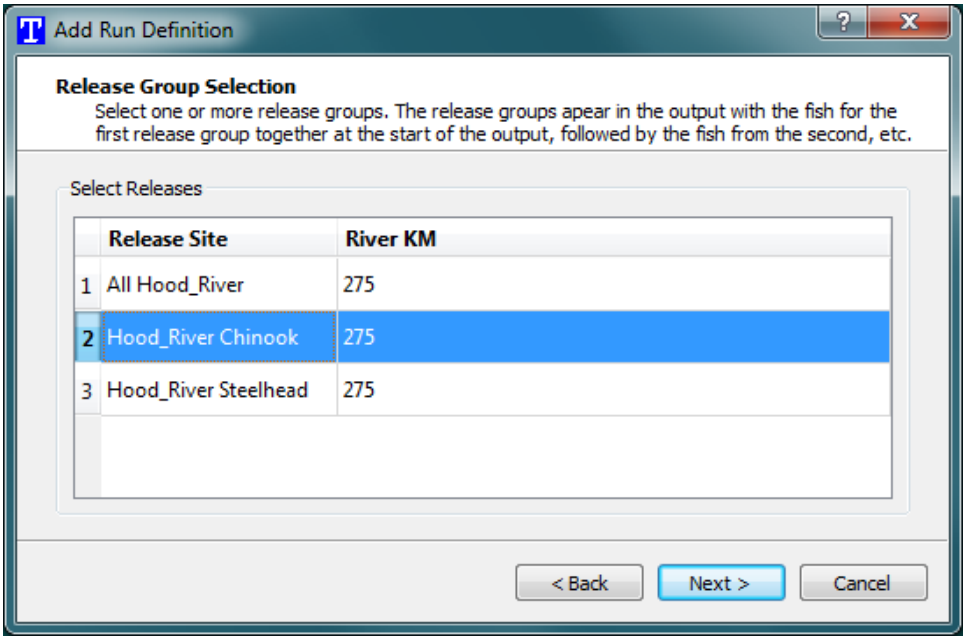


Figure 13. Selecting the releases for a run

Figure 14 shows the dialog for sites selection. The user must choose two or more previously defined sites. In this case, the user has selected all of the sites. As before, the user presses “Next” to continue.

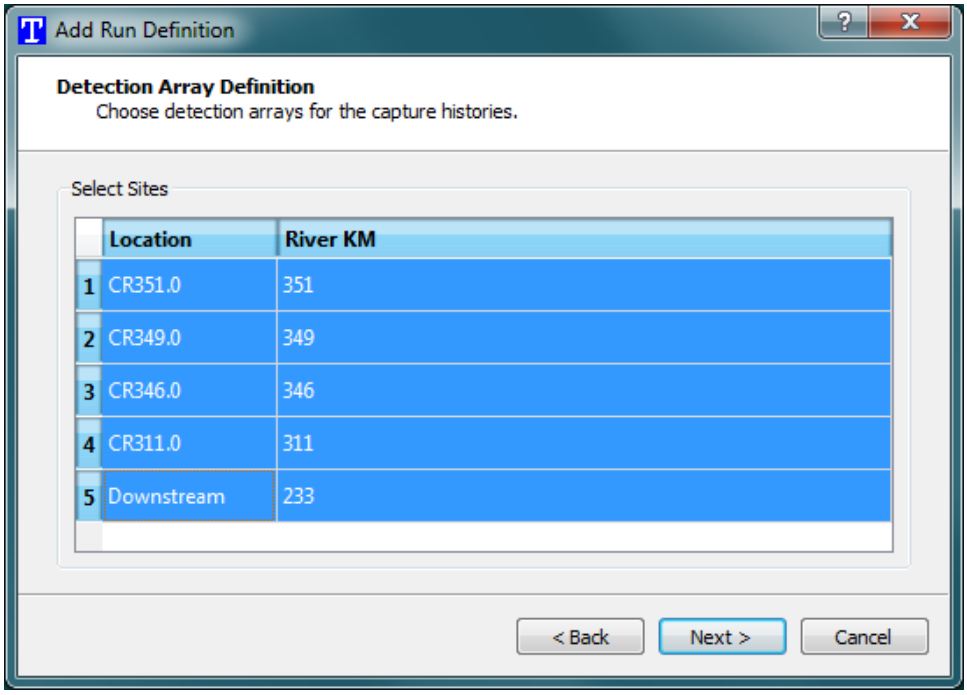


Figure 14. Selecting the sites for a run

The final step is to review the settings the user has selected, as shown in Figure 15. The summary shows the name of the run, the releases selected, and the sites selected. At any step while defining a run, the user can use the "Back" button to change a previous setting. Once the user is satisfied with the selections, the "Finish" button is clicked. The run then shows up on the "Data" tab of the TagPro dialog. The user can repeat the above process to define multiple runs.

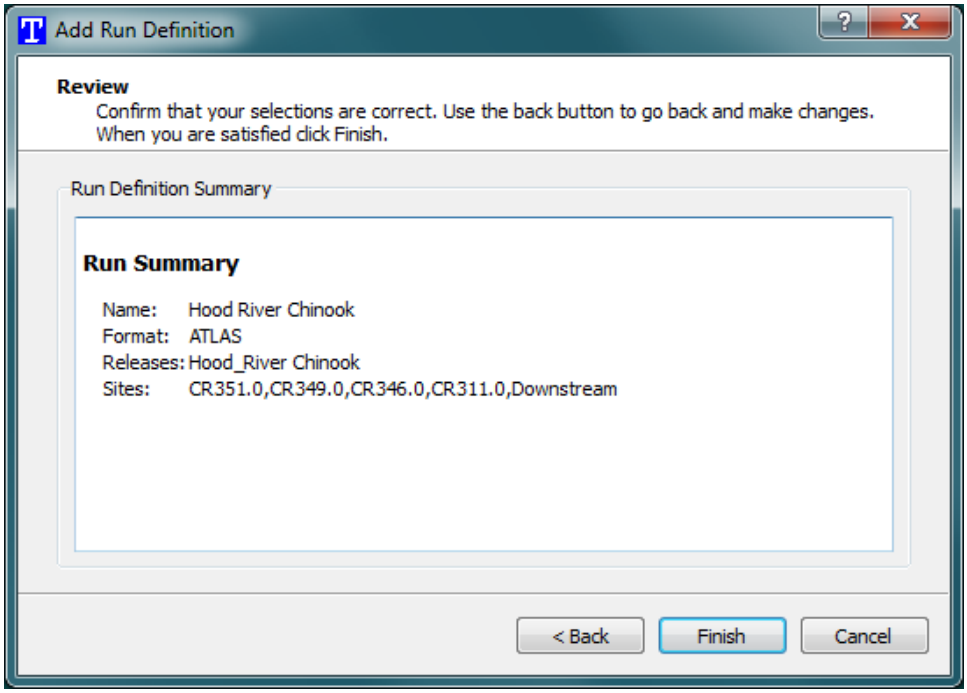


Figure 15. Summary of a run definition

In Figure 16, the user has defined an additional run—a second one using the Hood River Steelhead release instead of the Hood River Chinook.

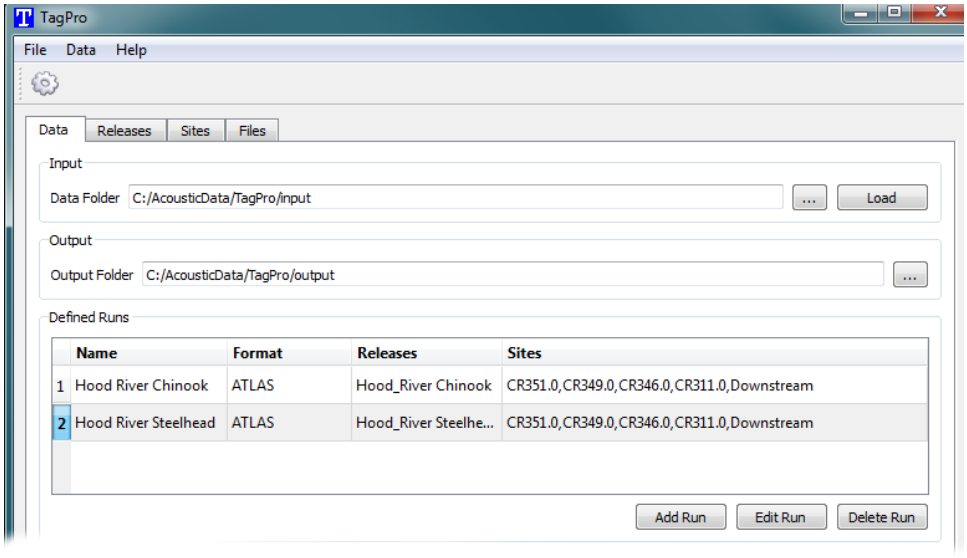


Figure 16. Portion of Main Dialog with two runs defined

5.1 Output Formats

There are two output formats available in the TagPro application: (1) ATLAS format and (2) Standard format.

The ATLAS format is the preferred format for survival analysis in Program ATLAS. The output file consists of one line per detection event. The Standard format can also be read by the Program ATLAS but contains additional information available in the input files for additional analysis such as looking at tagger effect, length or weight effect, or analyzing travel times.

Chapter 6: Executing the Runs

Prior to executing the runs, the user must select the folder where the output files are to be placed. This is done on the “Data” tab as shown in Figure 17.

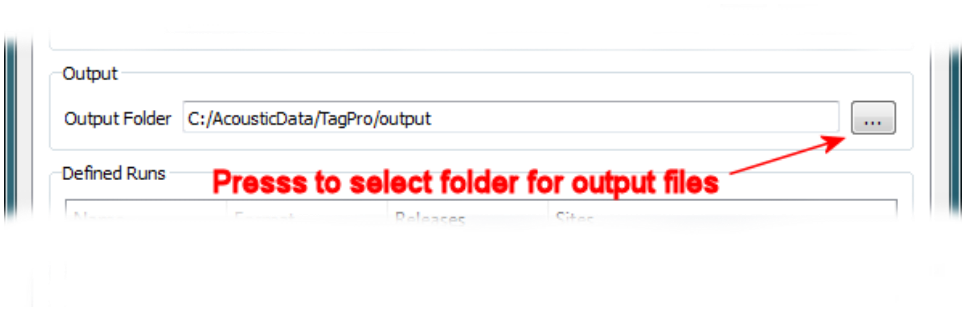


Figure 17. Selecting the folder for the output files

Once the runs have been defined and the output directory has been selected, the user may execute one or more runs by selecting them in the “Defined Runs” table, and pressing the gear icon in the upper left as shown in Figure 18.

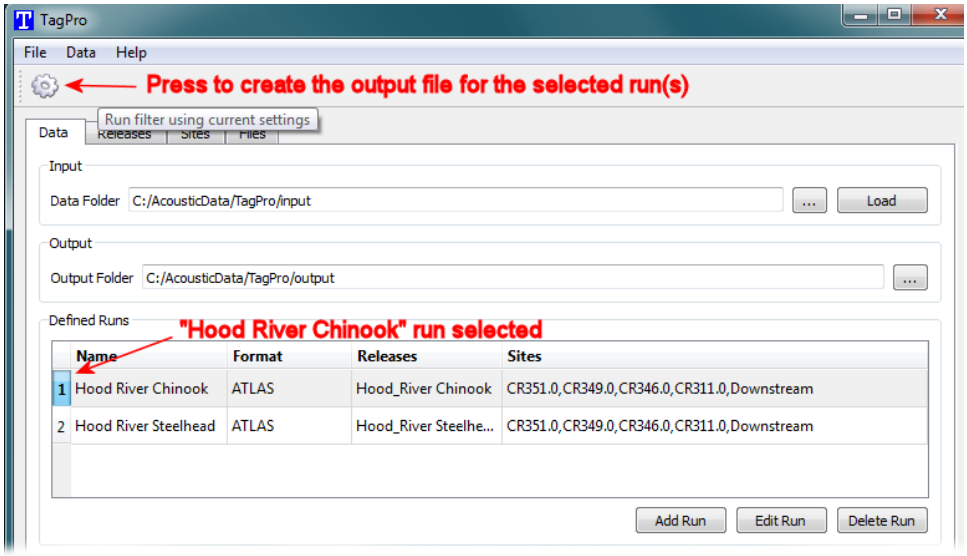
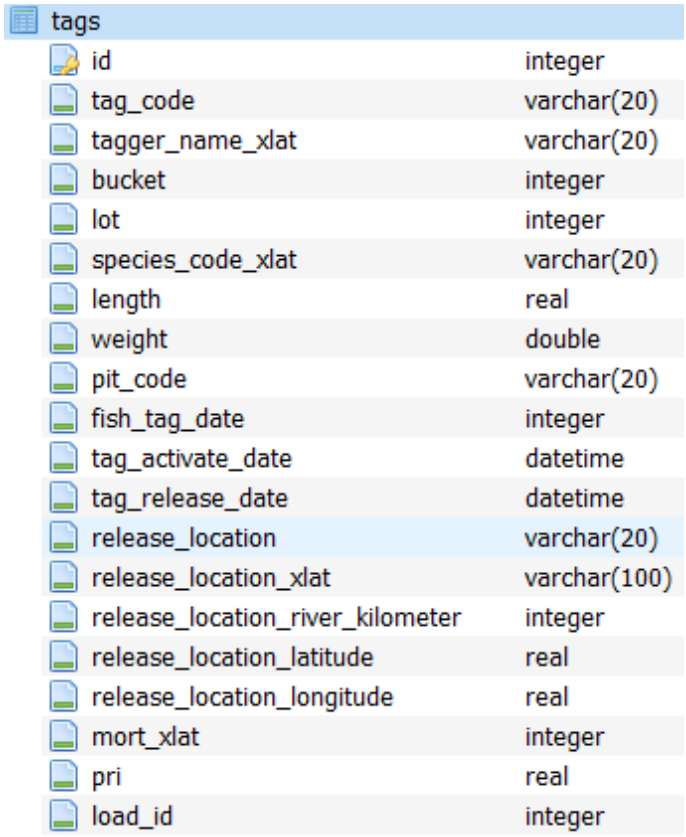


Figure 18. Executing the selected runs

In this example, the user has selected the "Hood River Chinook" run; pressing the gear icon will cause the selected run(s) to be executed and the output file(s) created. It may take several minutes for the output file to be created. The message "Run complete" will appear in the Status messages window when complete.

An output file is created for each run in the specified output directory. The name of the output file is the run name, followed by an underscore, followed by the format of the output ("ATLAS" or "Standard"). In the example in Figure 18, the output file will have the name "Hood River_ATLAS.csv."

Appendix: Input File Formats



tags	
id	integer
tag_code	varchar(20)
tagger_name_xlat	varchar(20)
bucket	integer
lot	integer
species_code_xlat	varchar(20)
length	real
weight	double
pit_code	varchar(20)
fish_tag_date	integer
tag_activate_date	datetime
tag_release_date	datetime
release_location	varchar(20)
release_location_xlat	varchar(100)
release_location_river_kilometer	integer
release_location_latitude	real
release_location_longitude	real
mort_xlat	integer
pri	real
load_id	integer

Figure A1. Proper column names and required value types for the *Tags* input file

nodes	
id	integer
node_code	varchar(20)
latitude	varchar(20)
longitude	varchar(20)
river_kilometer	integer
location	varchar(20)
location_xlat	varchar(20)
deploy_date	varchar(20)
recovery_date	varchar(20)
elev_or_depth	
load_id	integer

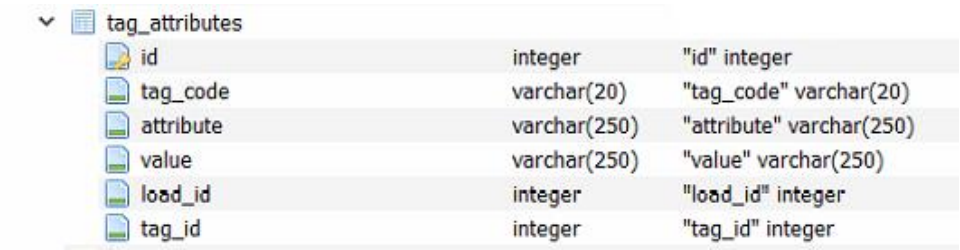
Figure A2. Proper column names and required value types for the *Nodes* input file

events	
id	integer
node_code	varchar(20)
tag_code	varchar(20)
first_computed_datetime	datetime
last_computed_datetime	datetime
hits	integer
load_id	integer
tag_id	integer
node_id	integer

Figure A3. Proper column names and required value types for the *Events* input file

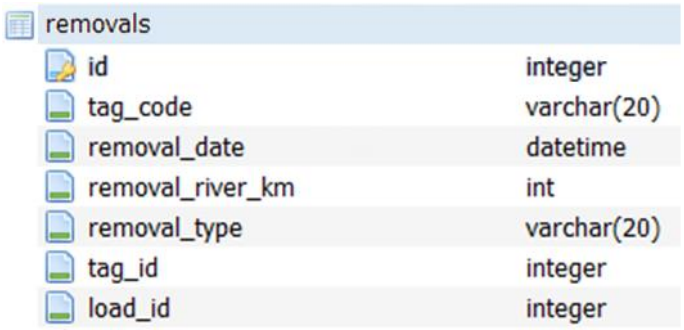
tag_life_data	
id	integer
tag_code	varchar(20)
lot	integer
tag_life_days	real
load_id	integer

Figure A4. Proper column names and required value types for the *Tag Life Data* input file



tag_attributes		
id	integer	"id" integer
tag_code	varchar(20)	"tag_code" varchar(20)
attribute	varchar(250)	"attribute" varchar(250)
value	varchar(250)	"value" varchar(250)
load_id	integer	"load_id" integer
tag_id	integer	"tag_id" integer

Figure A5. Proper column names and required value types for the *Tag Attributes* input file



removals	
id	integer
tag_code	varchar(20)
removal_date	datetime
removal_river_km	int
removal_type	varchar(20)
tag_id	integer
load_id	integer

Figure A6. Proper column names and required value types for the *Removals* input file