

NOTICE:

The information on the next few pages is my personal notes that I have made while trying to learn and use program MARK. Any mis-interpretation of information provided to me by other sources or mistakes I have made in these notes are completely my own. This is not intended to be any sort of guide book, or help reference. These are simply a few hints and tricks that I have learned through trial and error. Please do not reference nor distribute any of these materials. Again these are my personal notes and have not been reviewed by any one associated with CSU or any other university or agency.

Tricks to formatting data.

Bringing your data into MARK is a lot more difficult than it probably should be. I have found several tricks that have really help me.

FORMATTING DATA

First the actual format of the data. MARK data will need to be coded in a "capture-history" format. Basically a series of 0's and 1's corresponding to the presence or absence of an individual at that sample time. For example if you sample 5 years and fish tag number A3D is collected in years 3 and 4 only then the capture history would look like:

```
00110 /*A3D*/ 1;
```

An alternative format would summarize all the fish that had this capture history

```
00110 87;
```

Where 87 indicates that 87 fish have this capture history.

*NOTE that when you build your data entry files they need to look exactly like the ones above, no spaces in your capture history sequence, no stray tabs, etc.

Once you have this data in the coded format you are not quite finished. We need to tweak the data a little more before it will go into MARK. I usually take the data into EXCEL or some similar program and do two things. First we need add a column on 1 to each of our capture histories. What this says is that each of these capture histories is unique. MARK may report a warning saying that some of the capture histories are repeated but that is OK. If we were to get really fancy we could tell MARK how many fish had each combination of capture histories but I don't bother doing that. Second we need to "comment out" the tag number or MARK will think that this number is also a capture history. This is done just like in many programs by using the /* Insert your TAG NUMBER*/ commenting technique. We then need to add a semicolon at the end of each line. Here is an example:

```
1000100000 1 /* F4446 */ ;
```

Once you have this done in excel, just save the data as an Excel file to have a copy

of it, then click save as, chose plain text, and re-save the file and give it the necessary extension for MARK to read the data. This is done by putting the name of the file in quotation marks and including the .inp extension in the name. For example if the file was called pearl then we would enter the file as "pearl.inp" and then save it. This should save it as an inp file. Excel can be stubborn so if you have to save it as a text document, then open it in Notepad, and then re-save it as I outlined above.

DATA INTO MARK

Open MARK and start a new file.

Name your file a unique name each time or it will copy over your old files.

Press "Click to select file"

Bring your file into MARK, then click view file to make sure it looks ok.

Next I'll describe a Jolly-Seber type analysis, but the other modules work in a very similar way.

Chose the first Data type called "Recaptures Only", the information bubble will pop up explaining that this is the standard Cormack-Jolly-Seber model.

Set the encounter history, this is the number of periods that your data is collected over that we want to examine, in my case I had ten years of data so my encounter history was 10.

Click OK

A window will pop up indicating a results data base was created, click ok.

A window will open with the PIM chart available. PIMS represent each model parameter, here you see one for each "cohort/year class" across each sample year.

In the top of the screen click the run heading, then click the "predefined models" button. Using the predefined models is very limiting and also a wimpy way to program. Learn how to use the PIM's and write your own models. You will need to do

so to even begin to use MARK to its potential.

Next steps, on the current screen:

Name the output file

Press select models

Chose the models that you feel are biologically appropriate. What makes a biologically appropriate model? That depends on your study species (exploited yes or no, did mortality change during your sample time yes or no) and sampling events (did the weather change in the middle of your study?). Remember a "dot" model means that that parameter is fixed across all sampling occasions and a "t" model means that that parameter varies through time across all occasions. To fit something "in between", say with 3 of the 5 capture probabilities equal to each other and the other variable, then you have to use PIMS. MARK will list all possible models for the "." and "t" combination, whether they are biologically meaningful will be determined by you. Remember, the simplest model is a $\Phi(\cdot)$, $p(\cdot)$ model, fixed survival, fixed capture probability.

Click OK to run

A "Results Browser" window will now pop up. This outlines the fitting results from all of your models. Remember:

ϕ =survival

p = capture probability

When you look at the QAIC scores you see several models that are very close together in their values. I have heard different rules of thumb, but basically models that are with 3-5 AIC values are very difficult to distinguish. Now we turn and think about the model results from a biological standpoint and decide if several of the models make sense that way, if so, look at the parameter values for the different models and compare the results.

LOOKING AT RESULTS:

Click on the model with the lowest QAIC value, that should highlight the model in blue.

Now click on the fourth icon to the right of the garbage can icon, the little information bubble will say something like "view output from highlighted model in notepad"

This will give you a lot of information but what you are most interested in will look like this

Real Function Parameters of {Phi(.) p(t) Design Matrix coding}
95% Confidence Interval

Parameter	Estimate	Standard Error	Lower	Upper
1:Phi	0.8275935	0.0185978	0.7880469	0.8610623
2:p	0.1148672	0.0215291	0.0789289	0.1642510
3:p	0.1580115	0.0206095	0.1216692	0.2027039
4:p	0.1689232	0.0189020	0.1350326	0.2092617
5:p	0.1970314	0.0199986	0.1607435	0.2391767
6:p	0.1217537	0.0140401	0.0968016	0.1520548
7:p	0.1031626	0.0130543	0.0802380	0.1316990
8:p	0.0238751	0.0056799	0.0149432	0.0379403
9:p	0.0713352	0.0112886	0.0521364	0.0968811
10:p	0.0862194	0.0127959	0.0642224	0.1148262

This example model only had two parameters (Phi and p) as you see there is one phi parameter (we had MARK fix survival, the "dot" model) and the p's are different (we had MARK use a time dependent p).

Now the hardest part, deciding biologically how this model fits into what we know about the life history of an organism. We also want to fit the simplest model possible, particularly with data sets that have low capture probability. If you see parameters with nonsensical SE estimates, or estimates that bound back and forth through time between really large and really small values, then those estimates are not fitting very well and you should use a simpler model for that particular parameter. A simpler model may be using the PIMS to fix survival over a few intervals or to fix capture probability over a few intervals. For example, say you have 8 weeks of sampling, one sample per week, but you know that in weeks 2 and 3 that the weather was really cold and may have impacted your capture probability. A good idea would be to fix capture probability in those weeks equal to each other. This model is then simpler than a model that estimates capture probability in each of the periods, but is more biologically realistic than a model with fixed capture probability across all the sample intervals. I look for the most biologically reasonable model, and then examine how much AIC support is suggested for that model. The key is to find the simplest, most biologically reasonable model.

If you keep working with MARK and you start to build PIMs so you can fit models that aren't either fixed or fully time dependent, remember in looking at PIMS Time moves left to right (columns), cohort moves top to bottom (rows)