



ROSTER 2.9 USER'S MANUAL

RIVER-OCEAN SURVIVAL AND TRANSPORTATION EFFECTS ROUTINE

COLUMBIA BASIN RESEARCH

SCHOOL OF AQUATIC AND FISHERY SCIENCES

W UNIVERSITY *of* WASHINGTON

Program ROSTER: River-Ocean Survival and Transportation Effects Routine

Rebecca A. Buchanan, Jim Lady, Peter Westhagen, John R. Skalski

Columbia Basin Research
School of Aquatic and Fishery Sciences
University of Washington
1325 4th Avenue, Suite 1820
Seattle, WA 98101-2509

2007

Executive Summary

Program ROSTER (River-Ocean Survival and Transportation Effects Routine) analyzes joint juvenile and adult PIT-tag detection data from migrating salmon, and estimates a variety of performance measures related to salmonid migrations. Program ROSTER accounts for juvenile transportation, adult age structure, and known removals at detection sites, and provides maximum likelihood estimates and standard errors of juvenile inriver survival, ocean survival, adult inriver survival, smolt-to-adult ratios (SARs), transportation effects, and the adult age distribution. Several types of transportation effect measures are estimated, including dam-specific and systemwide measures, measures for untagged fish, and differential mortality (D). The user may test for effects of juvenile transportation on the upriver adult migration. Program ROSTER is suitable for analyzing data from spring and summer Chinook salmon, sockeye salmon, and steelhead.

Contents

1	Introduction	1
2	Running ROSTER	3
2.1	Program ROSTER	3
2.2	Obtaining ROSTER	3
2.3	Getting Started	3
2.4	Data	7
2.5	Estimation	8
2.5.1	Model Definition	9
2.5.2	Parameters Set by the User	11
2.5.3	Estimation	12
2.6	Reports	12
2.6.1	Data Summary Report	13
2.6.2	Sufficient Statistics Report	13
2.6.3	Parameters Report	14
2.6.4	Parameter Indices	15
2.6.5	Estimation Results Report	15
2.6.6	Covariance Matrix	15
2.6.7	Parameter Estimates and Performance Measures	16
2.7	Diagnostic Tools	16
3	ROSTER Theory	19
3.1	The Model Behind ROSTER	19
3.2	Extent of Transportation Effects	20
3.3	Tagged Versus Untagged Fish	23
3.4	Performance Measures	24
3.4.1	Transportation Effects	25

3.4.1.1	Site-Specific Transportation Effects Measures	25
3.4.1.2	Systemwide Transportation Effects Measures	27
3.4.2	Differential Mortality	29
3.4.2.1	Site-Specific Differential Mortality Measures	30
3.4.2.2	Systemwide Differential Mortality Measures	31
3.4.3	Juvenile Survival	32
3.4.4	Ocean Return Probabilities	32
3.4.5	Adult Return Age Distribution	33
3.4.6	Adult Upriver Survival	34
4	Fitting ROSTER and Model Selection	37
4.1	Guidance on Fitting ROSTER	37
4.2	Model Selection	39
4.2.1	Likelihood Ratio Tests	39
4.2.2	Akaike’s Information Criterion	40
5	Model Definition File Format	43
5.1	Validate Element	43
5.2	Fix Element	44
5.3	Equate Element	44
5.4	InitialSeeds Element	45
6	Data Sources	47
6.1	Estuary: Towed Array Data	47
6.2	Avian Predation Data: Detections From Bird Colonies	48
6.3	Harvest Data	48
7	Sample Data Sets	49
7.1	Example 1: 1999 Spring Chinook Salmon	49
7.2	Example 2: 2002 Spring Chinook Salmon	57

List of Figures

2.1	Data tab	4
2.2	Estimation tab	5
2.3	Schematic view	6
2.4	Toolbar icons	6
2.5	Loading data	7
2.6	Defining the model	10
2.7	Parameters set by user	12
2.8	Running estimation routine	12
2.9	Reports menu	16
2.10	Fletch parameters	17
2.11	Introduce noise to seeds	18
3.1	Schematic of migration paths	20
3.2	Schematic of migration paths of transported and inriver fish	21
7.1	Example 1 data summary	50
7.2	Example 1 default model file, top	52
7.3	Example 1 default model file, bottom	53
7.4	Example 1 results	55
7.5	Example 1 performance measures	56
7.6	Example 2 data summary	59
7.7	Example 2 default model file, bottom	60
7.8	Example 2 edited model file, bottom	62
7.9	Example 2 results	63

List of Tables

2.1	Detection history symbols used by ROSTER.	8
2.2	Sufficient statistics	13
3.1	Model parameters estimated by ROSTER	22

1.0 Introduction

This manual gives an overview of how to run Program ROSTER (Chapter 2), as well as background information on the theory behind ROSTER (Chapter 3). Some guidance on troubleshooting and fitting difficult data sets is offered in Chapter 4. The format of the model file is described in Chapter 5. Appropriate data sources are discussed in Chapter 6, and example analyses of two sample data sets are given in Chapter 7.

The statistical model used in ROSTER is presented in a manuscript currently in press in the *Journal of Agricultural, Biological, and Environmental Statistics*. A second paper discussing the different types of transportation effects measures estimated by ROSTER is available in the *North American Journal of Fisheries Management*. The two papers are listed below, and are available on the Columbia Basin Research Web page at <http://www.cbr.washington.edu/paramest/roster/>.

Buchanan, R.A., Skalski, J.R., and Smith, S.G. (2006), "Estimating the Effect of Transportation From Different Vantage Points and Management Perspectives," *North American Journal of Fisheries Management*, 26, 460-472.

Buchanan, R.A., and Skalski, J.R., "A Migratory Life-Cycle Release-Recapture Model for Salmonid PIT-Tag Investigations," *Journal of Agricultural, Biological, and Environmental Statistics*, in press.

2.0 Running ROSTER

2.1 Program ROSTER

ROSTER is designed to be easy to use while providing full computational capabilities to the expert user. The process can be broken down into three main steps: loading the data, defining the model file, and running the estimation routine.

2.2 Obtaining ROSTER

ROSTER can be downloaded as a pre-compiled binary for Windows from the ROSTER web site at <http://www.cbr.washington.edu/paramest/roster/>. The program is distributed as a zip file which will extract into a folder named Rosterx.x, where x.x is the version number. There are several library files (having the suffix .dll) distributed with ROSTER that are necessary for the program to run. These libraries should be kept in the same folder as the ROSTER executable file (having the suffix .exe).

2.3 Getting Started

The initial program view is a splash screen that displays for a few seconds and gives way to the main program interface. You can go directly to the program interface by clicking on the splash screen. The program interface has two tabs, labeled Data (Figure 2.1) and Estimation (Figure 2.2). The data set is loaded using the Data tab, and the Estimation tab is used to define the model file and run the estimation routine.

The program interface is the primary view of the program. It is here that the data set is loaded, the model is configured, and program status is displayed. Another view of the program is the schematic view (Figure 2.3), which gives a graphical display of the study environment and is available only when a data set has been loaded. To show the schematic view, use the View menu or the associated icon in the program toolbar. The View menu and associated toolbar icon are shown in Figure 2.4.

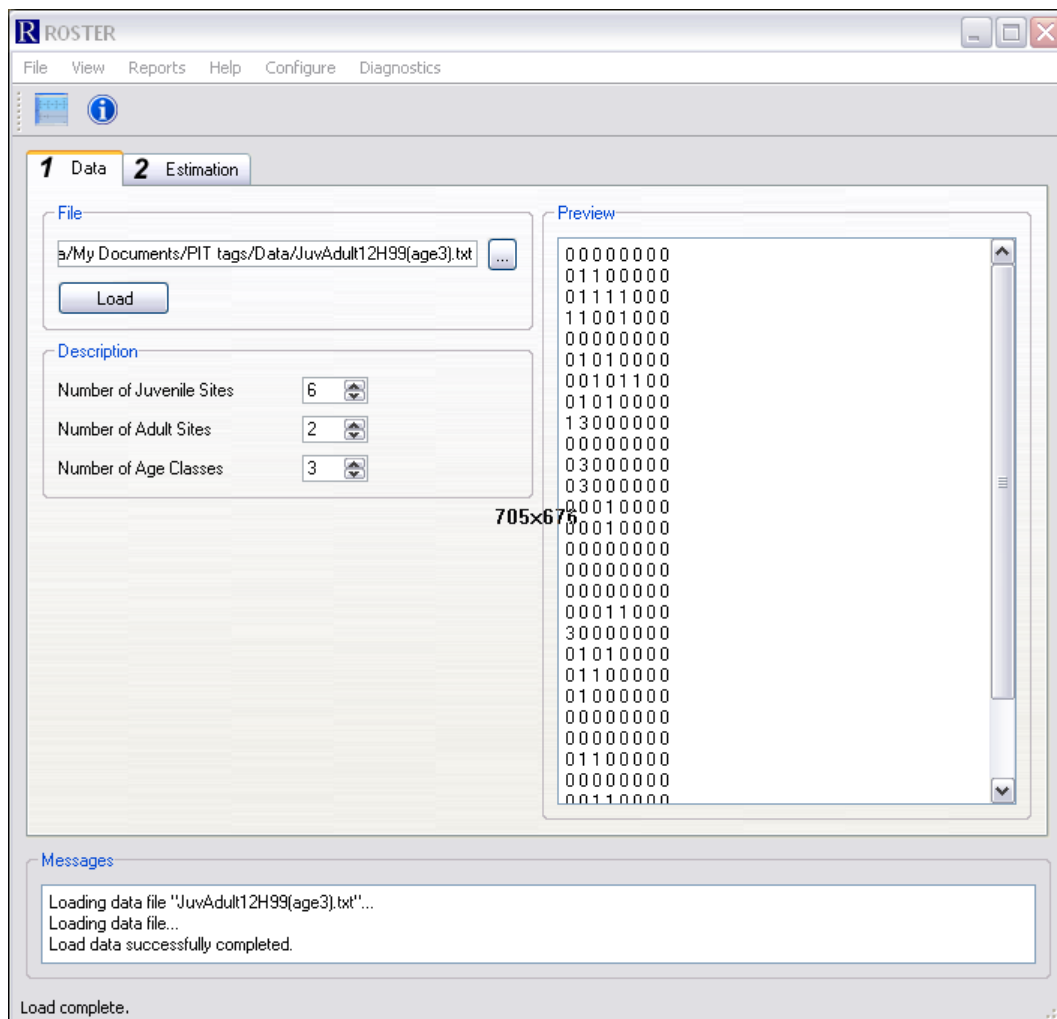


Figure 2.1: Program interface, showing the Data tab.

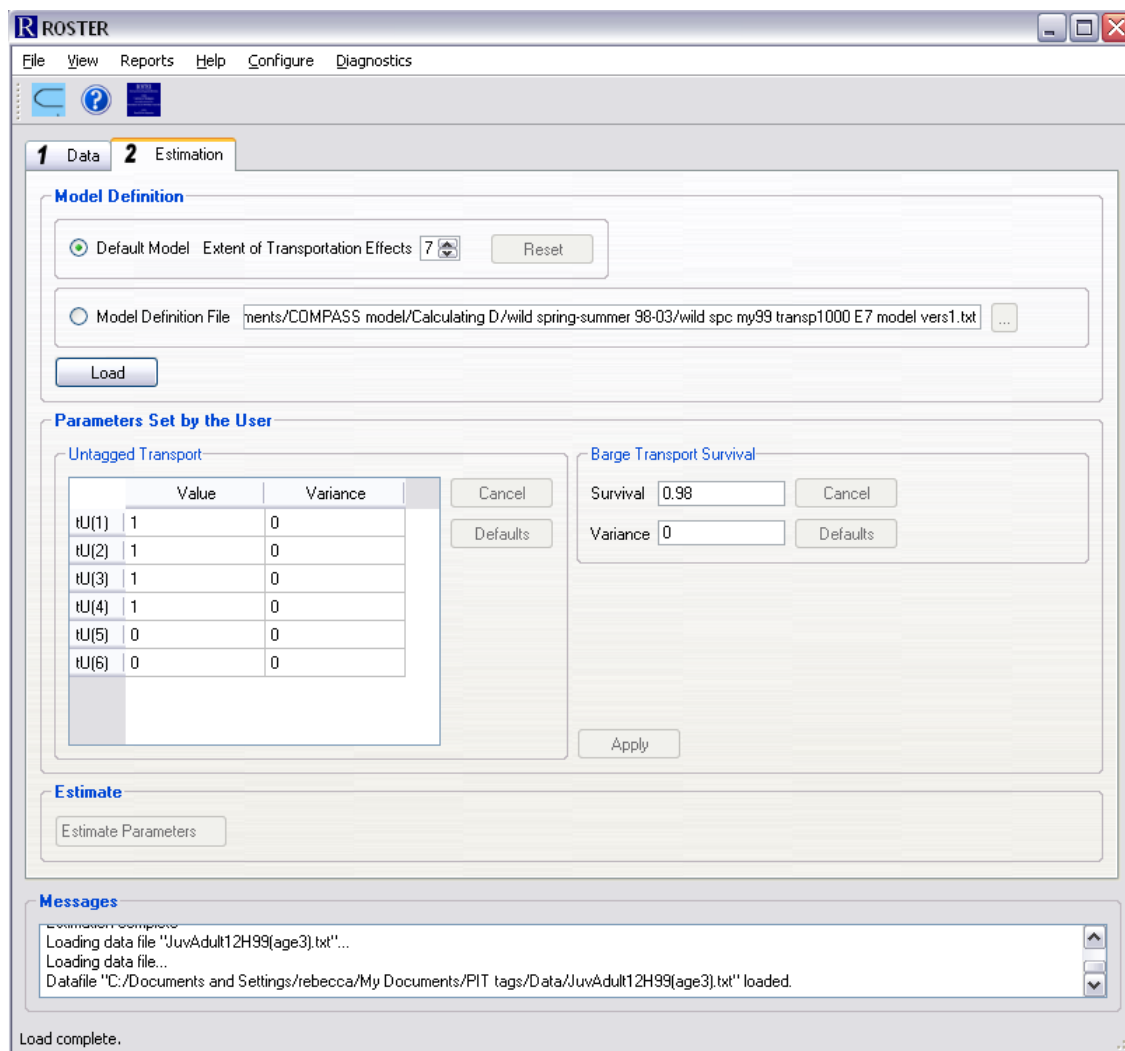


Figure 2.2: Program interface, showing the Estimation tab.

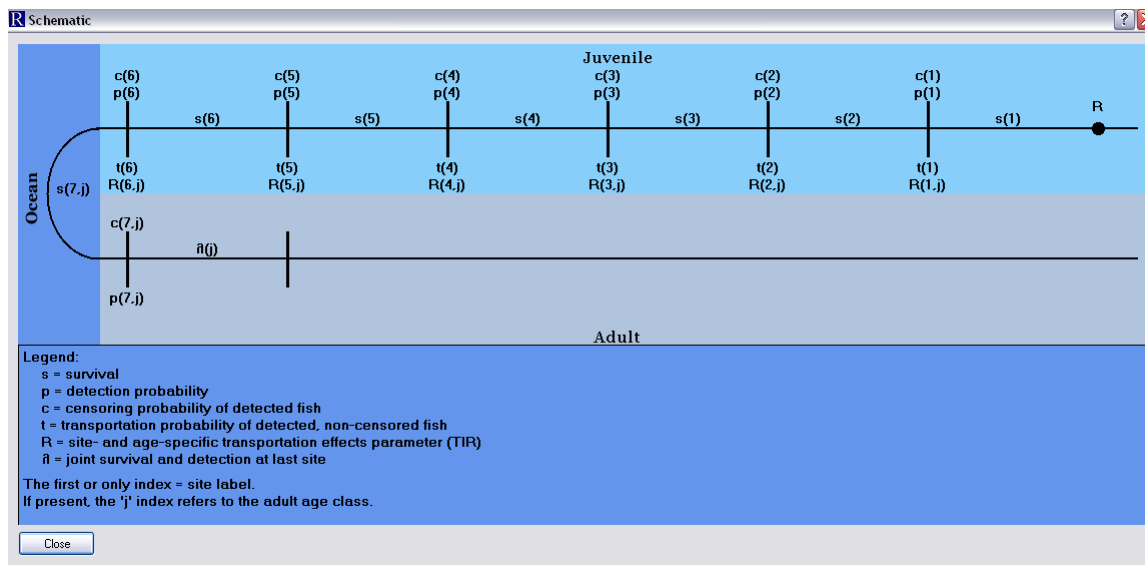


Figure 2.3: Schematic view - 6 juvenile sites and 2 adult sites.

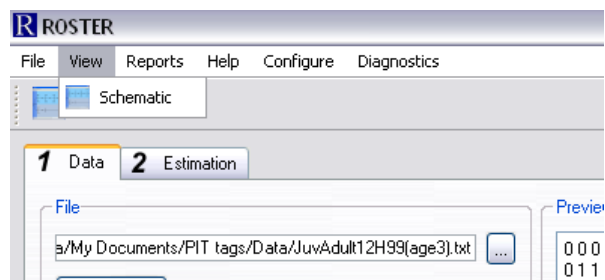


Figure 2.4: Toolbar icons.

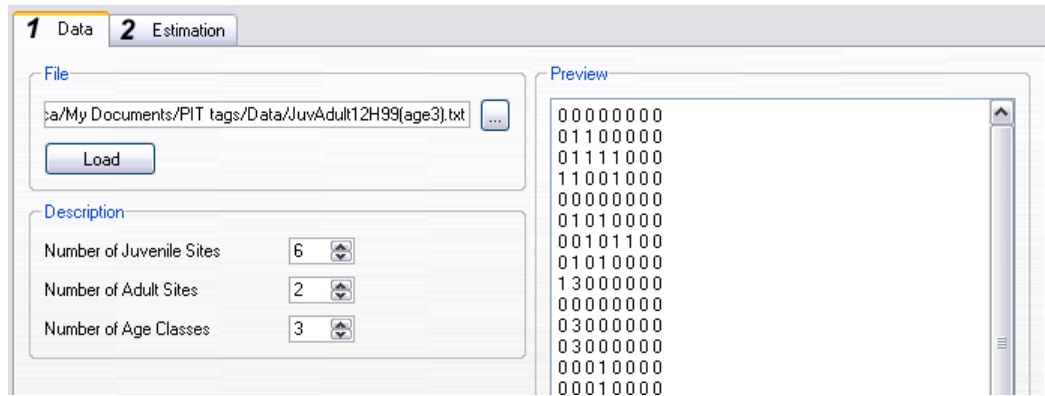


Figure 2.5: Step 1 - Loading data.

2.4 Data

The first step in running ROSTER is to load the data on the Data tab of the program interface (Figure 2.5). The Data tab has three sections: File, Description, and Preview. A Messages window appears at the base of the program interface. In the File section, the data file containing the capture histories is selected by clicking on the “...” button and navigating to the data file. In the Description section, certain parameters describing the capture histories are entered, including the number of juvenile sites in the capture histories, the number of adult sites, and the number of age classes for returning adults. A preview of the selected data file is provided to the right in the program interface as a quick reminder of the selected data file composition. When the load button is clicked, the data will be loaded and two reports will be generated, available under the Reports menu. The Data Summary report contains basic statistics from the data set, and the Sufficient Statistics report shows the statistics with a more technical presentation. These reports are described in more detail in Section 2.6.

The ROSTER data file consists of rows of detection (or capture) histories, each row representing the detection history for one fish. Each detection history consists of fields separated by white space where a field represents the observation record for the fish at a particular capture site. Detection histories may be generated from PIT-tag data available from the Pacific States Marine Fisheries Commission’s PTAGIS database system using UW PIT-tag processing software PitPro.

The fields in the detection histories are separated into two groups: juvenile detections and adult detections. A juvenile field in the detection history will have one of four values: 0, 1, 2, or 3 representing not detected, detected and returned to river, detected and removed from the migrating population, or detected and transported, respectively. An adult detection history field consists of either 0,

Table 2.1: Detection history symbols used by ROSTER.

Symbol	Life Stage	Meaning
0	Both	Not detected
1	Juvenile	Detected and returned to river
2	Juvenile	Detected and removed or detected but outcome unknown
3	Juvenile	Detected and transported
A	Adult	Detected as age class 1 adult and returned to river
a	Adult	Detected as age class 1 adult and removed
B	Adult	Detected as age class 2 adult and returned to river
b	Adult	Detected as age class 2 adult and removed
C	Adult	Detected as age class 3 adult and returned to river
c	Adult	Detected as age class 3 adult and removed
D	Adult	Detected as age class 4 adult and returned to river
d	Adult	Detected as age class 4 adult and removed
E	Adult	Detected as age class 5+ adult and returned to river
e	Adult	Detected as age class 5+ adult and removed

meaning not detected, or the letter *A*, *a*, *B*, *b*, *C*, *c*, *D*, *d*, *E*, or *e*. The letters indicate both detection and age class of the detected returning adult, with *A* indicating the first age class (ocean-1 fish), *B* the second age class, etc. Capital letters indicate that the adult was returned to the river after detection, and lowercase letters indicate known removal (i.e., “censoring”) of the adult upon detection. All potential symbols are listed in Table 2.1.

When running PitPro to process ROSTER data, the user must include two extra types of PTAGIS files, a recovery file and a mortality file. Without these extra files, PitPro will not be able to identify removed tags, and survival probabilities will not be adjusted for known removals. For more detail on how to prepare PTAGIS files for Program ROSTER, see PitPro at <http://www.cbr.washington.edu/paramest/pitpro/>.

The number of juvenile sites, adult sites, and adult age classes for the data set must be entered in the appropriate boxes before loading the data. The message window at the bottom of the program interface indicates when the data file has finished loading (2.1).

2.5 Estimation

Once the data set has been loaded, click on the Estimation tab in the program interface (Figure 2.2). The Estimation tab has three sections: Model

Definition, Parameters Set by the User, and Estimate. The Messages window remains at the bottom of the program interface.

2.5.1 Model Definition

The first step of the estimation process is defining and loading the model on the Estimation tab (Figure 2.6). ROSTER provides a default model, based on statistics compiled from the data set and the adult site to which transportation effects will be calculated (“Extent of Transportation Effects,” see Chapter 3.2). Alternatively, the expert user may modify the default model by saving the model file generated by ROSTER, and then editing it and loading the new model file. More information on this is given below and in Chapters 4.1, 5, and 7. The selected model is loaded by clicking on the Load button.

The model used in ROSTER is a statistical likelihood that is fit to the data to estimate juvenile, ocean, and adult survival probabilities of transported and non-transported fish, juvenile and adult detection and censoring probabilities, and transportation parameters. The most general, or “full,” statistical model estimates separate values for each of these parameters, distinguishing between transported and non-transported fish and among transport groups in classifying adult parameters. It is often both difficult and unnecessary to treat all parameters as unique. Instead, it may be sufficient to equate certain model parameters. In particular, it may be reasonable to equate adult parameters among all transported fish, or perhaps for all fish, regardless of juvenile migration method (transport or non-transport).

Program ROSTER is designed to run with as little input from the user as possible. A default statistical model is configured automatically by the program based on observed statistics from the data. This model will be a simplified version of the full model, tailored to the data set. The default model generated by ROSTER estimates unique upriver adult survival parameters for fish transported as juveniles and those not transported as juveniles until the site selected in the “Extent of Transportation Effects” window. The default model also estimates unique adult detection and censoring parameters for transported and non-transported fish at all adult detection sites up to the site identified as the extent of transportation effects. Additionally, the default model identifies which parameters to estimate and sets initial values for those parameters to be used by the estimation routine. In many cases, the default model will be sufficient for complete data analysis, in which case it is not necessary to save or edit a model definition file. To use the default model, make sure that the “Default Model” button is clicked on the Estimation tab, and click on the Load button. Alternatively, the default model may be used as a starting point for user-defined models.

If the user chooses to modify the default model, then it is necessary to save



Figure 2.6: Step 2 - Defining the model.

and edit the model definition file. The model file is used by ROSTER to define changes in the structure of the statistical model compared to the full model described above; the statistical model used by ROSTER is not itself shown in the model file. In general, the changes defined in the model file consist of (1) equating parameters, (2) fixing parameters to a particular number, or (3) specifying the initial values (“seeds”) to be used in fitting the model. The model file also specifies the number of juvenile and adult detection sites and adult age classes, as entered on the Data tab of the ROSTER interface, as well as the site identified as the extent of transportation effects on the Estimation tab. User-specified values for untagged transport parameters and barge transport survival are also listed in the model definition file.

After the default model is selected and loaded, it can be saved in a model file and edited. To save a copy of the model definition file for the default model, select “Save Model Definition...” from the File menu. The file can then be edited using an XML editor or any text editor. The user may edit the model file by changing the parameters that are equated, fixing parameters so that they are not estimated, removing parameters from the list of fixed parameters so that they are estimated, or setting initial values. For guidance on fitting models, refer to Chapter 4.1.

Once the model file has been edited and saved, it is available for use by the program. Click the button next to “Model Definition File” in the Model Definition section, and select the file by clicking on the “...” button. After selecting the desired model definition file, click on the “Load” button to load the model.

With the model loaded, two more reports are available under the “Reports” menu: the Parameters report shows initial values for estimation using the loaded model, and the Parameter Indices report shows the structure of the model. These reports are described in more detail in Section 2.6.

2.5.2 Parameters Set by the User

The second section of the Estimation tab allows the user to modify values of parameters that are always fixed in ROSTER, rather than estimated (Figure 2.7). These fixed parameters do not affect the model fit or estimates of the model parameters, but they are used in calculating estimates for certain performance measures. The first set of fixed parameters pertains to estimating transportation effect measures for untagged fish, under the heading Untagged Transport (see Chapter 3 for more details). The parameters to be set are the dam-specific transportation probabilities for untagged fish, representing the proportion of the collected untagged fish that are transported from each juvenile detection site (i.e., dam). The default values are 1 (100% transported) at transport dams, and 0 at other juvenile sites. The default variance is 0 for each parameter. The user may enter different values and positive variances if available. The second parameter set by the user is the Barge Transport Survival, or the survival probability of transported fish on the barge (or truck) during transportation. This parameter is used to estimate various measures of D , differential mortality. The default value of barge transport survival is 0.98 (98%), with a variance of 0. The user may alter the survival parameter or enter a positive variance for the survival parameter, if available.

Whether the user loads the default model or a model definition file, the loaded model should specify values for the untagged transport parameters and the barge transport survival. When the model is loaded, the “Parameters Set by the User” section of the ROSTER interface will show the values specified in the loaded model. The user may change the values of the untagged transport parameters and the barge transport survival and their variances by highlighting the appropriate box in the ROSTER interface and editing the values. Clicking the “Cancel” button will revert the parameters to their previously saved values, while clicking the “Defaults” button will revert the parameters to their default values. The user must tell ROSTER to apply the values from the interface if they have been changed, either by clicking the “Apply” button or by clicking “Estimate Parameters” and selecting “save” from the dialog box that opens. Either action will tell ROSTER to calculate performance measures using the new values for the untagged transport parameters and barge transport survival, but without overwriting the values saved in the model definition file. In order to save a version of the model definition file with the new values of the user-set parameters, the user must first click “Apply” and then select “Save Model Definition ...” under the File menu.

Parameters Set by the User

Untagged Transport

	Value	Variance
tU(1)	1	0
tU(2)	1	0
tU(3)	1	0
tU(4)	1	0
tU(5)	0	0
tU(6)	0	0

Buttons: Cancel, Defaults, Apply

Barge Transport Survival

Survival: 0.98 (Buttons: Cancel, Defaults)

Variance: 0 (Buttons: Defaults)

Figure 2.7: Set parameters for untagged transport and barge transport survival.

Estimate

Estimate Parameters

Messages

Iteration 49, function value = 119784.28170419
 Fletch iteration 49, function value = 119784.28170419
 Fletch iteration 50, function value = 119784.28170419
 Estimation complete

Estimation complete.

Figure 2.8: Running the estimation route.

2.5.3 Estimation

The final step of the estimation process is to run the estimation routine (Figure 2.8). Program status and any error messages are displayed in the message window at the bottom of the program interface. The function value reported in the message window is the negative log-likelihood value for the current iteration; this value depends on the parameter values being used and is minimized by the estimation routine. An extremely large function value such as $1.8e+308$ indicates an error in the model file or estimation routine. A final message of success or failure of the estimation routine is shown in the message window. If the estimation routine fails to converge, then an error message appears, and it is necessary to reload the model in order to examine reports or attempt to rerun the estimation routine.

2.6 Reports

Several types of reports are generated by ROSTER throughout the data loading and estimation process, and are available under the Reports menu

(Figure 2.9). Once the data set is loaded, the Data Summary and Sufficient Statistics reports are generated. After the model is loaded, the Parameters report and Parameter Indices report are available. After the estimation routine has been successfully run, the Estimation Results report and Covariance Matrix report are also available, as well as text files containing estimates of the model parameters and performance measures.

2.6.1 Data Summary Report

After the data set has been successfully loaded, a data summary is available by selecting “Data Summary” from the Reports menu. The data summary provides a useful report of basic statistics including size of the release group, number of detections at each juvenile site, number of detections at each adult site classified by juvenile transport group and adult age class, number of fish transported at each juvenile site, and number censored at each juvenile site and at each adult site classified by juvenile transport group and adult age class. The Data Summary report gives a quick, useful overview of the data set.

2.6.2 Sufficient Statistics Report

A more detailed and technical version of the Data Summary report, the Sufficient Statistics report is useful mostly for the expert user fitting a model other than the default model. The Sufficient Statistics report shows the statistics based on the data that the statistical model actually uses. Table 2.2 describes the nature of the sufficient statistics in this report.

Table 2.2: Statistics listed in Sufficient Statistics report. Detection sites are numbered consecutively. Site 0 is the initial release location, v is the number of juvenile sites, u is the number of adult sites, and w is the number of adult age classes. Site $v + 1$ is the first adult site, and site $v + u$ is the last adult site.

Statistic	Definition
a_i	Number of juveniles detected at site i (includes those transported from i or removed at i); $i = 1, \dots, v$
a_{ijC}	Number of non-transported fish detected at site i as an age- j adult; $i = v + 1, \dots, v + u$; $j = 1, \dots, w$
a_{ijT_k}	Number of site- k transport fish detected at site i as an age- j adult; $i = v + 1, \dots, v + u$; $j = 1, \dots, w$; $k = 1, \dots, v$
b_i	Number of juveniles detected at site i , re-released to the river, and detected at a later site (juvenile or adult); $i = 0, \dots, v$

Table 2.2 continued

Statistic	Definition
b_{iT}	Number of juveniles detected and transported from site i and detected later at an adult site; $i = 1, \dots, v$
b_{ijC}	Number of non-transport fish detected and re-released to the river at site i , and next detected as an age- j adult; $i = 0, \dots, v + u - 1$; $j = 1, \dots, w$
b_{ijT}	Number transported from juvenile site i and next detected as an age- j adult; $i = 1, \dots, v$; $j = 1, \dots, w$
b_{ijT_k}	Number of site- k transport fish detected and re-released to the river at site i as an age- j adult, and detected later (as age- j adult); $i = v + 1, \dots, v + u - 1$; $j = 1, \dots, w$; $k = 1, \dots, v$.
d_i	Number of juveniles censored (removed) at site i ; $i = 1, \dots, v$
d_{ijC}	Number of non-transported fish censored (removed) as age- j adults at site i ; $i = v + 1, \dots, v + u - 1$; $j = 1, \dots, w$
d_{ijT_k}	Number of site- k transport fish censored (removed) as age- j adults at site i ; $i = v + 1, \dots, v + u - 1$; $j = 1, \dots, w$; $k = 1, \dots, v$
h_i	Number of juveniles transported from site i ; $i = 1, \dots, v$
g_i	Number of fish <i>not</i> transported from any site $k \leq i$ that are detected <i>after</i> juvenile site i (includes fish transported from sites after site i); $i = 0, \dots, v - 1$
g_{ijC}	Number of non-transported fish detected <i>after</i> site i as age- j adults; $i = v, \dots, v + u - 1$; $j = 1, \dots, w$
g_{ijT_k}	Number of site- k transport fish detected <i>after</i> site i as age- j adults; $i = v + 1, \dots, v + u - 1$; $j = 1, \dots, w$; $k = 1, \dots, v$

2.6.3 Parameters Report

The Parameters report is available in a preliminary version after the model has been loaded, and in a completed version after successfully running the estimation routine. Before running the estimation routine, the preliminary report shows the Method of Moments estimates for each parameter, with all adult parameters showing the same estimate across transport and non-transport groups. The Method of Moments, or MoM, estimate is usually a good educated “guess” as a starting value for the parameter for the estimation routine. The first time the estimation routine is run using the default model, these MoM estimates are used as initial values or “seeds.” The Parameters report also indicates which parameters are fixed rather than estimated, and the value at which they are fixed. If the default model does not run successfully, the Parameters report may indicate where the problem is. For example, if the MoM estimate for a parameter is 0 or ∞ (“inf”),

then the default model will not run properly, and the user will need to modify the model file.

After the estimation routine has successfully completed, the Parameters report shows three values for each parameter: the MoM estimate, the seed (initial value) actually used, and the final maximum likelihood estimate (MLE) found by the estimation routine. Unless the user tells ROSTER otherwise, the MoM estimate and the seed will be equal. In many cases, the MoM estimate and the MLE will be equal as well.

Certain ROSTER features allow the user to alter the initial values used in running the model. In these cases, the seeds shown in the Parameters report will be different from the MoM estimates. It may still be the case that the MLE is equal, or nearly equal, to the MoM estimate for each parameter. Running the estimation routine many times with different starting values is a good way to check that the estimation routine has correctly found the true maximum likelihood estimates.

2.6.4 Parameter Indices

The Parameter Indices report is available after the model is loaded. It indicates the structure of the loaded model by showing which parameters are estimated and which parameters are fixed (not estimated). Estimated parameters show a number in the Index column; equated parameters share a common index.

2.6.5 Estimation Results Report

Upon successful completion of the estimation process, the program will generate and display an Estimation Results report which includes estimates of the parameters and performance measures. This report can also be shown by selecting “Estimation Report” under the Reports menu. Performance measures that exclude jacks are shown at the bottom of the report. Model parameters and performance measures are described in detail in Chapter 3.

2.6.6 Covariance Matrix

The Covariance Matrix report gives the variance-covariance matrix estimated for the model parameters by ROSTER. Each entry in the matrix is the covariance between two parameter estimates. Only parameters that are actually estimated, not fixed, are included in the matrix. The Covariance Matrix report may be saved and then imported into other software for examination.

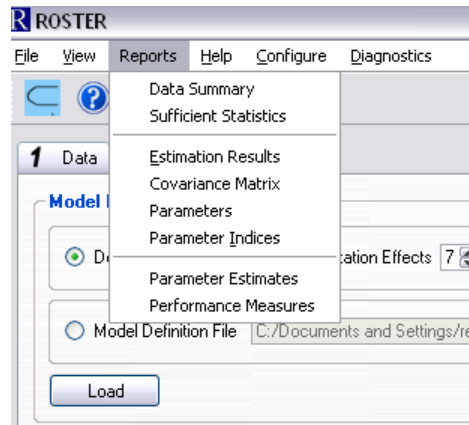


Figure 2.9: Reports menu.

2.6.7 Parameter Estimates and Performance Measures

The Parameter Estimates and Performance Measures reports list the estimates and standard errors of all possible model parameters and performance measures, respectively, in simple text files that may be saved and imported into other software. They are available under the Reports menu.

2.7 Diagnostic Tools

Program ROSTER includes several tools designed to help the user find the best model fit. In general, the initial values calculated by ROSTER will be very close, if not equal, to the actual MLEs. However, the Method of Moment estimators used by ROSTER as initial values will not always provide the best model fit, and in some cases they will not be the best initial values. Also, it is possible that the estimation routine misses the best model fit because of the step size used or because it ended too soon. These problems can be detected and avoided using ROSTER's diagnostic tools.

Program ROSTER uses the optimization routine FLETCH to find the MLEs. The precision, step size, and maximum number of iterations used by FLETCH can be changed under the Configure menu (Figure 2.10). In general, a smaller precision makes the stopping criterion more stringent for the optimization routine. The step size indicates how large a change in each parameter is considered between iterations. If the step size is too large, the optimizer may “step over” the true MLE and never find it. If the step size is too small, the optimizer may not be able to leave the initial values (seeds) even when the true MLE is elsewhere. Thus, it is wise to rerun

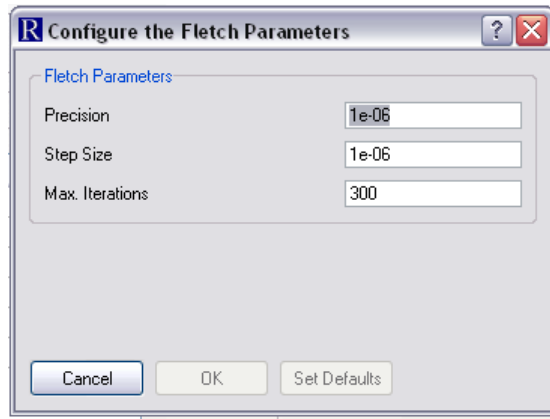


Figure 2.10: Configure the FLETCH Parameters.

the estimation routine several times, varying the step size with a fairly small precision. The default values for precision and step size are both $1e - 06$.

In addition to changing the FLETCH parameters, the user may use alternative seeds (initial values) to start the optimization. This can be useful; if different seeds lead to different final estimates, then the optimizer may be missing the correct MLE. The Parameters report is useful for keeping track of the seeds used and the final estimates found when using these tools. Under the Diagnostics menu is the “Parameter Seeds...” option, which allows the user to introduce random noise into the default MoM seeds (Figure 2.11). To introduce random noise into the parameter seeds, select “Introduce random noise to initial seeds,” select either Proportional (default) or Absolute type, and enter the amount of noise desired. If parameters are very close to boundaries (e.g., 0 or 1), then a large amount of noise may lead to sets of initial parameter values that are not viable. In these cases, running the estimation routine will lead to an error message. Rerunning the estimation routine may produce a set of random initial values that are viable, or it may be necessary to reduce the amount of random noise introduced to the seeds. In order to turn off the random noise in the seeds, unselect “Introduce random noise to initial seeds” and click “OK.” Alternatively, the user may specify seeds in the model file; see Section 5.4.



Figure 2.11: Introduce random noise to seeds, on either a proportional or absolute scale.

3.0 ROSTER Theory

3.1 The Model Behind ROSTER

Historically, juvenile and adult PIT-tag data have been analyzed separately, and records of transported smolts have been removed (censored) from PIT-tag data to estimate smolt survival. The results of these separate analyses can be inefficient and deceptive. A new PIT-tag release-recapture model¹ has been developed that models the life cycle of Pacific salmonids from the initiation of the juvenile migration as smolts to the end of the spawning migration as adults. This model jointly analyzes juvenile and adult PIT-tag data and uses both juvenile and adult data from transported individuals. Multiple performance measures for monitoring the juvenile and adult migrations and assessing recovery efforts are produced by the model, including estimates of juvenile inriver survival, ocean survival, adult inriver survival, transportation effects (e.g., T/I , D), SAR , and adult age composition.

This model is appropriate for spring and summer Chinook salmon, Sockeye salmon, and Steelhead. Because the model does not accommodate the overwintering behavior of fall Chinook salmon, data from this species should not be analyzed with this model. The life-cycle model primarily uses detections of PIT tags at dams, but can also use detections at hatcheries or spawning sites, if available. The model uses branching migration paths to capture the age structure of returning adults. Figure 3.1 shows a schematic of the migration paths used by the life-cycle model. Survival (S) and detection probabilities (p) can be estimated at each detection site; joint survival and detection (λ) can be estimated at the last site. Smolts in the release group move downstream along a single path, but branch into one of several paths in the ocean based on maturation rates and transport history. Each branch represents a particular adult age class; these are kept separate throughout the adult spawning migration. While Figure 3.1 shows only three age classes, ROSTER can accommodate up to five age classes.

In addition to multiple adult age classes, the life-cycle model also incorporates transportation and is able to estimate transportation effects directly. Adult detections of transported fish are used to compare the return probabilities of transported and non-transported fish. Figure 3.2 shows a schematic of the migration paths of the transported and inriver fish; each ocean and adult path represents multiple return years. The model also accounts for censoring of juveniles and adults due to known removals at dams.

Table 3.1 lists the parameters used in the release-recapture model and

¹Available online at <http://www.cbr.washington.edu/paramest/roster/>.

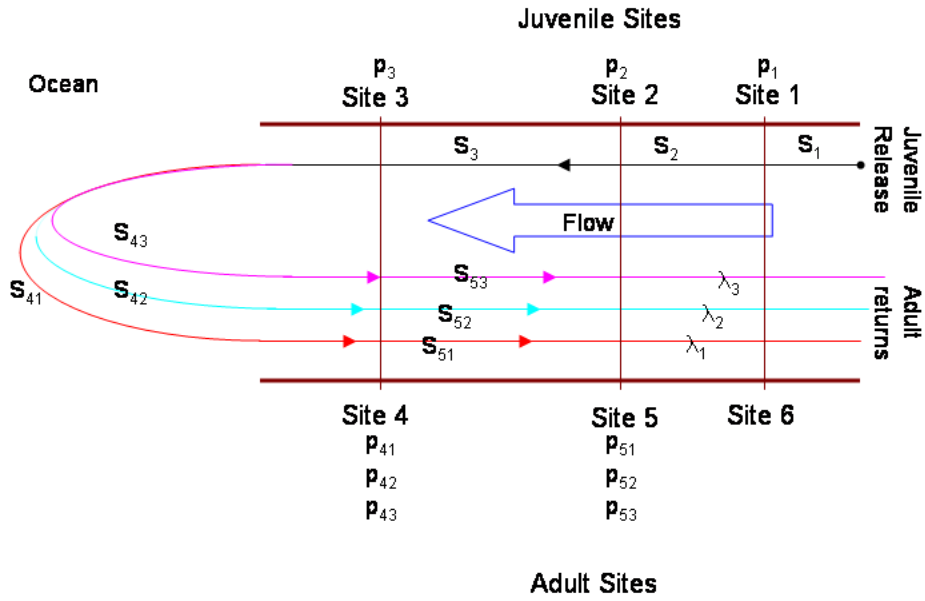


Figure 3.1: Schematic of migration paths for model with 3 juvenile sites, 3 adult sites, and 3 adult age classes.

estimated by ROSTER. The parameter estimates are maximum likelihood estimates, with the desirable properties of asymptotic unbiasedness, efficiency, and normality. All combinations of the basic model parameters (e.g., the performance measures in Section 3.4) are also maximum likelihood estimators.

3.2 Extent of Transportation Effects

It is possible that effects of juvenile transportation extend through part or all of the adult migration. For example, juvenile transportation may affect adult straying rates; this would result in different adult inriver survival probabilities for transported and non-transported groups. ROSTER has the ability to fit a model that estimates the effects of juvenile transportation on upriver adult survival, detection, and censoring probabilities. If transportation effects really end at the first adult site, however, the extra complexity in the model that estimates upriver effects will detract from the information that can be learned from the data set. Thus, the user can tell ROSTER which adult site should represent the extent of the effects of juvenile transportation, and using model selection tools (e.g., Likelihood Ratio Tests or AIC), determine the model with the appropriate extent of transportation effects.

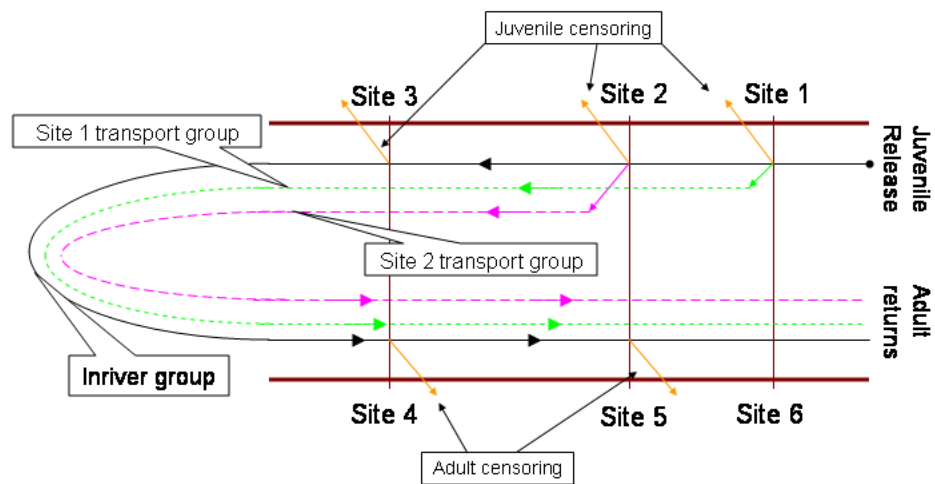


Figure 3.2: Schematic of migration paths of transported and inriver fish for model with 3 juvenile sites, 3 adult sites, and transportation at the first 2 juvenile sites. Each ocean and adult migration path represents multiple return years.

Table 3.1: Model parameters estimated by ROSTER. The number of juvenile detection sites is v , the number of adult detection sites is u , and the number of adult age classes is w .

Parameter	Definition
S_1	Probability of survival from release point to first detection site
S_i	Conditional probability of survival from detection site $i - 1$ to detection site i ($i = 2, \dots, v$) for non-transported fish
$S_{v+1,jC}$	Conditional joint probability of surviving from site v (last juvenile site) to site $v + 1$ (first adult site) and returning to site $v + 1$ in adult age class j ($j = 1, \dots, w$) for non-transported fish
S_{ijC}	Conditional probability of surviving from adult site $i - 1$ ($i = v + 2, \dots, v + u - 1$) to site i for non-transported fish in adult age class j ($j = 1, \dots, w$)
S_{ijT_l}	Conditional probability of surviving from adult site $i - 1$ ($i = v + 2, \dots, v + u - 1$) to site i for fish transported at site l ($l = 1, \dots, v$) as juveniles and returning in adult age class j ($j = 1, \dots, w$)
p_i	Conditional probability of detection at juvenile site i ($i = 1, \dots, v$), given survival to site i inriver (i.e., non-transported)
p_{ijC}	Conditional probability of detection at adult site i ($i = v + 1, \dots, v + u - 1$), given survival to site i , for non-transported fish in age class j ($j = 1, \dots, w$)
p_{ijT_l}	Conditional probability of detection at adult site i ($i = v + 1, \dots, v + u - 1$), given survival to site i , for fish transported from site l ($l = 1, \dots, v$) and returning in age class j ($j = 1, \dots, w$)
λ_{jC}	Conditional probability of non-transported fish in age class j ($j = 1, \dots, w$) surviving to and being detected at the final adult site (site $v + u$), given survival to the previous adult site (site $v + u - 1$)
λ_{jT_l}	Conditional probability of site- l transport fish ($l = 1, \dots, v$) in age class j ($j = 1, \dots, w$) surviving to and being detected at the final adult site (site $v + u$), given survival to the previous adult site (site $v + u - 1$)
c_i	Conditional probability of being removed (censored) at juvenile site i ($i = 1, \dots, v$), given detection at site i
c_{ij}	Conditional probability of being removed (censored) at adult site i ($i = v + 1, \dots, v + u - 1$) for age class- j fish ($j = 1, \dots, w$), given detection at site i
t_i	Conditional probability of being transported from site i ($i = 1, \dots, v$), given detection at that site, and no censoring
R_{ij}	Transport-inriver ratio (T/I) for fish transported from site i ($i = 1, \dots, v$) and returning to the first adult site in adult age class j ($j = 1, \dots, w$); defines survival of transported fish relative to survival of non-transported fish

The user tells ROSTER what the extent of the transportation effects should be by changing the value in the “Extent of Transportation Effects” window. The default value for the extent of transportation effects is the first adult site, because this relatively simple model is generally easier to fit than models that incorporate upriver transportation effects. Because juvenile and adult sites are numbered sequentially, the first adult site is one more than the number of juvenile sites. With 6 juvenile sites, the extent of transportation effects value will be 7 by default.

For example, for a data set with adult detections at Bonneville, McNary, Ice Harbor, and Lower Granite, ROSTER by default assumes that transportation effects end at Bonneville, and that adult survival, detection, and censoring probabilities upstream of Bonneville do not depend on the juvenile migration method (transportation or inriver). If the user wishes to investigate transportation effects on the adult migration up to McNary, then the value in the “Extent of Transportation Effects” window should be changed to the index of the second adult site (i.e., to 8 if there are 6 juvenile sites). If the user wishes to investigate transportation effects throughout the entire adult migration, then the extent of transportation effects value should be changed to the last adult site (Lower Granite, site 10 if there are 6 juvenile sites). After changing the value in the “Extent of Transportation Effects” window, the default model must be loaded in order for ROSTER to use the correct model.

The model parameters for adult survival, detection, and censoring vary between transport and non-transport groups up to, but not including, the site that is identified as the extent of transportation effects. Thus, if transportation effects are measured to the first adult site (Bonneville), then only ocean survival varies between transport and non-transport groups. If transportation effects are measured to the second adult site (e.g., McNary), then detection and censoring probabilities at Bonneville and survival between Bonneville and McNary vary with transport group. Upstream of the adult site identified as the extent of transportation effects, inriver survival and detection are assumed to be the same regardless of juvenile migration method (transportation or inriver). By default, ROSTER does not differentiate among transport groups in estimating upriver adult parameters. However, the expert user may modify the model definition file in order to compare, for example, upriver adult survival parameters for the different transport groups.

3.3 Tagged Versus Untagged Fish

Estimates of the model parameters and the performance measures are directly applicable only to the tagged fish in the release group. However, it is typical to assume that the tagged release group is representative of the untagged fish from the same populations that are migrating at the same time as the release

group. If the fish selected for tagging are assumed to be representative of the migrating group from the same population, then the estimated values of the model parameters and some of the performance measures may be applied to those untagged fish. (Note that the validity of such nonstatistical inference depends on the degree to which the tagged group represents the untagged population of interest, and is independent of the statistical model behind ROSTER.) Other performance measures must be modified for untagged fish. In particular, any performance measure that depends on the conditional transportation probabilities (i.e., t_i) must be modified for untagged fish, because tagged and untagged fish experience the transportation system differently. In general, all untagged fish that are bypassed at a transport dam are transported at that dam. However, some tagged fish are diverted from the bypass system back to the river rather than to the transport barges, for study purposes. This means that for tagged fish, the conditional transportation probability is usually less than 1 (i.e., $t_i < 1$), but for untagged fish, the conditional transportation probability is usually equal to 1. Letting t_i^U be the conditional transportation probability for tagged fish, we typically have $t_i^U = 1$. It is possible, however, that only a portion of untagged fish in the bypass system are transported, in which case t_i^U will be less than 1. For example, transportation may be operated only on alternate days, which would be consistent with $t_i^U = 0.5$.

The performance measures derived explicitly for untagged fish use the values t_i^U rather than the model parameters t_i , and are labeled with the superscript U to indicate “untagged”; otherwise, these performance measures are identical to those derived for tagged fish. The t_i^U values cannot be estimated from the tagging data, but must come from other sources. The default values used in ROSTER are $t_i^U = 1$ ($SE = 0$) for sites where tagged fish were transported, and $t_i^U = 0$ ($SE = 0$) for sites where no tagged fish were transported. The user may change both the t_i^U and the standard error values on the ROSTER interface; see Section 2.5.2 for more details. It should be noted that the direct statistical inference of performance measures using t_i^U is to fish in the (tagged) release group, had they been treated as untagged fish.

3.4 Performance Measures

The performance measures estimated by ROSTER are defined here, both heuristically and analytically in terms of the model parameters. In order to address the needs of as many users as possible, it has often been necessary to create several versions of the same type of performance measure. For example, consider a measure of transportation effects. It is possible to measure the effect of transportation from a single transport dam, either in the context of the entire transport system or else isolated from that system. Alternatively, the effects of the entire transport system may be measured, with all transport dams considered. The transport effect measure

may be either a relative measure (ratio) or an absolute measure (difference). The effect may be measured on return probabilities only to the adult site that has been identified in the model as the extent of the transport effects (E), or all the way to the final adult site ($v + u$). The measure may be calculated for tagged fish, or for untagged fish (that is, tagged fish treated as untagged fish), and finally, the measure may or may not include jacks. This gives us a multitude of different transport effect measures to choose from. In what follows, most of the different types of measures are shown explicitly. Each of the measures defined below that include multiple age classes also include jacks (age-1 adults, or ocean-1 fish). Not shown are separate measures that exclude the jacks. ROSTER reports measures both with and without jacks; the formulas for those that exclude jacks can be easily derived from the formulas that follow, by restricting the index j to $j > 1$.

In the following sections, v is the number of juvenile detection sites, u is the number of adult detection sites, and w is the number of adult age classes. The index E indicates the adult site where the model assumes the juvenile transport effects end (“Extent of Transportation Effects,” see Section 3.2). The default value of E is the first adult detection site, typically Bonneville. The index E follows the consecutive numbering format of juvenile and adult sites used by the model; thus, the possible values of E range from site $v + 1$ (first adult site) to site $v + u$ (last adult site). Some performance measures involving adult return probabilities are indexed by k . The index k is used to indicate the adult site to which returns are reported, with $k = E$ or $k = v + u$.

3.4.1 Transportation Effects

Several types of transportation effect measures are estimated by ROSTER, including both site-specific (equivalently, dam-specific) and systemwide measures. The site-specific measures assess the effects of transportation from a single dam, while the systemwide measures assess the effects of the entire transportation system. Buchanan et al. (2006)² discuss measures of transportation effects in more detail.

3.4.1.1 Site-Specific Transportation Effects Measures

All of the site-specific transportation effect measures compare the return probabilities of transported and non-transported (control) fish with a transport-inriver ratio, T/I . The different measures vary in the assumptions made about the control fish, and in whether they are applicable to tagged or untagged

²Buchanan, R.A., Skalski, J.R., and Smith, S.G. (2006), “Estimating the Effect of Transportation From Different Vantage Points and Management Perspectives,” *North American Journal of Fisheries Management*, 26, 460-472. Available online at <http://www.cbr.washington.edu/paramest/roster/>.

fish. Each measure is described and defined in terms of the model parameters below. In each case, the subscript i ($i = 1, \dots, v$) indicates the site (dam) being considered, and the subscript j ($j = 1, \dots, w$) indicates the adult age class being used.

Isolated Age-Specific T/I ($R_{ij(k)}$)

The isolated age-specific T/I ($R_{ij(k)}$) for juvenile site (or dam) i and adult age class j is the ratio of the age- j return probability from site i to adult site k for site- i transport fish to that of fish passing site i inriver, as if there were no further downstream transportation sites. The measure $R_{ij(k)}$ estimates the relative age- j transportation effect at site i isolated from the rest of the transportation system. The control fish for $R_{ij(k)}$ are treated as if they had no possibility of being transported at sites downriver of site i . The measure $R_{ij(k)}$ is valid for both tagged and untagged fish.

If $k = v + 1$ or $E = v + 1$ (i.e., if “Extent of Transportation Effects” = $v + 1$), then $R_{ij(k)} = R_{ij}$, the model parameter representing site- i transportation effects on the return probability to the first adult site for age- j fish. If both $k > v + 1$ and $E > v + 1$, then

$$R_{ij(k)} = R_{ij} \frac{S_{v+2,jT_i} \cdots S_{EjT_i}}{S_{v+2,jC} \cdots S_{EjC}}. \quad (3.1)$$

Note that $R_{ij(k)} = R_{ij(E)}$ for all $k > E$.

Isolated T/I ($R_{i(k)}$)

The isolated T/I ($R_{i(k)}$) for juvenile site i is the ratio of the return probability from site i to adult site k for site- i transport fish to that of fish passing site i inriver, as if there were no further downstream transportation sites. The measure $R_{i(k)}$ estimates the relative transportation effect at site i isolated from the rest of the transportation system. The control fish for $R_{i(k)}$ are treated as if they had no possibility of being transported at sites downriver of site i . The measure $R_{i(k)}$ is valid for both tagged and untagged fish. This measure is analogous to the R_{abcd} measure of Sandford and Smith (2002)³.

$$R_{i(k)} = \frac{\sum_{j=1}^w S_{v+1,jC} \cdots S_{kjC} R_{ij(k)}}{\sum_{j=1}^w S_{v+1,jC} \cdots S_{kjC}}, \quad (3.2)$$

where w is the number of adult age classes. Note that $R_{i(v+1)} = \frac{\sum_{j=1}^w S_{v+1,jC} R_{ij}}{\sum_{j=1}^w S_{v+1,jC}}$.

³Sandford, B.P., and Smith, S.G. (2002), “Estimation of Smolt-to-Adult Return Percentages for Snake River Basin Anadromous Salmonids, 1990-1997,” *Journal of Agricultural, Biological, and Environmental Statistics*, 7, 243-263.

Contextual T/I for Tagged Fish ($RC_{i(k)}$)

The contextual T/I for tagged fish ($RC_{i(k)}$) for juvenile site i is the ratio of the return probability to adult site k of site- i transport fish to that of fish passing site i inriver. The measure $RC_{i(k)}$ assesses transportation from site i in the context of the entire downriver transportation system. The control fish for $RC_{i(k)}$ may be transported at sites downstream of site i . Because $RC_{i(k)}$ depends on the proportions of the control fish that are transported at downstream dams, it is valid only for tagged fish. This measure is analogous to the relative recovery measure historically used to measure transportation effects (Ebel et al. 1973⁴).

$$RC_{i(k)} = \frac{R_{i(k)}}{\sum_{s=i+1}^v \left\{ p_s t_s R_{s(k)} \left[\prod_{x=i+1}^{s-1} (1 - p_x t_x) \right] \right\} + \prod_{s=i+1}^v (1 - p_s t_s)}. \quad (3.3)$$

Contextual T/I for Untagged Fish ($RC_{i(k)}^U$)

The contextual T/I for untagged fish ($RC_{i(k)}^U$) for juvenile site i has the same meaning as $RC_{i(k)}$, but is applicable to untagged fish (or, to tagged fish if they had been treated as untagged). That is, under the assumption that tagged and untagged fish respond identically to transportation, then $RC_{i(k)}^U$ is the ratio of the return probability to adult site k from site i of site- i transport fish to that of (untagged) fish passing site i inriver. The control fish for $RC_{i(k)}^U$ may be transported at sites downriver of site i . The rates at which the untagged control fish are transported at these sites are represented by the t_x^U values, for all juvenile dams $x > i$.

$$RC_{i(k)}^U = \frac{R_{i(k)}}{\sum_{s=i+1}^v \left\{ p_s t_s^U R_{s(k)} \left[\prod_{x=i+1}^{s-1} (1 - p_x t_x^U) \right] \right\} + \prod_{s=i+1}^v (1 - p_s t_s^U)}. \quad (3.4)$$

3.4.1.2 Systemwide Transportation Effects Measures

Systemwide transportation effect measures assess the effect of the complete transportation system on return probabilities (to adult site $k = E$ or $k = v + u$) of the entire release group. Rather than comparing return probabilities of transported fish to those of non-transported fish, systemwide measures compare the observed return probability of the release group to what the return probability would have been if there had been no transportation. Thus, systemwide measures are able to

⁴Ebel, W.J., Park, D.L., and Johnsen, R.C. (1973), "Effects of transportation on survival and homing of Snake River chinook salmon and steelhead trout," *U.S. National Marine Fisheries Service Fishery Bulletin*, 71, 549-563.

compare return probabilities under two strategies: transportation versus complete inriver migration. This comparison can be either a relative value ($R_{sys(k)}$ and $R_{sys(k)}^U$) or an absolute value ($R_{abs(k)}$ and $R_{abs(k)}^U$).

Systemwide T/I for Tagged Fish ($R_{sys(k)}$)

The systemwide T/I for tagged smolts ($R_{sys(k)}$) is the ratio of the return probability from the release site to adult site k of all smolts, transported or not, under the existing transportation system to what the return probability would have been without the transportation system, i.e., under complete inriver migration. The measure $R_{sys(k)}$ incorporates the proportion of the smolts reaching each dam that are transported. It combines the effects of transportation from all possible transport sites. It should be noted that $R_{sys(k)}$ does not compare the return probability of transported smolts to the return probability of non-transported smolts, but rather compares the return probability of the entire release group (transported and inriver) to what that return probability would have been if there had been no transportation system. Because $R_{sys(k)}$ depends on the proportion of smolts transported at each dam, it is valid only for tagged smolts.

$$\begin{aligned} R_{sys(k)} &= \frac{Pr[\text{Adult return to } k \mid \text{Transportation system}]}{Pr[\text{Adult return to } k \mid \text{No transportation system}]} \\ &= \sum_{i=1}^v \left\{ p_i t_i R_{i(k)} \prod_{x=1}^{i-1} (1 - p_x t_x) \right\} + \prod_{i=1}^v (1 - p_i t_i). \end{aligned} \quad (3.5)$$

Systemwide T/I for Untagged Fish ($R_{sys(k)}^U$)

The systemwide T/I for untagged smolts ($R_{sys(k)}^U$) is analogous to the value for tagged smolts ($R_{sys(k)}$), but is applicable to untagged smolts of the same stock that migrated at the same time as the tagged group, or to tagged smolts if they had been treated as untagged smolts. The measure $R_{sys(k)}^U$ is a relative measure comparing the return probability from the release site to adult site k of all smolts (of the same stock and timing), transported or not, under the existing transportation system to that of a non-transportation system, i.e., complete inriver migration. The measure $R_{sys(k)}^U$ depends on the proportions of untagged smolts transported at each dam, i.e., t_i^U for juvenile site i .

$$R_{sys(k)}^U = \sum_{i=1}^v \left\{ p_i t_i^U R_{i(k)} \prod_{x=1}^{i-1} (1 - p_x t_x^U) \right\} + \prod_{i=1}^v (1 - p_i t_i^U). \quad (3.6)$$

Absolute Systemwide Transportation Effect Measure for Tagged Fish ($R_{abs(k)}$)

The absolute measure ($R_{abs(k)}$) is the difference between the return probability from the release point to adult site k for the entire release group under the transportation system and what that return probability would have been without the transportation system. This measure is similar to $R_{sys(k)}$ in that it compares the observed return probability to the hypothetical return probability if there had been no transportation. Like $R_{sys(k)}$, $R_{abs(k)}$ is suitable only for tagged fish. Unlike $R_{sys(k)}$, $R_{abs(k)}$ reflects the absolute change in return probabilities due to transportation.

$$R_{abs(k)} = \left(S_1 \cdots S_v \sum_{j=1}^w S_{v+1,jC} \cdots S_{kjC} \right) (R_{sys(k)} - 1). \quad (3.7)$$

Absolute Systemwide Transportation Effect Measure for Untagged Fish ($R_{abs(k)}^U$)

The absolute measure ($R_{abs(k)}^U$) is analogous to $R_{abs(k)}$, but is applicable to untagged smolts of the same stock that migrated at the same time as the tagged group, or to tagged smolts if they had been treated as untagged smolts. The measure $R_{abs(k)}^U$ reflects the absolute change in return probabilities (to adult site k) due to the transportation system. This measure is similar to $R_{sys(k)}^U$ in that it compares the observed return probability to the hypothetical return probability if there had been no transportation. Like $R_{sys(k)}^U$, $R_{abs(k)}^U$ is suitable for untagged fish and uses transportation rates of untagged fish (i.e., t_i^U for juvenile site i).

$$R_{abs(k)}^U = \left(S_1 \cdots S_v \sum_{j=1}^w S_{v+1,jC} \cdots S_{kjC} \right) (R_{sys(k)}^U - 1). \quad (3.8)$$

3.4.2 Differential Mortality

Differential mortality, sometimes referred to as delayed mortality or delayed differential mortality, is the ratio of the post-Bonneville survival of transported fish to that of non-transported fish. In particular, the measures reported by ROSTER give the ratio of post-Bonneville survival to both site E (the adult site that defines the extent of the transportation effects) and to the final adult site, for transported and non-transported fish.

Differential mortality measures depend on the survival of transported smolts during transportation (i.e., on the barge or in the truck). The default barge transport survival is 98%, but the user can change this value and enter a variance

for barge survival on the Estimation tab of the program interface. See Section 2.5.2 for details.

Just as there are many different types of transportation effect measures depending on the focus (site-specific or systemwide), nature of the control group (downstream transportation possible or not), and applicable population (tagged fish or untagged fish), there are the same types of measures of differential mortality. ROSTER provides estimates of five types of differential mortality measures: two systemwide measures (for tagged and untagged fish), and three site-specific measures (one isolated measure, one contextual measure for tagged fish, and one contextual measure for untagged fish). For each type, the differential mortality measure is reported for returns to site $k = E$ and to the last adult site ($k = v + u$).

3.4.2.1 Site-Specific Differential Mortality Measures

Like the differences among the site-specific transportation effect measures ($R_{i(k)}$, $RC_{i(k)}$, and $RC_{i(k)}^U$), the differences among the three site-specific differential mortality measures depend on the nature of the control group and on whether they are applicable to tagged or untagged fish. The three measures are described and defined in terms of the model parameters below. In each case, the subscript i ($i = 1, \dots, v$) indicates the transport site (dam) being considered, and the subscript j ($j = 1, \dots, w$) indicates the adult age class being used. The subscript E ($E = v + 1, \dots, v + u$) indicates the adult site to which transportation effects have been estimated (default is the first adult site), and k indicates the adult site to which return probabilities have been calculated, with $k = E$ or $k = v + u$.

Isolated Differential Mortality ($D_{i(k)}$)

The isolated differential mortality measure ($D_{i(k)}$) for juvenile site i is the ratio of the return probability from Bonneville as a juvenile to adult site k for site- i transport fish relative to that of fish passing site i inriver, as if there were no further downstream transportation sites. Like $R_{i(k)}$, the control fish for $D_{i(k)}$ are treated as if they had no possibility of being transported at sites downriver of site i . The measure $D_{i(k)}$ is valid for tagged and untagged fish:

$$D_{i(k)} = \frac{R_{i(k)} \prod_{y=i+1}^v S_y}{S_B}, \quad (3.9)$$

where S_B is the survival probability on the barge during transportation.

Contextual Differential Mortality for Tagged Fish ($DC_{i(k)}$)

Contextual site-specific differential mortality for tagged fish ($DC_{i(k)}$) for juvenile site i is the ratio of the return probability from Bonneville as a juvenile to

adult site k for site- i transport fish relative to fish passing site i inriver. Like $RC_{i(k)}$, the control fish for $DC_{i(k)}$ may be transported at sites downstream of site i . Also like $RC_{i(k)}$, $DC_{i(k)}$ is valid only for tagged fish because it depends on downstream transportation probabilities:

$$DC_{i(k)} = \frac{RC_{i(k)}}{S_B} \left[S_B \sum_{y=i+1}^v \left\{ S_y p_y t_y \prod_{l=i+1}^{y-1} S_l (1 - p_l t_l) \right\} + \prod_{l=i+1}^v S_l (1 - p_l t_l) \right], \quad (3.10)$$

where S_B is the survival probability on the barge during transportation.

Contextual Differential Mortality for Untagged Fish ($DC_{i(k)}^U$)

Contextual differential mortality for untagged fish ($DC_{i(k)}^U$) for juvenile site i is analogous to $DC_{i(k)}$, but is applicable to untagged fish (or to tagged fish if they had been treated as untagged). That is, under the assumption that tagged and untagged fish respond identically to transportation and have identical survival, then $DC_{i(k)}^U$ is the ratio of the return probability from Bonneville (as a juvenile) to adult site k for transport site- i fish to fish that passed site i inriver. The control fish for $DC_{i(k)}^U$ may be transported at dams downriver of site i . The rates at which the untagged control fish are transported at those sites are represented by the t_x^U values, for all juvenile dams $x > i$, and may be entered by the user on the Estimation tab of the program interface, or else be left at the default values of either 1 or 0.

$$DC_{i(k)}^U = \frac{RC_{i(k)}^U}{S_B} \left[S_B \sum_{y=i+1}^v \left\{ S_y p_y t_y^U \prod_{l=i+1}^{y-1} S_l (1 - p_l t_l^U) \right\} + \prod_{l=i+1}^v S_l (1 - p_l t_l^U) \right], \quad (3.11)$$

where S_B is the survival probability on the barge during transportation.

3.4.2.2 Systemwide Differential Mortality Measures

Systemwide measures combine effects of transportation at all the transport dams in the system, comparing ocean return probabilities (to the extent of transportation effects) for all fish that were transported to all fish that were not transported. There is one measure for tagged fish and another measure for untagged fish (or for tagged fish had they been treated as untagged fish).

Systemwide Differential Mortality for Tagged Fish ($D_{sys(k)}$)

$$D_{sys(k)} = \frac{\left(\prod_{i=1}^v S_i \right) \left(\sum_{i=1}^v \left[\left(\prod_{y=1}^{i-1} (1 - p_y t_y) \right) p_i t_i R_{i(k)} \right] \right)}{S_B \sum_{i=1}^v \left(\left[\prod_{y=1}^{i-1} S_y (1 - p_y t_y) \right] S_i p_i t_i \right)}, \quad (3.12)$$

where S_B is survival on the barge during the transport.

Systemwide Differential Mortality for Untagged Fish ($D_{sys(k)}^U$)

$$D_{sys(k)}^U = \frac{\left(\prod_{i=1}^v S_i \right) \left(\sum_{i=1}^v \left[\left(\prod_{y=1}^{i-1} (1 - p_y t_y^U) \right) p_i t_i^U R_{i(k)} \right] \right)}{S_B \sum_{i=1}^v \left(\left[\prod_{y=1}^{i-1} S_y (1 - p_y t_y^U) \right] S_i p_i t_i^U \right)}, \quad (3.13)$$

where S_B is the survival probability on the barge during transportation.

3.4.3 Juvenile Survival

Juvenile inriver survival, S_J , is the probability of surviving from the first detection site (typically Lower Granite Dam) to the last juvenile detection site (typically Bonneville Dam) for non-transported fish.

$$S_J = \prod_{i=2}^v S_i. \quad (3.14)$$

3.4.4 Ocean Return Probabilities

Ocean Return Probability for Non-transported Fish (O_{NT})

The measure O_{NT} is the probability of returning to the first adult detection site, conditional on reaching the last juvenile detection site as an inriver (non-transported) fish. The measure O_{NT} is valid for both tagged and untagged fish.

$$O_{NT} = \sum_{j=1}^w S_{v+1,jC}. \quad (3.15)$$

Ocean Return Probability for Dam- i Transported Fish (O_{T_i})

The measure O_{T_i} is the probability of returning from juvenile site (or dam) i to the first adult detection site, conditional on being transported at juvenile site i . The measure O_{T_i} represents the combination of barge survival, estuary and ocean survival, and maturation. This measure is valid for both tagged and untagged fish.

$$O_{T_i} = S_{i+1} \cdots S_v R_{i(v+1)} \sum_{j=1}^w S_{v+1,jC}. \quad (3.16)$$

Smolt-to-Adult Ratio for Tagged Fish (SAR)

The measure SAR is the probability of returning from the first juvenile detection site to the final adult site. This measure incorporates the proportion of smolts transported and transportation effects, as well as adult age structure. It is valid only for tagged fish.

$$SAR = S_2 \cdots S_v R_{sys(v+u)} \sum_{j=1}^w [S_{v+1,jC} \cdots S_{v+u,jC}], \quad (3.17)$$

where $S_{v+u,jC} = \lambda_{jC}$ under the assumption of 100% detection at the final detection site.

Smolt-to-Adult Ratio for Untagged Fish (SAR^U)

The measure SAR^U is the probability of an untagged fish returning from the first juvenile detection site (for tagged smolts) to the final adult site. As with other measures applicable to untagged fish, this value is appropriate only for untagged fish that are represented by the tagged group. This measure depends on the transportation proportions of untagged smolts at each transport dam, i.e., t_i^U for juvenile site i .

$$SAR^U = S_2 \cdots S_v R_{sys(v+u)}^U \sum_{j=1}^w [S_{v+1,jC} \cdots S_{v+u,jC}], \quad (3.18)$$

where $S_{v+u,jC} = \lambda_{jC}$ under the assumption of 100% detection at the final detection site.

3.4.5 Adult Return Age Distribution

Adult Return Age Distribution for Tagged Fish ($A_{j(k)}$)

The measures $A_{j(k)}$ indicate the proportions of adults returning to site k (where $k = E$ or $k = v + u$) in each of the adult age classes (indexed by j , $j = 1, \dots, w$). These values incorporate transportation proportions and effects, and so are suitable only for the tagged fish in the release group. For $k = v + u$, $S_{kjC} = \lambda_{jC}$ under the assumption of 100% detection at the final adult detection site.

$$A_{j(k)} = \frac{S_{v+1,jC} \cdots S_{kjC}}{\sum_{m=1}^w S_{v+1,mC} \cdots S_{kmC}} \frac{\left[\sum_{i=1}^v \left(p_i t_i R_{ij(k)} \prod_{y=1}^{i-1} [1 - p_y t_y] \right) + \prod_{i=1}^v (1 - p_i t_i) \right]}{R_{sys(k)}}. \quad (3.19)$$

Adult Return Age Distribution for Untagged Fish ($A_{j(k)}^U$)

The measures $A_{j(k)}^U$ are analogous to the $A_{j(k)}$ measures, but for untagged smolts, representing the proportions of untagged adults (represented by the release group) returning to site k (where $k = E$ or $k = v + u$) in each of the adult age classes (indexed by j , $j = 1, \dots, w$). These measures depend on the proportions of untagged smolts entering the bypass systems at transport dams that are transported, i.e., t_i^U for juvenile site i . As with other measures for untagged fish, the direct inference for these values is to the tagged population, had they been treated as untagged. For $k = v + u$, $S_{kjC} = \lambda_{jC}$ under the assumption of 100% detection at the final adult detection site.

$$A_{j(k)}^U = \frac{S_{v+1,jC} \cdots S_{kjC}}{\sum_{m=1}^w S_{v+1,mC} \cdots S_{kmC}} \frac{\left[\sum_{i=1}^v \left(p_i t_i^U R_{ij(k)} \prod_{y=1}^{i-1} [1 - p_y t_y^U] \right) + \prod_{i=1}^v (1 - p_i t_i^U) \right]}{R_{sys(k)}^U}. \quad (3.20)$$

3.4.6 Adult Upriver Survival

Adult Upriver Survival for Tagged Fish (S_A)

The measure S_A is the survival from Bonneville (or the first adult site) to Lower Granite (or the last adult site) of all adults from the initial release group. This measure takes into account the proportions of smolts transported and the effects of transportation, as well as the adult return age distribution, but does not adjust for harvest or straying. It is suitable only for (tagged) fish in the release group.

$$S_A = \frac{\sum_{j=1}^w [S_{v+1,jC} \cdots S_{v+u,jC}] R_{sys(v+u)}}{\sum_{m=1}^w S_{v+1,mC} R_{sys(v+1)}}, \quad (3.21)$$

where $S_{v+u,jC} = \lambda_{jC}$ under the assumption of 100% detection at the final detection site.

Adult Upriver Survival for Untagged Fish (S_A^U)

The measure S_A^U is the survival from Bonneville (or the first adult site) to Lower Granite (or the last adult site) of untagged adults that are represented by the tagged release group. This measure takes into account the proportions of bypassed, untagged smolts transported (value t_i^U for juvenile site i), as well as the effects of transportation and the adult return age distribution, but does not adjust for harvest or straying.

$$S_A^U = \frac{\sum_{j=1}^w [S_{v+1,jC} \cdots S_{v+u,jC}] R_{sys(v+u)}^U}{\sum_{m=1}^w S_{v+1,mC} R_{sys(v+1)}^U}, \quad (3.22)$$

where $S_{v+u,jC} = \lambda_{jC}$ under the assumption of 100% detection at the final detection site.

Age-Specific Adult Upriver Survival for Tagged Fish (S_{Aj})

The measure S_{Aj} is the survival from Bonneville (or the first adult site) to Lower Granite (or the last adult site) of all age- j adults from the initial release group, including transported and non-transported fish. This measure takes into account the proportions of smolts transported and the effects of transportation, but does not adjust for harvest or straying. It is suitable only for (tagged) fish in the release group.

$$S_{Aj} = \frac{(\prod_{i=1}^v [1 - p_i t_i]) (\prod_{i=v+2}^{v+u} S_{ijC}) + \sum_{i=1}^v \left[p_i t_i R_{ij} \left(\prod_{l=1}^{i-1} [1 - p_l t_l] \right) (\prod_{l=v+2}^{v+u} S_{ljT_i}) \right]}{\prod_{i=1}^v [1 - p_i t_i] + \sum_{i=1}^v \left[p_i t_i R_{ij} \left(\prod_{l=1}^{i-1} [1 - p_l t_l] \right) \right]}, \quad (3.23)$$

where $S_{v+u,jC} = \lambda_{jC}$ and $S_{v+u,jT_i} = \lambda_{jT_i}$ under the assumption of 100% detection at the final detection site.

Age-Specific Adult Upriver Survival for Untagged Fish (S_{Aj}^U)

The measure S_{Aj}^U is the survival from Bonneville (or the first adult site) to Lower Granite (or the last adult site) of all untagged age- j adults represented by the initial release group, including transported and non-transported fish. This measure takes into account the proportions of smolts transported (value t_i^U for juvenile site i) and the effects of transportation, but does not adjust for harvest or straying.

$$S_{Aj}^U = \frac{(\prod_{i=1}^v [1 - p_i t_i^U]) (\prod_{i=v+2}^{v+u} S_{ijC}) + \sum_{i=1}^v \left[p_i t_i^U R_{ij} \left(\prod_{l=1}^{i-1} [1 - p_l t_l^U] \right) (\prod_{l=v+2}^{v+u} S_{ljT_i}) \right]}{\prod_{i=1}^v [1 - p_i t_i^U] + \sum_{i=1}^v \left[p_i t_i^U R_{ij} \left(\prod_{l=1}^{i-1} [1 - p_l t_l^U] \right) \right]}, \quad (3.24)$$

where $S_{v+u,jC} = \lambda_{jC}$ and $S_{v+u,jT_i} = \lambda_{jT_i}$ under the assumption of 100% detection at the final detection site.

4.0 Fitting ROSTER and Model Selection

4.1 Guidance on Fitting ROSTER

In most cases, the default model provided by ROSTER should be sufficient to analyze the data. In some cases, however, the user may need to modify the model file in order to properly fit the model to the data. Below are some suggestions on using ROSTER. These steps require familiarity with the format of the model file; see Chapter 7 for an example of a model file.

1. Begin by loading the file with all the age classes observed.
2. Try running the default model with the value in the “Extent of Transportation Effects” window set to the first adult site (e.g, 7 if there are 6 juvenile sites).
3. If the model in step 2 does not fit, then it is necessary to edit the model file. Save the model file as either a XML or a text document (under the File menu), then open and edit the model file. The necessary edits will depend on the data set. Some common changes are described here.
 - (a) Examine the Parameters report, which lists the default initial values that ROSTER uses (the MoM estimates). If any of the MoM estimates *other* than for the T/I parameters (R_{ij}) are equal to 1, 0, or “inf” (∞), then the default model file or the data file will need to be modified as follows:
 - i. Detection parameters (p_{ij} , p_{ijC} , or p_{ijT_k}) and “last reach” parameters (λ_{jC} or λ_{ijT_k}) that are very high (e.g., 1 or above) cannot be easily estimated, and so must be fixed in the model file. Rather than fixing them to 1, they should be fixed to a value very close to 1, such as 0.999999.
 - ii. Survival parameters (S_{ij} , S_{ijC} , or S_{ijT_k}) that are equal to 1 or very close to 1 may need to be fixed to 1. In general, ROSTER can estimate these parameters, but the particular structure of the data may make this impossible. Try fitting the model both with and without these very high survival parameters fixed to 1 or 0.999999. If possible, it is preferable to leave these parameters free (not fixed), but in some cases that option may not work.
 - iii. If the MoM estimate for a detection parameter is equal to 0, that means that this parameter cannot be estimated for this data set; it is not necessarily the case that there was no detection at the particular site, but may be the case that no fish detected at the site were also

detected at a later site. Rather than modifying the model file, it is necessary to remove the field corresponding to that detection site from the data file. To do this, it is necessary to run the raw PTAGIS data files through PitPro again, omitting the problematic detection site. If the problematic detection site is the first or last juvenile or adult site, then the meaning of the performance measures S_J (juvenile survival) or S_A (adult survival), respectively, as well as SAR , will be altered.

- iv. If the MoM estimate for a survival parameter is either 0 or “inf”, it is likely that the detection parameter for an adjacent site cannot be estimated. Check if a detection parameter has 0 for its MoM estimate; if so, remove the detection site from the data file by running the data through PitPro again, reload the modified capture history file into ROSTER, and try fitting the default model again.
 - v. If the MoM estimate for a parameter is “nan”, then it may be necessary to specify a numeric initial value (seed) for the parameter in the model file; see Section 5.4 for examples.
- (b) If the model still does not fit following the edits described here, then examine the Data Summary report and note the sizes of the observed age classes. Age classes with very few detections will be difficult to analyze, and may need to be removed. This can be done by:
- i. Censoring the capture histories of fish observed in these age classes, with the censoring occurring at the final juvenile site at the which the fish was detected.
 - ii. Alternatively, the detections in the problematic age class can be recoded to fall in an adjacent age class (i.e., two age classes are collapsed to one). This option is not ideal because age classes cannot properly be collapsed unless there is no difference in adult upstream parameters between the two age classes. However, because adult detection and survival estimates are often very high, it may be logical to collapse two adult age classes.
 - iii. Remove from the data file the capture histories of the fish detected in the problematic age class. This option is reasonable only if there are very few fish to be removed (e.g., 1 or 2), because it may bias results.
4. If the model in one of the above steps fits, then it may be worthwhile to explore a more flexible model structure.
- (a) Try using different detection or censoring probabilities for all transport groups (including non-transported or control fish) for the first adult site. This can be done by deleting the section of the model file that equates these parameters. Note that if any of these parameters were fixed to a

particular value (e.g., 0 or 0.999999), then this must be adjusted if the parameters are allowed to differ.

- (b) If the model in step 4a fits, then try increasing the site identified as the extent of transportation effects. This is most easily done by selecting the default model to be used, increasing the number in the “Extent of Transportation Effects” window, loading the model, and then saving and editing the model as described above, if necessary. Alternatively, the expert user may alter the edited model file so that transport and non-transport groups have separate adult parameters up to the adult site that is identified as the extent of transportation effects. It may also be interesting to distinguish among groups transported at different sites.

It is generally not necessary for the user to pick initial values (seeds) for ROSTER, because ROSTER calculates the Method-of-Moments estimates to use as seeds. However, the user may select a different seed if desired, and adjust the model file accordingly. An example of how to do this is shown at the end of Chapter 5.

4.2 Model Selection

In most cases, models with different structures will fit the data. For example, some models may equate certain parameters that are left unique in other models. In general, the “best” model captures all valid differences among parameters without unnecessarily estimating separate values for parameters that are effectively equal. There are several tools that may be used to select the best model, including Likelihood Ratio Tests and Akaike’s Information Criterion (AIC). Additionally, common sense should be used in determining whether it is really necessary to estimate unique values for a set of parameters.

4.2.1 Likelihood Ratio Tests

Likelihood Ratio Tests (LRTs) are statistical tests used to compare two models in which one model is “nested” within the other. The model nested in the other may be considered a special case of the more general model. The more general (“full”) model typically has more parameters than the special case (“reduced model”), or estimates separate values for parameters that are equated in the reduced model. The null hypothesis for an LRT is that the two models are equivalent, in which case the reduced model should be used.

Program ROSTER provides the information necessary to perform an LRT between two nested models. If Model 1 is nested within Model 2, then the user fits

both Model 1 and Model 2 in ROSTER, and notes the final function value for each model. Let FV_1 be the function value for Model 1 (the reduced model), and FV_2 be the function value for Model 2 (the full model). Because Model 2 is more general than Model 1, $FV_2 \leq FV_1$. The LRT test statistic is

$$X = 2(FV_1 - FV_2). \quad (4.1)$$

Define d to be the difference in the number of parameters between Models 1 and 2. For a Type I error of size α , compare X to $\chi^2_{d,1-\alpha}$, the upper $100 \times (1 - \alpha)$ percentile point from the chi-square distribution with d degrees of freedom. If $X > \chi^2_{d,1-\alpha}$, then reject the null hypothesis, and use the full model (Model 2). Otherwise, use the reduced model (Model 1).

For example, the sample data set `data1.txt` (in the sample data folder that comes with ROSTER) has three transportation sites and two adult detection sites. We can estimate transportation effects for the three transportation sites, but where do those effects end? At the first adult site (Bonneville), or at the second adult site (Lower Granite)? We can fit models with either assumption by setting the extent of transportation effects to either 7 or 8, and loading and running the default model. Let the model with transportation effects ending at site 7 be Model 1, and the model with transportation effects ending at site 8 be Model 2. The difference between these two models is that in Model 2, the adult detection and censoring parameters at site 7, and the “last reach” lambda parameters, are estimated separately for transported fish and non-transported fish (within each age class), while in Model 1, these parameters are equated for transported and non-transported groups. The model equating these parameters (Model 1) is the reduced model, while the model estimating unique parameters for transported fish (Model 2) is the full model. The full model yields a negative log-likelihood value of $FV_2 = 478537.60088236$ with 47 parameters, and the reduced model yields a negative log-likelihood value of $FV_1 = 478551.17171487$ with 40 parameters. This gives the LRT test statistic $X = 27.14167$ with $d = 7$ and critical value $\chi^2_{7,0.95} = 14.06714$ for $\alpha = 0.05$. Because $X > \chi^2_{7,0.95}$, we reject our null hypothesis that adult parameters are the same for transported and non-transported fish, and Model 2 should be used.

4.2.2 Akaike’s Information Criterion

Models that are not nested cannot be compared using LRTs. For example, consider comparing two models that are identical except in how they handle adult detection probabilities at the first adult site. Model 1 equates these detection probabilities across all transport groups and the non-transport group, while Model 2 specifies unique values for the transport and non-transport groups. So far, Model 1 is nested within Model 2, and model selection could be performed using an LRT. Suppose, however, that the detection probability for the first age class of the

non-transport group was 100%, but for the first age class of the transported fish, it was smaller, say 80%. In this case, Model 2 would need to fix the detection probability for the first age class of the non-transport group to 1, but leave the comparable parameter for the transport group free. Because the parameter for the non-transport group is fixed in Model 2 but not in Model 1, Model 1 is *not* nested in Model 2, and model selection cannot be done via LRTs. Instead, we can use Akaike's Information Criterion (AIC), which is based on likelihood values and the number of parameters in each model. The AIC is not a formal test, and should not be used to compare models with widely different likelihood structures, but is a reasonable tool for model selection in ROSTER. The AIC can be used to compare nested models as well, but in general it is better to use LRTs in such a case.

To compare models using AIC, calculate the AIC value for each model, and select the model with the smaller AIC. As with the LRT defined in Section 4.2.1, let FV_i be the ROSTER function value for Model i (here, $i = 1$ or 2). Also, let K_i be the number of parameters estimated in Model i . The number of parameters estimated is listed near the top of the Estimation Results report. Then the AIC for Model i is

$$AIC_i = 2FV_i + 2K_i. \quad (4.2)$$

As an example, consider the models in Section 4.2.1. Although these models are nested, it is still possible to compare them via AIC. Model 1 had a function value of $FV_1 = 478551.17171487$ and $K_1 = 40$ parameters, with an AIC of $AIC_1 = 957182.3$. Model 2 had a function value of $FV_2 = 478537.60088236$ with $K_2 = 47$ and an AIC of $AIC_2 = 957169.2$. Because $AIC_2 < AIC_1$, Model 2 is favored over Model 1, as indicated by the LRT.

5.0 Model Definition File Format

The model definition file created by ROSTER conforms to Extensible Markup Language (XML) 1.0 as found at www.w3.org/xml. The model definition file can be modified using an XML editor or any other text editor. All terminology used in the following, such as “tag,” “element,” and “attribute,” is from the XML 1.0 specification.

Typically a custom model definition may be created by using ROSTER to save the default model definition to a file. This file can then be used as a starting point and edited to create a new model definition. The first line of the file must be:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
```

The root element must be **roster**. The allowable child elements are

- **validate**
- **fix**
- **equate**
- **initialSeeds**

These top-level elements can appear in any order in the file.

5.1 Validate Element

The **validate** element is not actually a part of the model definition, but gives ROSTER the ability to determine if the model definition is appropriate for the data currently loaded.

The **validate** tag must be specified once in the model definition file. The tag takes no attributes, and there are three required child elements, each of which must occur once (in any order).

- **juvenileSites**
- **adultSites**
- **ageClasses**

They specify the number of juvenile detection sites, the number of adult detection sites, and the number of age classes, respectively. The content for each must be a positive integer number. The following model definition fragment specifies that the model is defined for data with six juvenile detection sites, four adult detection sites, and two age classes.

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
...
<validate>
<juvenileSites>6</juvenileSites>
<adultSites>4</adultSites>
<ageClasses>2</ageClasses>
</validate>
...
```

5.2 Fix Element

The **fix** element allows a parameter to be removed from the model and replaced by a fixed constant. It has one required attribute **value** which takes a decimal number for its value. The **fix** element requires one or more **param** elements that have a parameter name as their content. The following fragment replaces two parameters with the constant “1.0”.

```
...
<fix value="1.0">
<param>R11</param>
<param>R12</param>
</fix>
...
```

5.3 Equate Element

The **equate** element allows two or more parameters to be represented by a single parameter in the model. The tag takes no attributes, and requires two or more **param** elements as described above for the **fix** element. The following fragment causes three model parameters to be represented by a single parameter:

```

...
<equate>
<param>sC711</param>
<param>sC712</param>
<param>sC713</param>
</equate>
...

```

5.4 InitialSeeds Element

The `initialSeeds` element is used to specify initial parameter seeds to be used instead of the default seeds. It allows only one child element: “seed” which may occur one or more times.

The `seed` element requires two child elements that must each occur once in any order: `param` and `value`. The following will cause ROSTER to use 0.7 as the initial seed for the parameter *s1*, and 0.9 as the initial seed for the parameter *pC83*.

```

...
<initialSeeds>
  <seed>
    <param>s1</param>
    <value>0.7</value>
  </seed>
  <seed>
    <param>pC83</param>
    <value>0.9</value>
  </seed>
</initialSeeds>
...

```


6.0 Data Sources

The primary source of the PIT-tag data to be used with ROSTER is detections in juvenile bypass systems and adult fish ladders at dams on the Columbia and Snake rivers. These data may be downloaded from the Pacific States Marine Fisheries Commission PTAGIS database, and formatted for ROSTER using UW PIT-tag processing software PitPro, available at <http://www.cbr.washington.edu/paramest/pitpro/>. It is also possible to use spawning grounds or hatcheries as the first and/or last detection sites, as long as detection effort is constant throughout the period when the tagged fish are either leaving or arriving at those sites, respectively. Only a single spawning ground location or hatchery should be used as the first detection site. Multiple spawning grounds and hatcheries may be used as the final field in PitPro, resulting in data pooled across location for the final detection site.

The model used in Program ROSTER is appropriate for spring and summer Chinook salmon, Sockeye salmon, and Steelhead. Because the model does not accommodate the overwintering behavior of fall Chinook salmon, data from this species should not be analyzed with Program ROSTER.

6.1 Estuary: Towed Array Data

Detections of PIT tags are available from the towed array in the Columbia River estuary (code TWX in PTAGIS). These data have been used elsewhere to estimate juvenile survival to the Bonneville tailrace, and may allow for direct comparison of post-Bonneville survival of transported and non-transported individuals. However, there are several practical considerations that preclude using TWX observations in ROSTER at this time. First, detection probabilities at the towed array are typically very low, with $< 1\%$ to 4% of tagged fish previously detected at Bonneville subsequently detected at the towed array in recent years. This means that the standard errors on survival estimates based on TWX detections will be very large, and the estimation routine may have difficulty fitting the model because of the very small detection probabilities. Second, the model used in ROSTER assumes that there is no juvenile detection site downstream of the release site of transported fish. Thus, TWX observations will appear to be adult detections, and the transport-inriver ratio parameters (R_{ij}) will reflect the relative effect of transportation on survival only to the towed array, rather than to the first adult site. Also, the performance measures will report “ocean” survival only to the estuary as juveniles (towed array), rather than to the first adult site. Thus, ROSTER is not currently designed to incorporate TWX observations.

6.2 Avian Predation Data: Detections From Bird Colonies

Using PIT-tag detections from bird colonies in ROSTER is appealing because it would provide extra detection sites, and potentially help identify and quantify sources of mortality; however, use of such data is inappropriate for ROSTER. It is problematic to use detections from mid-Columbia bird colonies (e.g., Crescent Island) due to uncertainty about location of the predation event (e.g., upstream or downstream of McNary Dam). Misspecification of the reach where predation occurred may result in biased estimates of reach survival. Detections from Rice Island and East Sand Island, downstream of Bonneville Dam, are inappropriate for use in ROSTER for the same reasons that detections from the towed array are inappropriate: these detections will be viewed as adult detections, and transportation parameters and performance measures will be inappropriately estimated. Thus, it is inappropriate to use avian predation detections with ROSTER.

6.3 Harvest Data

Migrating adults pass through several fisheries in the lower Columbia River, in particular the Zone 5 and Zone 6 fisheries. Incorporating harvest rates from these fisheries would help identify sources of mortality and provide tools to explore effects of different harvest management strategies. If harvest rate estimates are independently known, it is possible to adjust reach survival estimates to reflect only non-harvest mortality. If harvest rates are unknown, however, it is impossible to adjust for harvest, even with PIT-tag recoveries from the fisheries, unless the tag recovery rate is also known. Without an estimate of the tag recovery rate, it is impossible to separate the harvest rate from the tag recovery rate. Thus, data from PTAGIS alone are insufficient for incorporating harvest data into ROSTER.

7.0 Sample Data Sets

Two sample data sets are included with ROSTER. The first data set (`data1.txt`) contains PIT-tag detection history data from hatchery-raised spring Chinook salmon (*Oncorhynchus tshawytscha*) released in the Snake River Basin upstream of Lower Granite Dam in 1999. The second data set (`data2.txt`) contains PIT-tag detection histories from hatchery-raised spring Chinook salmon released in the Snake River upstream of Lower Granite Dam in 2002. For both data sets, release and detection data were obtained from the PTAGIS database maintained by the Pacific States Marine Fisheries Commission, and formatted for ROSTER by UW software PitPro. Here, the steps required to fit the ROSTER model to these data sets are shown.

7.1 Example 1: 1999 Spring Chinook Salmon

This data set is `data1.txt` in the sample data folder that comes with ROSTER.

Records were available for $N = 178,260$ spring Chinook salmon. Juvenile detections were available from Lower Granite (LGR), Little Goose (LGO), Lower Monumental (LMO), McNary (MCN), John Day (JD), and Bonneville (BON) dams. Adult detections were available in the years 2000, 2001, and 2002 from the adult fish ladders at Bonneville and Lower Granite dams. Fish were transported from LGR, LGO, and LMO.

To load this data set into ROSTER, navigate to the data file in the sample data folder, enter the settings on the Data tab as shown in Figure 2.1, and click the “Load” button. After loading the data, the Data Summary is available from the Reports menu. Figure 7.1 shows the first part of the Data Summary output.

The default model is sufficient for this data set. Set the model definition settings as shown in Figure 2.2. The value in the “Extent of Transportation Effects” window should be set to 7 (the first adult site). Click the “Load” button to load the default model. For future reference, refer to this model as Model 1.1.

Before running the estimation routine, it may be useful to examine the structure of the default model file (also see Section 5). The model file is in Extensible Markup Language (XML) format; hence, the default file extension is `.xml`. Click on “Save Model Definition ...” under the File menu, and save the file; this is file `model11_1.xml` in the sample data folder. Open the saved model file using an appropriate XML editor or any text editor. There are several types of code in the

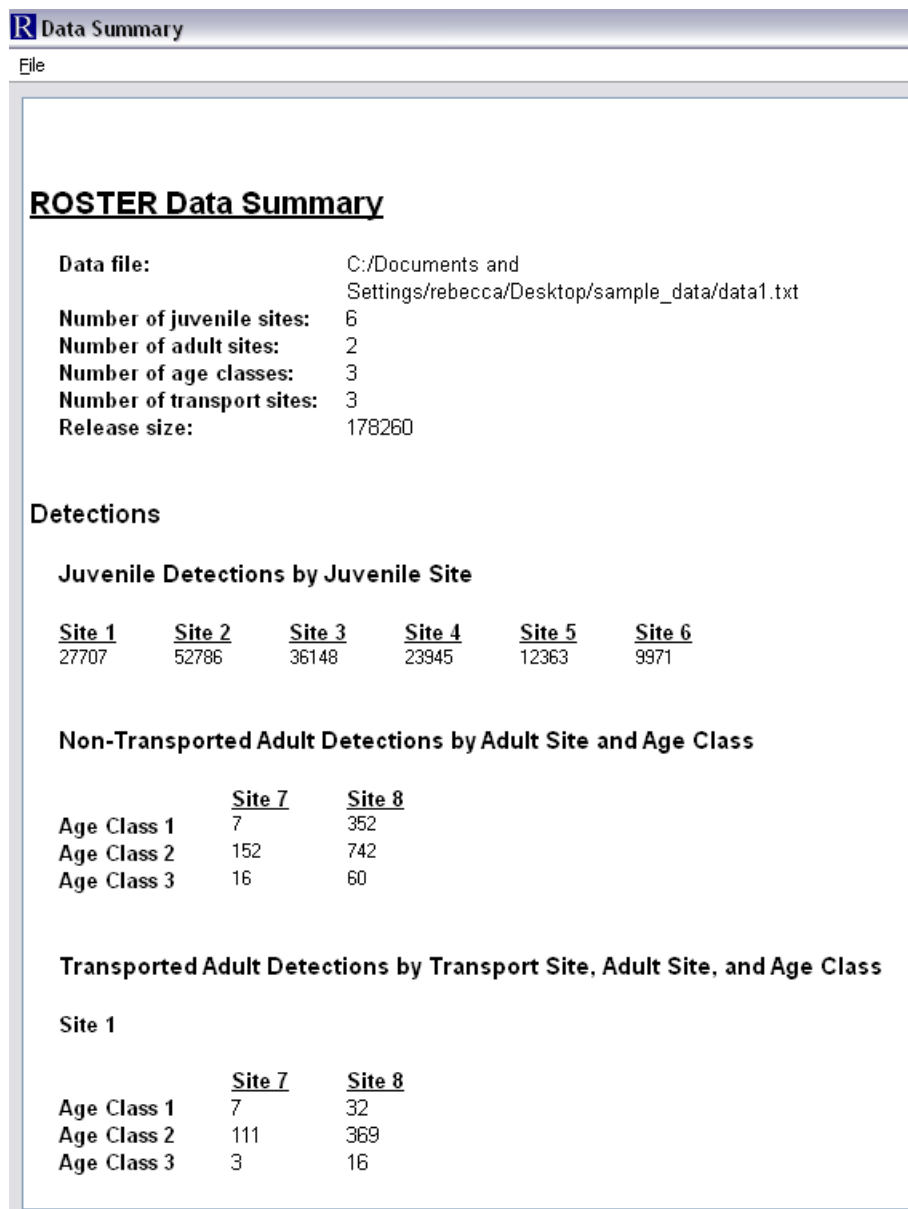
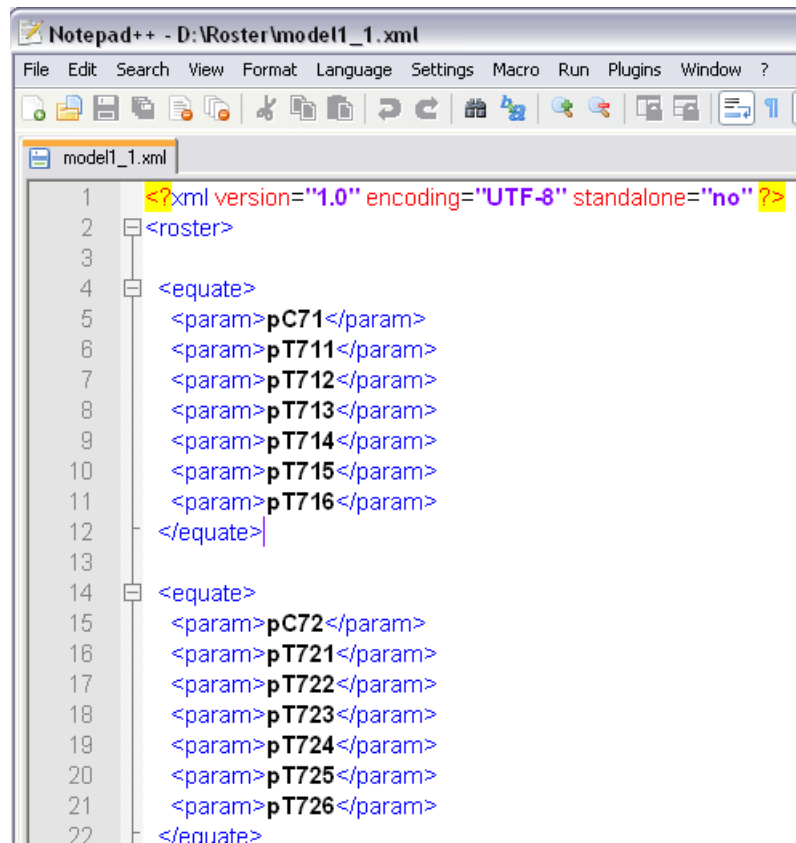


Figure 7.1: The top portion of the Data Summary report for the first sample data set (data1.txt).

default model file. After the heading comes a section where certain adult parameters are equated (Figure 7.2). The first set of parameters equated are the detection parameters at (adult) site 7 for age class 1. There are seven parameters listed as equated, referring to the seven possible treatment groups: “control” ($pC71$), and transport groups from each of the six juvenile dams ($pT711$ through $pT716$). The naming of model parameters for the adult portion of ROSTER follows a convention. The first character (here, p) indicates the type of parameter (p = detection probability, c = censoring probability, and s = survival probability). The second character indicates the treatment group the parameter describes (C = control or non-transported, T = transported). The third character is a number indicating the detection site; recall that detection sites are numbered consecutively, so if there are six juvenile sites, then “7” indicates the first adult site. The fourth character is a number indicating the adult age (or ocean) class, typically 1, 2, or 3. If the parameter is for a transport group, then there will be a fifth character indicating the transport site. Thus, $pT713$ is the detection probability at site 7 for age class 1 fish that were transported at site 3 as juveniles. In order to incorporate the maximum amount of flexibility, ROSTER names parameters for transport groups from all juvenile sites, even those that are not transport dams. For example, the file `model11_1.xml` shows parameters $pT714$, $pT715$, and $pT716$ even though sites 4, 5, and 6 (McNary, John Day, and Bonneville, respectively) were not transport sites. However, these parameters for non-existent transport groups are not actually used in fitting the model, so the user may ignore them. In general, it is safest to keep them equated to the “control” parameters (e.g., $pC71$) in the model definition file.

The default model in file `model11_1.xml` shows that adult detection and censoring parameters are equated across transport and non-transport groups for detection site 7 and all age classes, and that the “last reach” (λ) parameters are also equated for each age class. Figure 7.3 shows the next section of code, where certain parameters are fixed to either 0 or 1. In particular, transportation parameters for non-transport sites (sites 4, 5, and 6) are fixed to 0, as are adult censoring parameters for sites and age classes where no adult censoring occurred. T/I parameters (R_{ij}) are fixed to 1 for sites i for which there was no transportation. Next is a line indicating the extent of transportation effects entered on the ROSTER interface (i.e., 7), followed by the values for the parameters set by the user (barge survival and untagged transport parameters). Last is a necessary section showing the number of juvenile and adult sites, and the number of adult age classes. The model will not load correctly if this section, or the transportation effects extent line, is missing.

Having examined the different parts of the default model, run the estimation routine: click the “Estimate Parameters” button on the Estimation tab. The message window should indicate that the estimation completed after about 53 iterations, and the Estimation Results report should open. Figure 7.4 shows the



```
1 <?xml version="1.0" encoding="UTF-8" standalone="no" ?>
2 <roster>
3
4 <equate>
5   <param>pC71</param>
6   <param>pT711</param>
7   <param>pT712</param>
8   <param>pT713</param>
9   <param>pT714</param>
10  <param>pT715</param>
11  <param>pT716</param>
12 </equate>
13
14 <equate>
15   <param>pC72</param>
16   <param>pT721</param>
17   <param>pT722</param>
18   <param>pT723</param>
19   <param>pT724</param>
20   <param>pT725</param>
21   <param>pT726</param>
22 </equate>
```

Figure 7.2: The top part of the default model file (model1_1.xml) for the first sample data set.

```

89     <param>lambdaT34</param>
90     <param>lambdaT35</param>
91     <param>lambdaT36</param>
92 </equate>
93
94 <fix value="0">
95     <param>t4</param>
96     <param>t5</param>
97     <param>t6</param>
98     <param>cC71</param>
99     <param>cC73</param>
100 </fix>
101
102 <fix value="1">
103     <param>R41</param>
104     <param>R42</param>
105     <param>R43</param>
106     <param>R51</param>
107     <param>R52</param>
108     <param>R53</param>
109     <param>R61</param>
110     <param>R62</param>
111     <param>R63</param>
112 </fix>
113
114 <transportEffectExtent>7</transportEffectExtent>
115
116 <bargeSurvival>
117     <value>0.98</value>
118     <variance>0</variance>
119 </bargeSurvival>
120
121 <untaggedTransport>
122     <values>1 1 1 0 0 0</values>
123     <variances>0 0 0 0 0 0</variances>
124 </untaggedTransport>
125
126 <validate>
127     <juvenileSites>6</juvenileSites>
128     <adultSites>2</adultSites>
129     <ageClasses>3</ageClasses>
130 </validate>
131
132 </roster>

```

Figure 7.3: The bottom part of the default model file (model1_1.xml) for the first sample data set.

first portion of the estimation results window. Parameters that were fixed (not estimated) in the default model (e.g., transportation parameters for non-transport sites, t_4 , t_5 , and t_6) show the value to which they were fixed, and have NA as their standard error. In general, any parameter with standard error shown as NA was not estimated. The performance measure estimates are shown after the model parameter estimates, first with jacks included and then excluding jacks. ROSTER reports transportation effect measures and differential mortality measures on adult returns to two sites: the adult site that is the extent of the transportation effects used in the model (site E), and the final adult site. In this case, these sites are site 7 (BON) and site 8 (LGR), respectively. The performance measure estimates using return rates to the different sites are slightly different, but generally not significantly (Figure 7.5).

Now fit the model with transportation effects extended to the second and final adult site. On the Estimation tab of the ROSTER interface, change the value in the “Extent of Transportation Effects” window to 8, and click the “Load” button. For future reference, refer to this model as Model 1.2. Before running the estimation routine, save the model file and examine its structure. This is the file `model1_2.xml` in the sample data folder. Notice that in this model file, the adult parameters for the non-transported fish (e.g., $pC71$, along with $pT714$, $pT715$, and $pT716$) are equated, while the parameters for transported fish (e.g., $pT711$, $pT712$, and $pT713$) are equated separately. This model estimates different detection, censoring, and lambda parameters for transported fish compared to non-transported fish.

Return to the Estimation tab, and click on “Estimate Parameters.” After 64 iterations, the estimation routine should be complete and a new Estimation Results report should open. In the results report for this model, notice that the adult detection and censoring parameters at site 7 and the lambda parameter are reported separately for non-transported fish and for each transport group. However, the values are the same among all the transport groups. This is because the default model equates adult parameters among the transport groups, but distinguishes between transported and non-transported fish, up until the site that is defined as the extent of transportation effects. In this case, the extent of transportation effects is the last adult site, so all adult parameters are estimated separately for transported and non-transported fish. Also notice that the transportation effect and differential mortality performance measures are now reported only to site 8, because this site is both the extent of transportation effects and the final adult site.

It may be reasonable to expect that the lambda parameters will be smaller for transported fish than for non-transported, if transported fish have higher straying rates and so lower perceived adult survivals. This appears to be the case in two of the three age classes, although the differences are not significant. However, it seems less likely that transported and non-transported fish would have different detection and censoring probabilities at the adult sites. Thus, it may be useful to

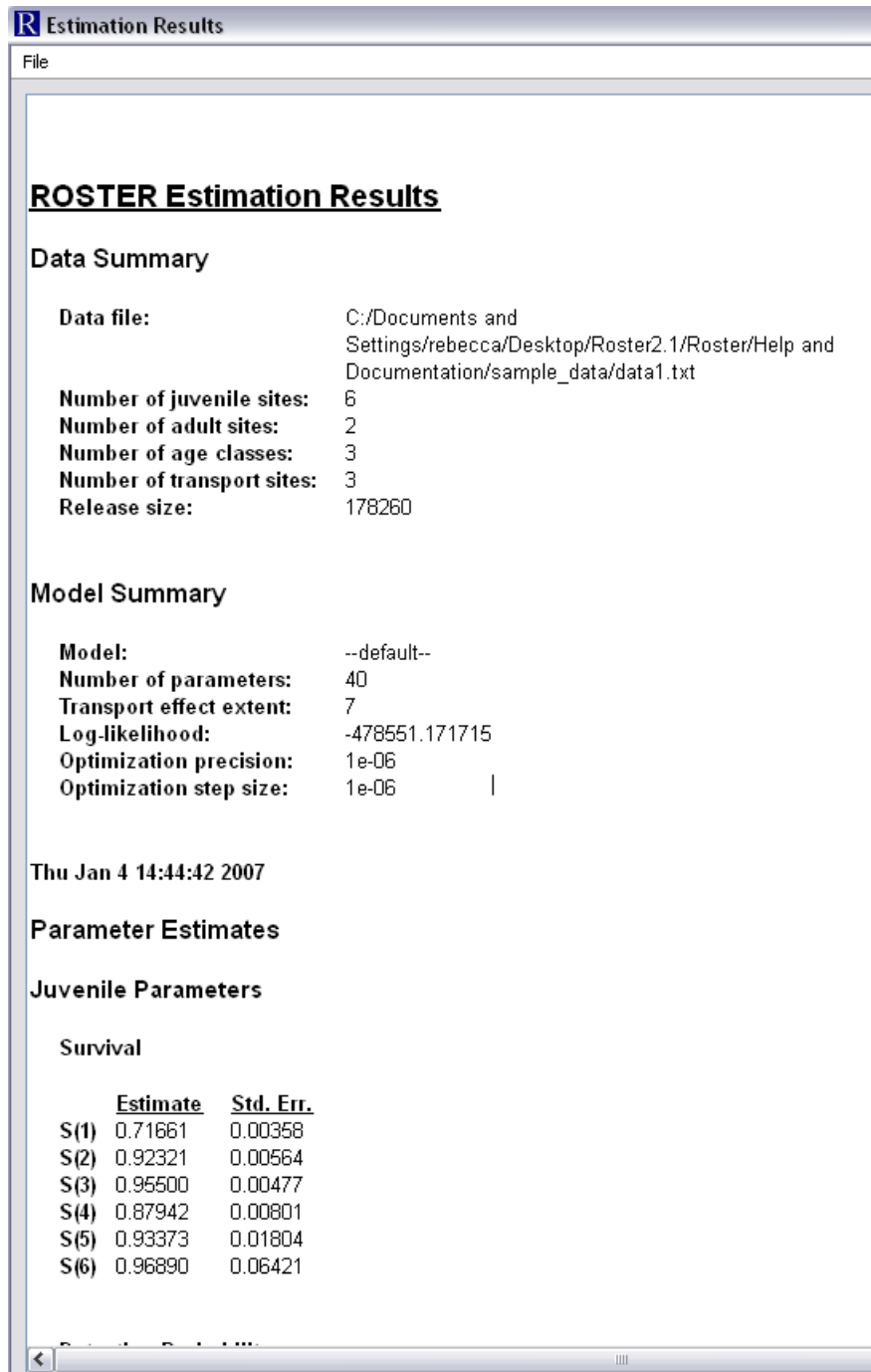


Figure 7.4: The top portion of the estimation output for the first sample data set (data1.txt), analyzed with the default model and extent of transportation effects set to 7.

R Estimation Results		
File		
Isolated Site-Specific T/I Ratio (Ri)		
<u>to site 7</u>	<u>Estimate</u>	<u>Std. Err.</u>
Juvenile transport site 1	1.77751	0.10933
Juvenile transport site 2	1.70440	0.11699
Juvenile transport site 3	0.70514	0.17269
<u>to site 8</u>	<u>Estimate</u>	<u>Std. Err.</u>
Juvenile transport site 1	1.75373	0.09611
Juvenile transport site 2	1.68259	0.10735
Juvenile transport site 3	0.69719	0.17042
Contextual Site-Specific T/I Ratio for Tagged Fish (RCi)		
<u>to site 7</u>	<u>Estimate</u>	<u>Std. Err.</u>
Juvenile transport site 1	1.64881	0.09556
Juvenile transport site 2	1.71372	0.11735
Juvenile transport site 3	0.70514	0.17269
<u>to site 8</u>	<u>Estimate</u>	<u>Std. Err.</u>
Juvenile transport site 1	1.63083	0.08676
Juvenile transport site 2	1.69204	0.10783
Juvenile transport site 3	0.69719	0.17042

Figure 7.5: Part of the performance measure estimates in the Estimation Result report for the first sample data set (data1.txt), analyzed with the default model and extent of transportation effects set to 7.

explore a model that estimates unique lambda parameters for transported and non-transported fish, but common detection and censoring parameters. This is not the structure of the default model, so it is necessary to modify a model definition file. With the extent of transportation effects set to 8 and the default model loaded, click on “Save Model Definition ...” under the File menu to save the file. Open the file using the appropriate software, and edit it so that all the detection and censoring parameters are equated within each age class. Also, remove the parameters $cT711$ and $cT731$ from the list of parameters set to 0. Save the model as Model 1.3, select “Model Definition File” on the ROSTER interface, select the model just saved, and load the model. Alternatively, use `model11_3.xml` from the sample data folder. Run the estimation, and a new Estimation Result report should open after 57 iterations. Notice that adult detection and censoring parameters are still listed separately for the non-transported group and for each transport group, but the estimates are the same across all the groups now. The lambda estimates still differ for the non-transported group and the transport groups.

Which of the three models considered is the best fit or most appropriate? These three models are nested, with Model 1.1 nested in Model 1.3, and both nested in Model 1.2. This means that we can use LRTs to compare these three models, as described in Section 4.2.1. In fact, this data set and models 1.1 and 1.2 are highlighted in the discussion of LRTs, where it was found that the added structure between the non-transported and transported groups was significant ($p = 0.0003$), and Model 1.2 was selected (Model 2 in Section 4.2.1) over Model 1.1. What about Model 1.3, where transportation effects are estimated only in the lambda parameters, with the detection and censoring parameters equated over juvenile migration method (inriver or transportation)? The function value for Model 1.3 is $FV_3 = 478545.01244264$ with 43 parameters, so the LRT test statistic for models 1.2 and 1.3 is $X = 14.82312$ with $d = 4$. The critical value for this test (at the 95% significance level) is $\chi^2_{4,0.95} = 9.4877$. Because $X > \chi^2_{4,0.95}$, we reject the null hypothesis of common detection and censoring parameters between non-transported and transported groups, and conclude that Model 1.2 should be used.

7.2 Example 2: 2002 Spring Chinook Salmon

This data set is `data2.txt` in the sample data folder that comes with ROSTER. Records were available for $N=303,096$ spring Chinook salmon released in the Snake River upstream of Lower Granite Dam in 2002. Juvenile detections were available from Lower Granite (LGR), Little Goose (LGO), Lower Monumental (LMO), McNary (MCN), John Day (JD), and Bonneville (BON) dams. Adult detections were available in the years 2003, 2004, and 2005 from the adult fish ladders at Bonneville, McNary, Ice Harbor, and Lower Granite dams. Fish were

transported from LGR, LGO, LMO, and MCN.

To load this data set into ROSTER, first set the data description settings to 6 juvenile sites, 4 adult sites, and 3 adult age classes, and click the “Load” button on the program interface. Figure 7.6 shows part of the Data Summary report for this data set. Set the value in the “Extent of Transportation Effects” window to 7 and load the default model.

As with the first sample data, the default model for this data set (file `model12_1.xml` in the sample data folder) equates detection and censoring parameters for each age class and detection site, regardless of juvenile migration method (i.e., transport versus non-transport). Because this data set has more than two adult detection sites, the default model in `model12_1.xml` also equates adult survival parameters for each age class and site. Unlike the first data set, this data set includes transportation groups that were not detected as adults in certain age classes. In particular, no fish transported from site 3 (LMO) was detected as an age-3 adult, and no fish transported from site 4 (MCN) was detected as either an age-1 or an age-3 adult. Thus, the default model fixes the corresponding age- and site-specific T/I parameters (i.e., R_{33} , R_{41} , and R_{43}) to 1; see Figure 7.7.

Now run the estimation routine by clicking “Estimate Parameters” on the Estimation tab (make sure the model is loaded first). The estimation routine stops after about 30 iterations, and the Estimation Results report opens. The report shows “NA” for the standard errors of the parameters that were fixed in the model file (e.g., all adult censoring parameters except for site 7, age class 2), indicating that these parameters were not estimated. Also, the report shows “NA” for the point estimates of parameters R_{33} and R_{43} , which is appropriate because these parameters were not estimated. However, the report also shows “NA” for the standard error of the survival probability s_6 , the detection probability p_6 , and other estimated parameters, and the survival parameter for site 9, age class 1 (i.e., s_{91C}) is estimated at 2.62247 with “NA” for the standard error, when its estimate should never be much greater than 1. These problems indicate that the estimation routine did not converge, even though it stopped and did not give a warning or error message. There is a problem with the model file.

Examination of the Parameters report shows that several survival estimates were very high, including s_5 , s_{C91} , and s_{C93} and the associated survival parameters for transported fish. It is possible that because ROSTER had to estimate these values that are approximately 1, it could not also estimate certain other parameters. Accordingly, it is worth fixing these parameters to 1 (or 0.999999) and trying again to estimate the parameters. Figure 7.8 shows part of the edited model file with these survival parameters fixed; the edited model file, named `model12_2.xml`, is in the sample data folder. Select this model definition file for use in the program interface, and click the “Estimate Parameters” button. After 68

R

Data Summary

File

ROSTER Data Summary

Data file:

C:/Documents and Settings/rebecca/Desktop/sample_data/data2.txt

Number of juvenile sites:

6

Number of adult sites:

4

Number of age classes:

3

Number of transport sites:

4

Release size:

303096

Detections

Juvenile Detections by Juvenile Site

Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
42491	51170	70787	59500	22010	23560

Non-Transported Adult Detections by Adult Site and Age Class

	Site 7	Site 8	Site 9	Site 10
Age Class 1	203	214	214	202
Age Class 2	1124	1063	1067	1038
Age Class 3	29	23	22	21

Transported Adult Detections by Transport Site, Adult Site, and Age Class

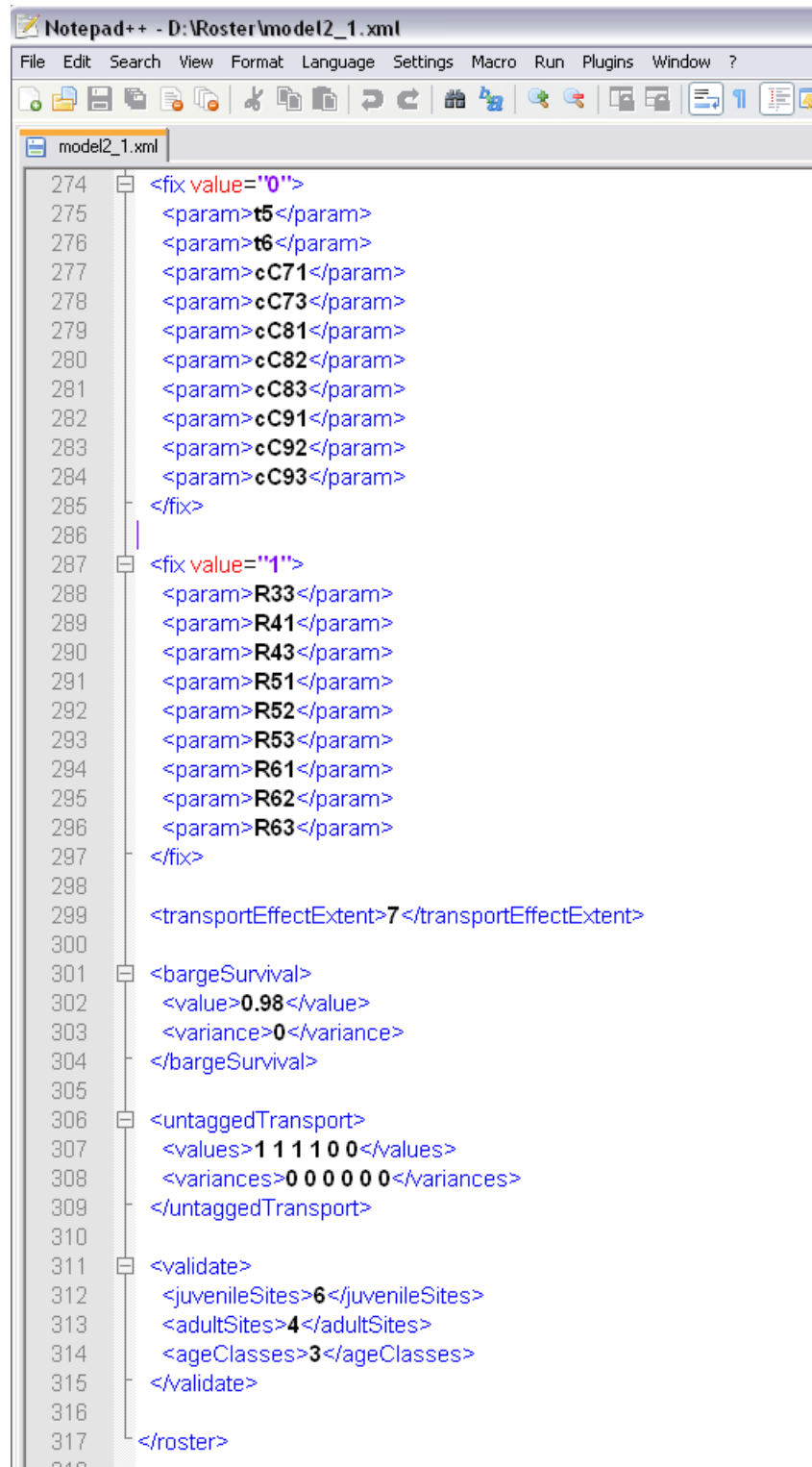
Site 1

	Site 7	Site 8	Site 9	Site 10
Age Class 1	23	24	22	23
Age Class 2	129	113	109	105
Age Class 3	6	4	3	4

Site 2

	Site 7	Site 8	Site 9	Site 10
Age Class 1	25	28	28	28
Age Class 2	111	103	103	99
Age Class 3	3	2	3	3

Figure 7.6: The top of the Data Summary report for the second sample data set.



```
274 <fix value="0">
275   <param>t5</param>
276   <param>t6</param>
277   <param>cC71</param>
278   <param>cC73</param>
279   <param>cC81</param>
280   <param>cC82</param>
281   <param>cC83</param>
282   <param>cC91</param>
283   <param>cC92</param>
284   <param>cC93</param>
285 </fix>
286
287 <fix value="1">
288   <param>R33</param>
289   <param>R41</param>
290   <param>R43</param>
291   <param>R51</param>
292   <param>R52</param>
293   <param>R53</param>
294   <param>R61</param>
295   <param>R62</param>
296   <param>R63</param>
297 </fix>
298
299 <transportEffectExtent>7</transportEffectExtent>
300
301 <bargeSurvival>
302   <value>0.98</value>
303   <variance>0</variance>
304 </bargeSurvival>
305
306 <untaggedTransport>
307   <values>1 1 1 1 0 0</values>
308   <variances>0 0 0 0 0 0</variances>
309 </untaggedTransport>
310
311 <validate>
312   <juvenileSites>6</juvenileSites>
313   <adultSites>4</adultSites>
314   <ageClasses>3</ageClasses>
315 </validate>
316
317 </roster>
```

Figure 7.7: The bottom part of the default model file (model2_1.xml) for the second sample data set.

iterations, the optimization routine ends, and the revised Estimation Results report is available. Note that the parameters $s5$, $s91$, and $s93$ all have point estimates equal to 1, with NA for standard error. The NA standard error indicates that these values were fixed to 1, rather than estimated at 1. This time, the estimates $\hat{s}6 = 0.95144$ with $\hat{SE} = 0.06117$, and $\hat{p}6 = 0.17882$ with $\hat{SE} = 0.01151$, are given for $s6$ and $p6$, respectively (Figure 7.9). The user should explore whether the estimation can be performed by fixing some subset of the survival parameters $s5$, $s91$, and $s93$ to 1 (or 0.999999).

It is generally not necessary for the user to pick initial values (seeds) for ROSTER, because ROSTER calculates the Method-of-Moments estimates to use as seeds. However, the user may select a different seed if desired, and adjust the model file accordingly. The syntax for doing this is shown in Chapter 5. The complexity of the statistical model used by ROSTER means that the success of the estimation routine may be sensitive to using appropriate seeds.

The model fit for the second sample data set assumes that transportation effects end at the first adult detection site (i.e., Bonneville). We can explore whether the data indicate upriver adult effects of juvenile transportation by fitting models that estimate different adult survival, detection, and censoring parameters for transported versus non-transported fish. On the Estimation tab of the user interface, select the default model, and then change the extent of transportation effects to 10; load the model. Before running the estimation routine, save the model and examine it; this model is `model12_3.xml` in the sample data folder. Notice that in this model file, the adult parameters for the non-transported group are equated, while the adult parameters for the four transport sites are equated separately. This model estimates unique adult survival, detection, and censoring parameters for transported versus non-transported fish all the way to the final detection site.

Return to the Estimation tab, and click on “Estimate Parameters.” An error window pops up, announcing an “estimation numerical error,” and the messages window also announces a numerical error. The model file will have to be edited. To determine how to edit the model file, first click on the “Load” button to reload the default model, and then examine the Parameters report. Again, we see that the seed for $s5$ is 1, while the seeds for the parameters $sC91$, $sC93$, $sT911$, and $sT931$ (and the parameters equated to them) are all at or above 1. We need to fix these parameters to 1 (or 0.9999999) in the model file. The file `model12_4.xml` in the sample data folder has the necessary modifications. Load that model, and click on “Estimate Parameters” again. This time, the estimation routine converges after 71 iterations. Upon examination, the Estimation Results report provides reasonable results (i.e., probability parameters are estimated between 0 and 1, and standard errors of estimated parameters are not NA), so we conclude that the model converged. It may be wise to run the estimation routine again, after introducing random noise in the “Parameter Seeds...” option under the Diagnostics menu. Try


```

284     <param>cC93</param>
285     </fix>
286
287     <fix value="1">
288         <param>R33</param>
289         <param>R41</param>
290         <param>R43</param>
291         <param>R51</param>
292         <param>R52</param>
293         <param>R53</param>
294         <param>R61</param>
295         <param>R62</param>
296         <param>R63</param>
297         <param>s5</param>
298         <param>sC91</param>
299         <param>sC93</param>
300     </fix>
301
302     <transportEffectExtent>7</transportEffectExtent>
303
304     <bargeSurvival>
305         <value>0.98</value>
306         <variance>0</variance>
307     </bargeSurvival>
308
309     <untaggedTransport>
310         <values>1 1 1 1 0 0</values>
311         <variances>0 0 0 0 0 0</variances>
312     </untaggedTransport>
313
314     <validate>
315         <juvenileSites>6</juvenileSites>
316         <adultSites>4</adultSites>
317         <ageClasses>3</ageClasses>
318     </validate>
319
320 </roster>

```

Figure 7.8: The bottom part of the edited model file for the second sample data set (model2_2.xml), with survival parameters s_5 , s_{91} , and s_{93} fixed to 1.

R Estimation Results		
File		
Parameter Estimates		
Juvenile Parameters		
Survival		
	<u>Estimate</u>	<u>Std. Err.</u>
S(1)	0.72145	0.00249
S(2)	0.94346	0.00431
S(3)	0.98416	0.00419
S(4)	0.83232	0.00458
S(5)	1.00000	NA
S(6)	0.95144	0.06456
Detection Probability		
	<u>Estimate</u>	<u>Std. Err.</u>
P(1)	0.19432	0.00106
P(2)	0.26846	0.00127
P(3)	0.40481	0.00168
P(4)	0.41823	0.00228
P(5)	0.15769	0.00121
P(6)	0.17882	0.01215
Censoring Probability (conditional on detection)		
	<u>Estimate</u>	<u>Std. Err.</u>
C(1)	0.10770	0.00150

Figure 7.9: Part of the Estimation Results report for the second sample data set, using model “model2_2.xml” (shown in Figure 7.8) in the Sample Data Set folder.

entering a very small amount of proportional noise, say 0.01. You may receive an error message when running the estimation routine with noise in the seeds, due to the complexity of the model. If this happens, you will need to reload the model, and try clicking the “Estimate Parameters” button again. Eventually, you should find a set of random seeds that leads to convergence at the same function value as initially seen without noise. This is one way to confirm that the estimation routine has found the correct MLE.

We have fit two models for the second sample data set. Model 2.2 assumed transportation effects ended at the first adult site (Bonneville), and was fit with a function value of $FV_{2.2} = 823983.91550963$ and 49 parameters. Model 2.4 estimated transportation effects to the final adult site (Lower Granite), and was fit with a function value of $FV_{2.4} = 823963.98198837$ and 63 parameters. Model 2.2 is a simpler model than Model 2.4, and in the general case, the structure of Model 2.2 is nested in Model 2.4. However, the data set required fixing certain detection parameters to 0.999999 in the more complex Model 2.4, while they were left free in the simpler Model 2.2 (ROSTER fixed them automatically when the default model with extended transportation effects was loaded; see the file `model2_4.xml`). This means that even though Model 2.2 is a simpler model than Model 2.4, it is not actually nested in Model 2.4. Thus, we cannot compare the two models with a LRT, but must instead use AIC. The AIC for Model 2.2 is $AIC_{2.2} = 1648066$, and the AIC for Model 2.4 is $AIC_{2.4} = 1648054$, indicating that the greater flexibility in Model 2.4 enables a considerably better fit to the data than Model 2.2. It may be useful to compare analogous survival parameters for transported and non-transported fish using z-tests; this requires finding the covariance between the two survival estimates from the Covariance Matrix report. It may also be useful to explore models with the extent of transportation effects set at either 8 or 9 to find the most parsimonious model.

8.0 Acknowledgments

This research was supported by the National Oceanic and Atmospheric Association (NOAA) under contract number AB133F-04SE-0584, and by the Bonneville Power Administration under project number 1989-107-00 and contract number 29651. We thank PSMFC and PTAGIS for providing access to PIT-tag data.

