Filtering Acoustic Signal Transmissions (FAST) Program

Prepared by:
Peter Westhagen
Columbia Basin Research
School of Aquatic and Fishery Sciences
University of Washington

Version: 0.9 (beta)

Note: This current release of the program has been thoroughly tested on both autonomous node and cabled node data. The current release has been validated and is currently in beta testing for useability.

The Filtering Acoustic Signal Transmissions (FAST) program is designed to process JSATS acoustic-tag data from juvenile salmonid tagging studies in order remove invalid detections. The filter works by first removing tags not in found the tagging data, then by removing multipath detections, and finally by removing fish with detection histories which do not meet a minimum frequency and duration threshold (PRI filter).

The inputs to the program are tagging and nodes data, and raw detections data from ATS autonomous nodes, SC autonomous nodes, or cabled node data. The output is events, hits, and a log file, described below.

This software development was funded by Portland District, Army Corps of Engineers, under a Government Order with the U.S. Department of Energy, Contract DE-AC05-76RL01830.

1 Data Formats

1.1 Tagging Data

Tagging data provide release date and nominal PRI for each fish. Any tag code encountered during filtering that is not present in the tagging data file is ignored.

The file format is CSV with header optional.

Table 1. Tag data field specification

Field #	Description	Format	Example	Notes
1	Tag Code	*G7xxxxxxx* (where x is hex digit)	G72610B1E	
2	Release Date	*yyyy-mm-dd hh:mm:ss*	2012-10-15 21:38:12	Not currently used
3	Nominal PRI	*Number*	3, 3.0	

1.2 Node Data

Any node code (node id + date) encountered during filtering that is not provided in the node data file is ignored.

The file format is CSV with header optional.

Table 2. Node data field specification

Field #	Description	Format	Example	Notes
1	Node Code	node id + date	12090_100720	
2	Location	*String*	CR1086.0_04	Not currently used

1.3 Detection Data

The program recognizes three types of detection data: ATS, SC, and cabled node data. The program searches for valid files, those which match the expected format, recursively through folders under the user-defined data folder. A raw data file is considered valid based a match of the expected location of a correctly formatted tag code, within the first 100 rows, according to the following:

Data Type	Expected Tag Code Field
ATS	6
SC	1
Cabled	6

This process allows nondetection data files, such as summary reports, configuration files, etc., to be present in the data folder.

1.3.1 ATS Auto Node Data

ATS data files are processed as follows. The node code is read from the file based on convention. The node code is expected to follow the letters RX at the start of the filename. For example, the filename RX11149_120522.csv would indicate a node code of 11149_120522. The resulting node code must be present in the node data for the file to be accepted.

The file is then parsed, searching for any row with a valid tag code and valid date. A tag is considered valid if it has the format G7xxxxxxx (i.e., G7 followed by 7 hexadecimal digits) (e.g., G72610B1E). The tag code must be present the tagging data to be considered valid.

Dates are valid if the format matches *mm/dd/yyyy hh:mm:ss.mmmmmm*, where the precision is given in microseconds (e.g. 05/21/2012 11:13:13.540598). The signal strength is expected to be in column 11. All other fields are ignored.

When a row meets these requirements, the tag code, node code, reception time, and signal strength are stored for processing. All other rows are ignored.

1.3.2 SC Auto Node Data

SC data files are processed as follows. The node code is read from the file name based on convention. The node code should immediately follow RXN in the filename. For example, the filename RXN06009_111006.DAT.processed would indicate a node code of 6009_111006 (leading 0's are ignored). This node code, without the leading zeros, must be present in the node data.

The file is then parsed, searching for any row with a valid tag code. A tag is considered valid if it has the format G7xxxxxxx (i.e. G7 followed by 7 hexadecimal digits) (e.g. G72610B1E). The tag code must be present in the tagging data to be considered valid.

The detection time is read from fields 5 and 6. In SC data, field 5 is a hexadecimal string representing the number of seconds since 2000-01-01 00:00:00. Field 6 is additional seconds, also a hexadecimal number in units of 12 microseconds.

When a row meets these requirements, the tag code, node code, converted reception time, and signal strength are stored for processing. All other rows are ignored.

1.3.3 Cabled Data

Cabled data is scanned for rows with valid tag codes, which is defined as having the format G7xxxxxxx, where x is a hexadecimal digit (e.g. G72610B1E). The tag must found be in the tagging data file to be accepted. The hydrophone name is taken from the 3rd column of the input and must also be found in the node data file.

The date is found by searching in a given row for a string formatted as mm.dd.yy (e.g., 11.20.2012). The signal time, in sub-microseconds (7 digits), is read from column 5 and combined with the calendar date. The signal strength expected to be in column 8.

When a row meets these requirements, the tag code, node code, reception time, and signal strength are stored for processing. All other rows are ignored.

2 Output

The program outputs individual accepted hits, and events, which summarizes the hits into groups of hits that fall within a restricted timeframe. Both files are output in CSV format. A log file is output along with the main program output to provide detailed processing information and error reporting associated with the run that produced the output. The convention used in naming the files records the data type, number of hits, and hit interval in the filename. For example, the result of processing ATS data with a 4-hit, 12-interval PRI setting will result in the following output:

> fast_ats_4_12_events.csv
> fast_ats_4_12_hits.csv
> fast_ats_4_12_log.txt

Each of these output files is described below.

2.1 Events

An event is a grouping of hits with gaps no larger than the PRI filtering window size. A window size is roughly the number of intervals times the PRI for a fish. For example, with 12 intervals and a PRI of 3, the window will be about 36 seconds. The event will have hits such that no gaps between hits are larger than this. The numbers above are rough estimates, and the actual window will vary depending on estimated PRI and other considerations. Event dates are the first and last dates of the hits or messages that make up the event, truncated to the second.

Table 3. Event output field description

Field	Description
Node Code	Node code as found in the node data
Tag Code	Tag code is the acoustic tag code for the event
Event ID	Event id is a unique id across the entire run for the event
First Date	First date is the date of the first hit for the event (*yyyy-mm-dd hh:mm:ss*)
Last Date	Last date is the date of the last date for the event (*yyyy-mm-dd hh:mm:ss*)
Number of Hits	The number of hits represented summarized by the event

2.2 Hits

Hits are the individual detection records which make up an event. Time precision is to the sub-microsecond (1e-7 seconds), which matches the precision from the cabled data input.

Table 4. Hit output field descriptions

Field	Description
Node Code	Node code as found in the node data
Tag Code	Tag code is the acoustic tag code for the event
Event ID	Id of the event the hit belongs to
Time	The time of the hit (*yyyy-mm-dd hh:mm:ss.xxxxxxx*)
SNR	Signal to noise ratio for the hit

2.3 Log File

There are two log files: a system log file and a run log file. The *system* log file keeps track of log messages during the duration of program execution. It may be viewed by selecting **Output=>View log file** from the main menu. The *run* log file records messages for one run and provides a record of the settings and status message for a given run.

3 Settings

Settings can be accessed through the menu by selecting **Edit=>Advanced settings**.

3.1 PRI Filter Settings

There are separate settings for autonomous node filtering and for cabled filtering. In both cases the parameters are:

- Hits – The minimum threshold for hits during the time window
- Intervals – The number of PRI intervals which defines the time window, or duration, for the hits. This window will be roughly PRI * Intervals long.

The default PRI settings for autonomous node filtering are 4 hits and 12 intervals. For cabled node data filtering, the default is 6 hits in 12 intervals.

3.2 Output Settings

The program can output Events, Hits, or both. By checking Output Estimate PRIs, the estimated PRI of the initial hit in an event will be output as a final column. When events are merged this value will be overwritten with the estimated PRI merged in event.