4/19/2012

User Guide:
Spreadsheet Model for Ungulate Population data

Model developed by:
G. White and B. Lubow, Colorado Division of Wildlife, Ft. Collins, CO
White, G.C. & Lubow, B.C. (2002) Fitting population models to multiple sources of observed data. *Journal of Wildlife Management,* **66,** 300-309.

User guide developed by:
T. Morrison Wyoming Cooperative Fish and Wildlife Research Unit, 2012

TABLE OF CONTENTS

1. OVERVIEW

The 'Spreadsheet Model' is actually a collection of several different models, each with different assumptions about how juvenile and adult survival rates vary over time. The goal is to identify which of these models is the simplest and best-fits the data. Each model estimates the predicted post-hunt population sizes using classification data, harvest data, survival (if available) and abundance (if available). To do the estimation, the model relies on an optimization program in Excel called 'Solver' to minimize the weighted difference (called the 'penalties' and 'deviances') between field collected data and model-estimated data. After minimizing these values in each of the models, we select the simplest, best-fitting model by comparing relative AICc scores.

2. SPREADSHEET MODEL VS. POP-II

Compared to POP-II, the Spreadsheet Model provides a more statistically rigorous way to estimate population sizes and to compare different models. While this takes some of the 'art' out of modeling ungulate populations, it results in more defensible and repeatable estimates. POP-II allows nearly an order of magnitude greater number of decisions and assumptions to be made by the modeler than does the Spreadsheet Model, illustrating the relative degree of subjectivity in the two models (Table 1). There is still some subjectivity and flexibility in the Spreadsheet Model, which is useful and important for making adjustments to problematic datasets (see *Trouble-Shooting*, Section 8). Nonetheless, there is less need for the user to "tinker" with the model, because the results are optimized based on the data biologists and wardens collect in the field.

The first important difference between the Spreadsheet Model and POP-II is the Spreadsheet Model has no MSI (mortality severity indices) values. In POP-II, MSI's are used to adjust survival rates and allow biologists to align the estimated and observed classification ratios. This introduces one of several semi-subjective steps into the modeling process. The Spreadsheet Model, by contrast, directly estimates juvenile and adult survival rates. Any variation in observed data caused by variation in winter severity should be reflected directly in these estimated survival rates and in the fit of the model (more on this later). The Spreadsheet Model also handles missing data – for example, if classification data were not collected in a particular year, you would still be able to estimate population sizes – and it can more easily handle multiple years of data (*i.e.,* >5 years) at a time. It is arguably wasteful that POP-II cannot always easily incorporate all available data. Finally, the Spreadsheet Model is easy to update with new data or models, which will hopefully allow the department to incorporate increasingly complex understanding of the dynamics of ungulate populations in years to come.

POP-II vs Spreadsheet Comparison

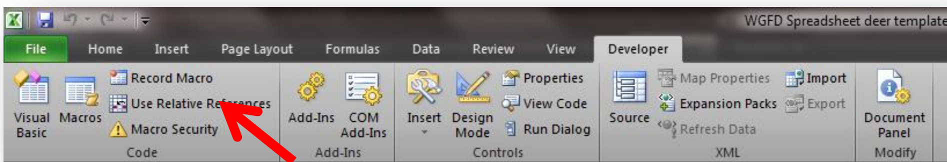
		POP-II	Spreadsheet	# Parameter estimates & assumptions for 15 yr dataset: POP-II	TSJ,CA Model
Model Components	Initial population proportions by age/sex	X		30	
	Pre-season mortality by age/sex	X		30	
	Post-season adult mortality by age/sex	X		30	
	Harvest effort adjustment by age/sex	X		30	
	Annual pre-season mortality severeity index	X		15	
	Juvenile harvest data	X			
	Adult harvest data	X	X		
	Annual post-season mortality severity index	X		15	
	Age-class specific reproductive rate	X		45	
	Wounding loss adjustment	X	X	6	3
	Sex ratio at birth	X	X	1	1
	Initial beginning of bio year total population size	X		1	
	Initial posthunt population size of adults		X		2
	Classification data without SE	X			
	Classification data with SE		X		
	Abundance estimates with SE		X		
	Abundance estimates or trend counts without SE	X			
	Annual juvenile survival and SE		X		15
	Annual adult survival and SE		X		1
General Features	Last age class has 100% mortality	X			
	Juv:Adult Female ratio is exact	X	X		
	Explicit parameter contraints in Solver		X		
	Model improves the longer the time span of data		X		
	Option to consider more than one model		X		
	Prevents overparameterization		X		
	Statistical optimization of model estimates		X		
	Statistical comparison of models		X		
	Model visually fit by modeller	X			
				TOTAL = 203	**TOTAL = 22**

Table 1. *A summary of the major similarities and differences between the Spreadsheet Model and POP-II. Field observed data are shaded in red. Parameters shaded in blue can be either field estimated and/or model estimated (Spreadsheet Model only). The two rightmost columns sum the total number of assumptions and parameter estimates that go into POP-II and the* ***TSJ,CA*** *Spreadsheet Model (i.e., the most complex Spreadsheet Model) for a typical deer dataset consisting of 15 years of harvest and classification data.*

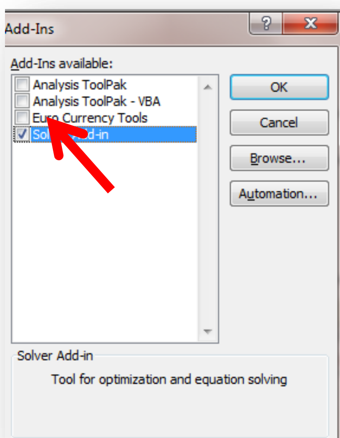
3. GETTING STARTED

To use the Spreadsheet Model, you need several software components:

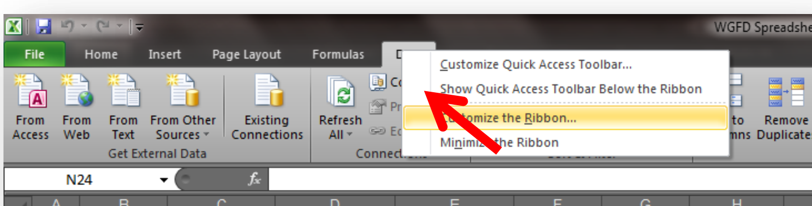
1. **Microsoft Excel** (Excel 2010 is currently the best)
2. **Solver**. Solver is a free add-in program used to find optimal values, given a certain model structure and input data. Solver should be available in all recent MS Excel software (Excel 2003 or later). In Excel 2010, Solver is under the DATA tab at the top of Excel spreadsheets. If it does not appear, you may have to go to the Add-In pop-up window (under the DEVELOPER Tab at the top of Excel 2010)....



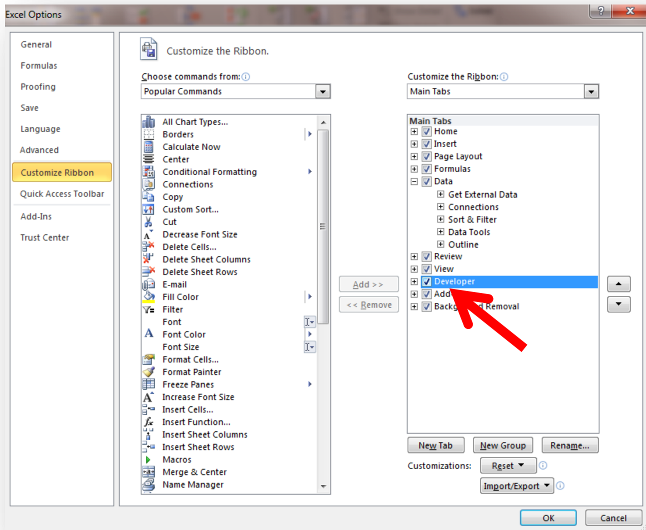
And click on the box next to Solver Add-In…



If you don't see the DEVELOPER tab, then right-click on one the tabs on top (e.g. 'Data'). Select the 'Customize Ribbon'.

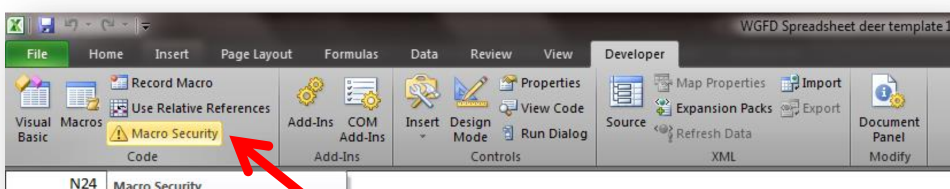


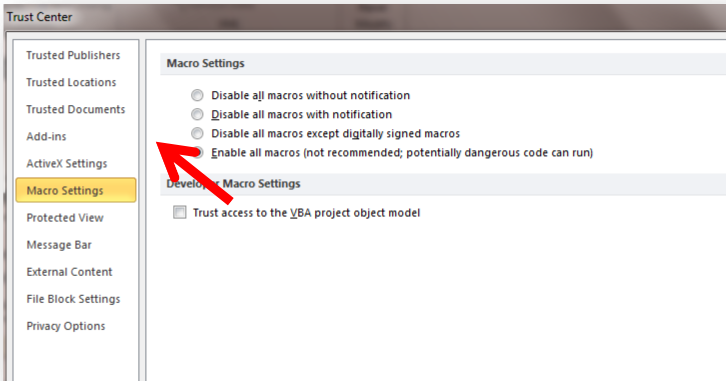
On the right-hand scroll window, click the DEVELOPER box, and click OK.



Solver should now appear in your Data tab.

3. **Enable Macros**. Macros are short programs embedded within Excel to perform repetitive tasks. Often, they are disabled in standard settings in Excel because they can pose security risks and carry viruses if you do not know their sources. We have written several short macros to help compile results at the end of the modeling in the Summary sheet. To use them, you must enable Macros. When you first open the file, Excel will ask at the top of the program whether to enable macros – click 'Enable'. If this does not work, you can manually allow all Macros by going to the Macro Security button on the DEVELOPER tab.





Select the 'Enable all Macros' option. After working in the Spreadsheet model, make sure to change this back to 'Disable all macros' so that other Excel files are not at risk.

4. QUICK AND DIRTY STEPS TO USING SPREADSHEET MODEL

(See Section 7 for detailed instructions)

a. **Data Input.**
   i. Enter data in the DataInput tab. Red columns indicate the raw field data and estimates. Grey columns are derived values. Raw classification counts will be converted automatically into ratios with standard errors. Use the tab button to move through the different data input cells. Cells that do not have field data/estimates should be left blank (rather than entering 0).
   ii. If classification data are not available, enter a reasonable value in the derived ratios columns – we recommend using the average value from at least the previous 5 years. Color these cells Yellow.
   iii. Below the final year of data, add an additional year of classification data to project the population into the future for season setting purposes. Color these cells Blue. We recommend using average values from the previous 5 years.

b. **Setting up models.**
   i. **CJ,CA model:**
      a. Copy-paste-special-values from the generic initial population box into the "Optim Box".
      b. Open Solver; Check that the constraints are appropriate.
      c. Run Solver until model estimates stabilize (usually 2-3 times)
   ii. **SCJ, SCA model**:
      a. Modify juvenile and adult survival optimization cells for any year(s) you think should have time-varying survival estimates. First delete the formula in these cells and input a reasonable starting value. Change the

color of these cells to green to keep track of years that have time-varying survival rates.

- **b.** Count the total number of parameters in the model (i.e., the number of cells that are now in green) and input this number in cell **L4**. This number will be the four parameters in the "Optim Box" plus any juvenile/adult survival rates you allowed to vary by year.
- **c.** Check that the model is not over-parameterized by making sure cell **L5** is greater than 0 (i.e., that DF>0).
- **d.** Copy-paste-special-values from the generic initial population box into the "Optim Box".
- **e.** Open Solver. In the pop-up window, modify the "By changing variable cells" field to add references to any new optimization cells added in step "a". In the "Constraints" window, change the constraints by highlighting each constraint separately and clicking "change". In the new pop-up window, update the references for the constraints in the cells you have allowed to vary by year.
- **f.** Run Solver repeatedly until model estimates stabilize (usually 2-3 times)

iii. **TSJ,CA** and **TSJ,CA,MSC models**: modify

- **a.** Copy-paste-special-values from the generic initial population box into the "Optim Box".
- **b.** Open Solver. In the pop-up window, modify the "By changing variable cells" field to include any new years of data recently added. In the "Constraints" window, change the constraints to include new years of data by highlighting each constraint separately and clicking "change". In the new pop-up window, update the references to include all years of data.
- **c.** Run Solver repeatedly until model estimates stabilize (usually 2-3 times)

- **c.** **Examining results**. Do the estimates seem reasonable? Look at the deviances and penalties – is there one (or a few) year(s) that accounts for most of the deviance/penalties? If so, it suggests the field and model estimates are very different and the model may not be a very good one. Consider either constraining that particular parameter differently or skipping to a more complex model in which that parameter varies independently.
- **d.** **Compare model results.** Examine the model results in the Summary tab. The lowest relative AICc indicates the best model.
- **e.** **Create summary report.** Click on the 'check-box' next to the best model. This will collate the model results from your top model into a report that can be printed.

5. MODEL DESCRIPTIONS AND NOTES

Below we describe the four models that are currently used within the Spreadsheet Model (see tabs '**CJ, CA**', '**SCJ, SCA**', '**TSJ, CA**' and '**TSJ, CA, MSC**' in the Excel file). These models are all exactly the same *except* in how they treat juvenile and adult survival rates. Survival rates can either be constant, semi-constant or time-specific. The 'best' model, and the one that provides the best population estimates, is the model with the lowest AICc score (see Section 4). Additional models could easily be added in the future if there was sufficient need and justification. (***Beware*****:** the second model (**SCJ, SCA**) is the trickiest and requires the most thought on the part of the modeler)

CJ, CA: CONSTANT JUVENILE SURVIVAL AND CONSTANT ADULT SURVIVAL

This is the simplest model, both in terms of structure and in terms of setting it up for analysis. The model assumes juvenile survival is constant (a single value) across all years and male and female survival is the same for both sexes and also constant (though a different value from juvenile survival) across all years. When setting up the model, the user can program it with a range of survival values for juveniles and adults. Within that range, Solver will choose the best (*i.e.,* optimal) survival value for all years for juveniles and another for adults. There are only four parameters in this model: (1) juvenile survival, (2) adult survival, (3) initial male population size and (4) initial female population size. While the assumption of constant juvenile and constant adult female and male survival is biologically unlikely, this model is useful because it uses a small number of parameters and can be applied to sparse datasets.

When should you expect CJ,CA to be the best model? This model likely works best when classification and harvest are the only field data available (i.e., no survival or abundance estimates were collected), when there are only a few years of data available, and/or when the true juvenile and adult survival rates are relatively constant across years. Thus, it is more likely this model will be the best fit (*i.e.,* have lowest AICc score) for sparsely monitored elk than for deer or pronghorn, all else being equal.

SCJ, SCA: SEMI CONSTANT JUVENILE SURVIVAL, SEMI-CONSTANT ADULT SURVIVAL

This model is the most flexible of the 3 models, but it requires the most thought and care on the part of the modeler. This model assumes some (probably most) years have constant juvenile and adult survival rates (one parameter for each constant value), similar to the **CJ, CA** model above. However, **if the modeler has access to either a field estimate of survival from a collaring or mark-recapture study,** *or* **if the biologist has strong reason to believe a particular year produced abnormally high or low true survival rates (e.g. because of a bad weather year), they can set those years to have time-varying survival rates**. 'Time-varying survival rates' implies a particular year will have a unique survival estimate, each of which is estimated by Solver. When using this model it is important for the modeler to document in their JCR narrative why survival in specific years was allowed to be year-specific. A couple of things to note:

- The model will always ‘fit’ better with a greater number of year-specific survival rates. However, your AICc value will be penalized (will increase) when you add each extra year-specific parameter. Thus, the challenge with this model is to set only the years that have the most abnormal survival estimates as the ones that vary.
- You have to be careful not to ‘over-parameterize’ this model -- this occurs when the number of parameters (cells colored in green) exceeds or equals the number of data points ($k \geq n$). When this happens, Degrees of Freedom (DF in cell **L5**) is less than 1.
- Adult survival rates should only be year-specific (allowed to vary) in years juvenile survival is allowed to vary. This is because there is strong biological evidence juvenile survival is more variable than adult survival. However, juvenile survival can vary while adult survival remains constant.
- To allow particular years to have variable survival, you simply add these specific cells to the ‘by changing variable’ field in **Solver** (see section below for full explanation) and color the cells green (to mark them as optimization cells).
- For each survival rate that is year-specific, remember to add constraints in **Solver** (see Section 7)

When should you expect SCJ,SCA to be the best model? In terms of the number of parameters being estimated, this model is of medium complexity. The model likely works best when the modeler has access to a moderate or large number of years of classification and harvest data and when some (but not all) years have abnormally high or low true survival estimates.

TSJ, CA: Time-Specific Juvenile survival, Constant Adult survival

This model is relatively easy to set up, but it tends to have the greatest number of parameters of any model. This model assumes juvenile survival varies every year and holds adult survival constant across all years. This model is biologically realistic on the basis juveniles tend to have much higher variability in survival than adults among ungulates (Gaillard et al. 2000).

When should you expect TSJ,CA to be the best model? This model likely works best when the modeler has access to a moderate or large number of years of classification and harvest data and when juvenile survival estimates vary considerably from year to year but adult survival is relatively constant. This model may be particularly useful for deer and pronghorn, but depending on how much data is available and how variable adult survival is from year to year, it could also apply to elk.

TSJ, CA, MSC: Time-Specific Juvenile survival, Constant Adult survival, Male Survival Coefficient

This model is equivalent to the one above (TSJ,CA), but includes one additional parameter that makes male survival a constant proportion of female survival (the 'male survival coefficient'). This model still assumes juvenile survival varies every year and that adult survival is constant across all years. This model is biologically realistic on the basis juveniles tend to have much higher variability in survival than adults among ungulates, but natural male survival is thought to be lower than female survival.

When should you expect TSJ,CA,MSC to be the best model? This model is most likely in areas where you suspect male survival is different than female survival. In particular, herd units that have high natural predation, such as those in the northwest part of Wyoming, may be most likely to have differences between adult male and adult female survival rates.

6. SELECTING THE BEST MODEL

A. WHAT IS AICC?

Before describing how to set up the models, it is important to understand how we decide which model is best. Generally we're interested in identifying and using the simplest model (*i.e.,* the one with fewest parameters) that simultaneously does an adequate job of fitting the data. To minimize the subjectivity in this decision, we use a statistical property called AICc to compare the relative support for each model. AICc stands for 'corrected Aikaike Information Criterion' (Burnham and Anderson 2002). The spreadsheet automatically calculates an AICc for each model (in cell **L6**). The best model is the one with the ***lowest*** AICc value.

AICc gives you a relative metric of support for each model that balances these 3 components:

(1) model fit (*L* = total summed deviance and penalties)
(2) the number of data points (*n*)
(3) the number of parameters used in the model (*k*)

$$AICc = L + (2k) + \frac{2k(k+1)}{n-k-1}$$

Thus, the AICc value depends on a trade-off between model fit (*L*), the number of parameters being estimated (*k*) and the number of data points (*n*) (also called sample size). Each model has a different AICc score because the fit and the number of parameters are different for each model (note: the number of data points, *n*, is the same across models). Also, *each time you add a new year of data, AICc values will change*. This may change which model is best, so you have to check AICc scores each year new data is added.

B. MODEL FIT (*L*)

As the name suggests, model fit is a measure of how well the model fits the data. In practice this is calculated as the difference between your field estimates and your model estimates. Indeed, the whole goal of the model boils down to finding the model estimates that minimize the difference between field estimates and model estimates. This difference is the sum of 'Penalties' and 'Deviance' (cell **I5**).

C. DATA POINTS (*N*)

The number of data points, *n*, is calculated automatically in the spreadsheet model; however, you should have some idea where this number comes from. *n* is calculated as:

n = No. of years of Male:Female ratio field estimates (column **E**)
+ No. of Juvenile survival field estimates (column **J**)
+ No. of adult survival field estimates (column **N**)
+ No. of years of Harvest data (column **U**)
+ No. of years of Population size estimates (column **V**)

D. MODEL PARAMETERS (*K*)

The number of parameters is the number of independent values you are trying to estimate. These should be colored in green in each model. Each of the three models uses a different number of model parameters. Indeed, the models are mainly distinguished by whether or not the juvenile and adult survival parameters vary over time. The number of parameters is calculated automatically, ***except*** in the **SCJ, SCA** model. For this model, the modeler must calculate the number of parameters because they are the ones who determine which years should vary independently and which years should stay the same.

K = No parameters being optimized (cells in green)

CJ, CA model: ***k*** = 4 (cells **Q3:Q6**)
SCJ, SCA model: ***k*** = 4 + sum of all juvenile (Col **J**) or adult (Col **N**) survival cells that vary by year
TSJ, CA model: ***k*** = 3 + sum of all juvenile survival cells (Col **J**)

Here is an example of a simple Spreadsheet Model (from the **SCJ, SCA** model) showing how *n*, *k* and *L* are calculated.

Figure 1. *The number of data points ('n', red arrow at top, columns shaded in pink) is 29 because there are 11 years of Male/Female data, 2 years of winter juvenile survival rate field estimates, 2 years of adult survival rate field estimates, 11 years of harvest data and 3 years of Post-hunt population field estimates. The number of parameters being estimated ('k', green arrow, columns shaded in green) is 8 because there are 4 parameters on top, 2 time-varying winter juvenile survival model estimates and 2 time-varying adult survival model estimates.*

A warning about over-parameterization: Simple is usually better.

With the more complex models (**SCJ, SCA** and **TSJ, CA**), you have to be careful not to over-parameterize the model, which occurs when the number of parameters exceeds or equals the number of data points ($k \geq n$). In other words, you run out of 'degrees of freedom' which is invalid statistically. Degrees of freedom are basically the difference between the number of parameters being estimated (k) and the number of data points (n). As an example, suppose you collected 50 years of harvest data for a particular population and wanted to use these to estimate 50 survival rates. You would have zero degrees of freedom because n=50 and k=50 and you'd have no way of estimating all of these survival rates without additional information. Having a large number of degrees of freedom (*i.e.,* 30) as opposed to small number (*i.e.,* 3) will always benefit your AICc score. However, the only way to have more degrees of freedom for a given number of data points is to estimate fewer parameters. When you estimate fewer parameters, your model fit (L) will likely be worse. The upshot is that your AICc score may be better or worse with more degrees of freedom. The only hard and fast rule is your degrees of freedom cannot exceed the number of things being estimated (n-k > 0).

7. DETAILED INSTRUCTIONS FOR USING SPREADSHEET MODEL

DATA INPUT

In the "DataInput" tab, enter new or existing data. Field data are entered in columns highlighted in red and include: Classification counts (required), Harvest data (required), Population Objective, Survival data and SE (if available), Abundance or Line-transect estimates and SE (if available) and Trend counts (if available). These input cells were set up to link to the three model spreadsheets in the other tabs so you do not need to enter data multiple times. Trend count data are not used in the model, but they may be useful for adjusting the model as a measure of minimum abundance.

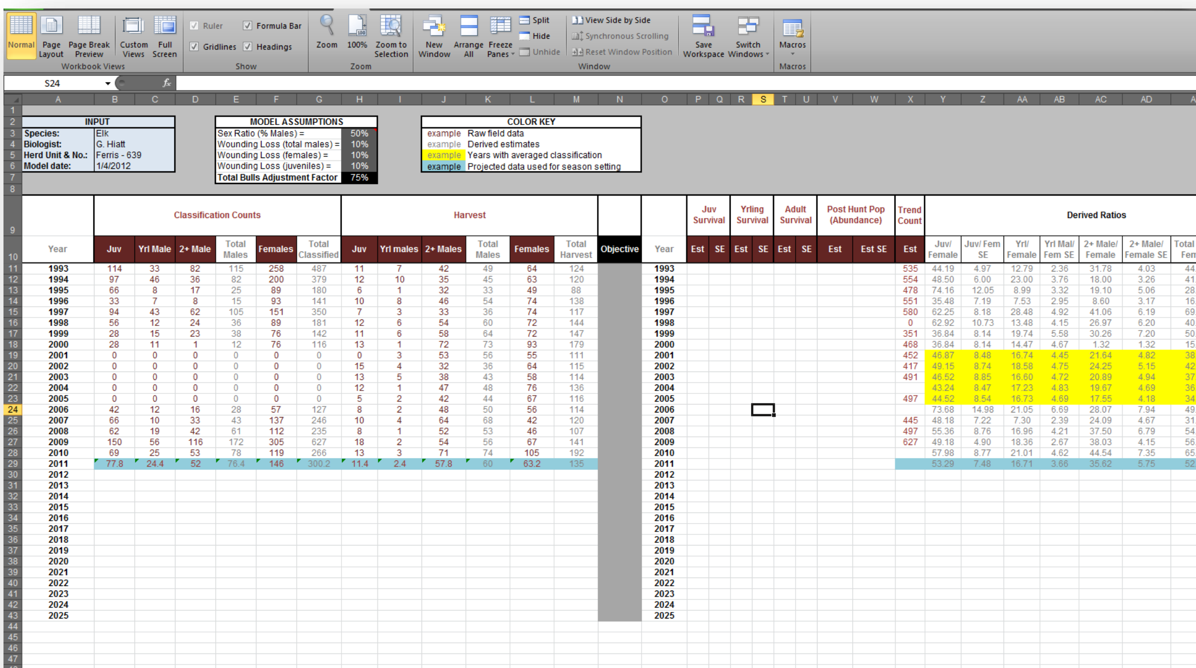


Figure 2. *DataInput tab. Note the red columns are raw field data. Yellow cells are average data for years without classification counts. Blue cells references projected data in the upcoming year used for season setting purposes.*

<u>Three Important Notes about Data Entry:</u>

1. **Cells that lack field data should be left blank (rather than entering '0').** Entering 0 will cause ratio formulas to divide by 0, which generates the **"#DIV/0"** error.

2. **When first entering data, always start your initial year of data on the existing top row (row 3), then modify Years in Column A. <u>Never delete rows of data</u>.** The DataInupt tab is linked to the model pages and when you delete rows of data it will cause references issues for these model tabs. When entering field data for a new herd unit, it is best to simply delete old field data (type in red) on a cell-by-cell basis, then correct the year (Column A) to correspond to your initial year of data. Finally, enter (or copy and paste) your field data, leaving the cells between your final year of data and the last year in the datasheet (likely 2025) empty.
3. **The DataInput sheet is protected so users can only change cells with field data (in red) or derived ratios in grey.** While not recommended, you can turn this protection off by clicking on the 'Review' tab along the top of Excel and selecting 'Unprotect Sheet'. If you try modifying protected cells, you will get an error message.

<u>Steps and notes for Data Input</u>

- First, in the blue 'Input' box type the species, herd unit, biologist and date of the model. This box is linked to the other sheets so you only have to enter this information here.
- For classification data, enter the raw count data collected during classification counts. The ratios and SE's will be calculated automatically in the derived ratio columns (columns Y through AF). If you have calculated the ratios and prefer to enter these instead of the raw count data, you can directly enter these in the Derived Ratios columns after unprotecting the datasheet (see instructions above). SE's must be greater than 0.
- If no classification data were collected in a particular year, you must still enter a reasonable value for this year in the 'Derived Ratio' columns. We recommend taking the average of at least the previous 5 years of ratio and SE estimates from other years. As an example, if we did not manage to collect 'Juv/Female' classification data in year 2002, we would (i) unprotect the worksheet, (ii) select the 'Juv/Female' cell for 2002, and (iii) delete the existing formula and replace with the formula '=AVERAGE(Y7:Y11)'. This takes the average Juv/Female ratios between 1997-2001 (Y7:Y11). Continue averaging any additional cells from that year (SE, Total Male/Female, etc). **Any cells/rows that have average values should be highlighted in YELLOW so these years can be easily identified in subsequent years.**
- SE for classification data are calculated automatically in columns Z, AB, AD and AF of the DataInput Tab. The formula uses the individual as the sampling unit and assumes sampling occurs with replacement (Czaplewski 1983; page 127, equation 10).
- Enter harvest, survival and abundance (or line-transect) estimates in their respective columns.
- Modify the Total Bulls Adjustment Factor (for elk), the Over-summer adult survival rate (pronghorn) or sightability factor (for deer) in cell **H7**, if necessary.
- Below the final year of data, add an additional year of classification data to project the population into the future for season setting purposes. Color these cells Blue. We recommend using average values from the previous 5 years.

Where can you get field estimates of survival? Including a few years of survival estimates can greatly improve your model. However, these estimates are often expensive to collect – typically you need about 100 collared animals to get a reasonable estimate or survival. Fortunately, there is reasonable evidence survival rates are correlated across herd units. So, if there are field estimates of survival from an adjacent herd, even within a nearby state, adding this data to your model can greatly improve it. If there have been survival studies in an adjacent herd we recommend adding the data to your model. We hope to conduct future studies on the topic of correlated survival rates to help improve this recommendation for herds not adjacent to one another.

SETTING UP MODELS:

- The new data you just entered should show up in the four model sheets in the row for the appropriate year. Check that the data appears correctly for each year.
- In the SCJ,SCA, TSJ,CA and TSJ,CA,MSC models, any new years of data allowed to vary should be updated so they are colored ***green***. In the CJ,CA model, the only optimization cells are in the 'Parameters' box at the top of the spreadsheet, so will not change.
- In the TSJ,CA and TSJ,CA,MSC models, you have to type in starting values for juvenile survival for each new year of data you have added – again, these should be highlighted green. The other two models have formulas that will fill in values automatically.
- **Check that all green optimization cells have reasonable starting values before you run Solver.** If they do not, type or copy in reasonable values. Values will change when you run Solver, but it helps to have starting values somewhat close to the real values. The Generic Initial Population box (V3-V6) calculates rough guesses of the initial starting values. These values are not used directly in the model; rather, they give you reasonable starting values that can be copied into the green optimization cells (Q3-Q6) when first running a model.

NOTE: When copying the starting values from the Generic Initial Population box to the Optim box, make sure to **'Paste Special: *values-only*'** rather than simply 'pasting'. If you 'Paste', you will copy the formulas from the Generic Initial Population box rather than the values. To 'Paste Special: *values-only*', either right click when you are pasting and select 'Paste Special' or use the Paste Special button at the top of Excel.

Initial Adult male (**Q5**) and Initial Female (**Q6**) cells are the approximate population sizes of Adult males and Females during the *first year* of data, divided by 10,000. This division helps the model run more quickly and efficiently. If these values are unknown, just make your best guesses. (See Section 11, *Cell Definitions*, for explanation of these formulas).

- In the Model Selection box at the top of the spreadsheet, Excel will automatically calculate the number of parameters being estimated (*k*, cell L4) for the CJ, CA and TSJ, CA models and will automatically calculate the sample size (*n*, cell L3) for all models. **Importantly, for the SCJ,SCA model you must manually calculate the number of parameters being estimated**. This is the number of green optimization cells in the model, which equals 4 (for cells Q3:Q6) *plus* any juvenile or adult survival cells you allow to vary in Columns I and M. Again, the number of parameters is calculated as:

k = No. of green optimization cells in a model

CJ, CA model: ***k*** = 4 (cells **Q3:Q6**)
SCJ, SCA model: ***k*** = 4 + sum of all juvenile (Col **J**) or adult (Col **N**) survival cells that vary by year
TSJ, CA model: ***k*** = 3 + sum of all juvenile survival cells (Col **J**)

Sample size (also again) is the same in all models and is automatically calculated as:

n = No. of years of Male:Female ratio field estimates (col **E**)
+ No. of Juvenile survival field estimates (col **J**)
+ No. of adult survival field estimates (col **N**)
+ No. of years of Harvest data (col **U**)
+ No. of years of Population size estimates (col **V**)

- Update the Overall, 5-year and 3-year averages in rows 44-46 at the bottom of the datasheet to include the most recent years of data by simply changing the formula in cells **B44-46** and then copying this group of cells to the rest of those rows. These data are reported in the Summary Tab to help summary the model results.

RUNNING MODELS:

- Click on Solver. Solver is a free add-in program used to find optimal values, given a certain model structure and input data. Solver should be available in all recent MS Excel software (Excel 2003 or later). In Excel 2010, Solver is under the DATA tab at the top of Excel spreadsheets. If it does not appear, you may have to go to the Add-In pop-up window (under the DEVELOPER Tab at the top of Excel 2010) and check the box next to Solver. We have noticed that certain security settings will cause Solver to disappear from the top ribbon each time you close Excel. At the moment, we haven't figured out how to avoid this annoyance.

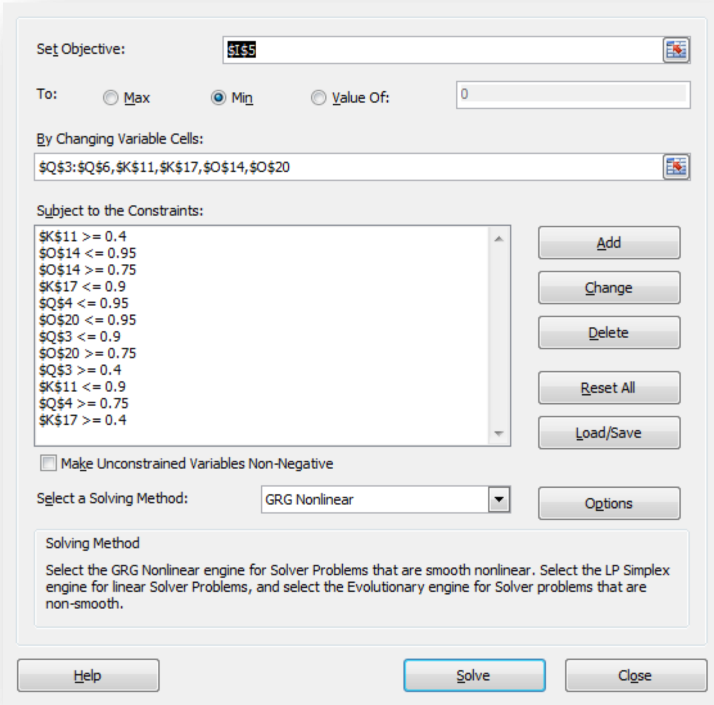


Figure 3. *Solver window, from a relatively complicated SCJ,SCA model. Other models will typically use only a few constraints.*

- Open Solver. **Solver is probably the most likely place to make a mistake when first learning how to use the Spreadsheet Model.** The Solver user-interface is not especially user-friendly and it behaves differently than other things in Excel. When running new models (except in the CJ,CA model), you will have to update the reference cells in two components of Solver: the “Changing variable cells” field and the “Subject to the Constraints” box. When changing these fields, **take care not to use the arrow keys to select reference cells**—this will inevitably cause you to reference the wrong cells. A much better practice is to always use the mouse to select the reference cells.
- The top field, called “Set Objective” or “Set Target Cell” (depending on your version of MS Office), should automatically reference the Total Model penalties and Deviances (highlighted in yellow in cell I5). **<u>Do not change this field</u>.** This is the cell Solver tries to minimize (make sure the ‘Min’ radio button is selected). This field should say: I5. In Solver it is very easy to have the cursor in the “Set Objective” box and accidently click around a cell in the spreadsheet. If you do this, the reference cell in this box will change, so be very careful to check to make sure this references the right cell before running Solver.
- “Changing variable cells” should reference the green optimization cells in the spreadsheet. If you’ve added new data in the spreadsheet that varies by year, you need to add these cells to this field. Which cells can vary depends on the model (in fact this is the crux of the differences between the models; see below for instructions).

- The “Subject to the Constraints” field refers to reasonable constraints on the survival optimization cells. For example, for mule deer we would expect juvenile survival to be between 0.4 and 0.9 in most years. The constraints box allows us to constrain the model so the model estimated survival rate is between these values. Each model will have a different number of optimization cells and so each may have different references for constraints.

What are “reasonable” survival constraints? Constraints can have a large impact on the population estimates, so they should be chosen with care. Constraints should ***not*** be used as a way to tweak the model to get the results you expect. This is bad practice. Species and herd units will likely have different constraints. For example, deer survival is generally more variable than elk, so will have a larger range in the constraints. Likewise, certain herd units may have more variable survival than others – for example, herds subjected to grizzly predation may have lower survival than those that do not – so constraints may differ by herd units for a given species. For the time being, we recommend using the “Standard Ranges of Parameters” sheet as a guide for constraints. We hope to make these recommendations more precise in future versions of the User’s Guide.

SOLVER EXAMPLE

In the SCJ, SCA model: “Changing variable cells” and “Subject to Constraints” boxes should reference the cells in the Parameters box (cells Q3:Q6) as well as any survival rates allowed to vary by year (Columns I and M). Before running Solver, **any cell allowed to vary by year needs to be added to the “Change variable cells” and to the “Subject to Constraints” boxes.** For example, if the SCJ, SCA datasheet looks like this:

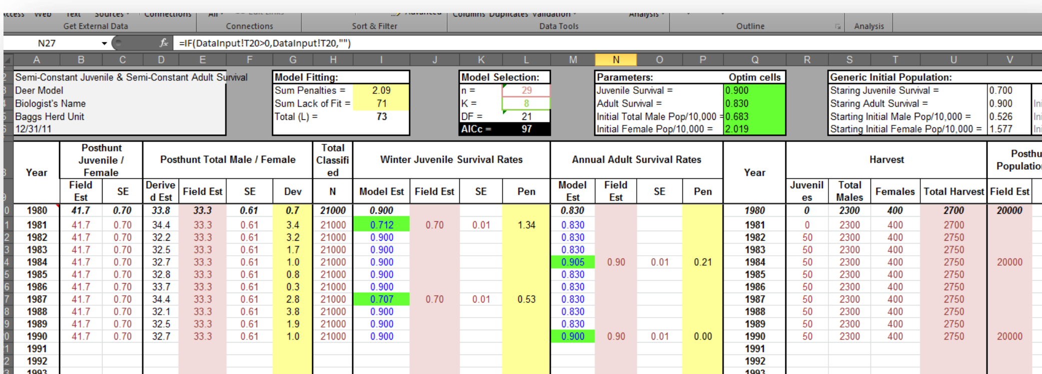


Figure 4. Example of the Spreadsheet Model.

…then there are 8 green optimization cells (**Q3:Q6**, **I11**, **I17**, **M14**, **M20**). Juvenile survive is allowed to vary in 2 years (here, cells **I11** and **I17**) and adult survival is allowed to vary in 2 years (here, cell **M14** and **M20**). In the Solver pop-up window, we would want to make sure all 8 cells are referenced in the “Change variable cells” and to the “Subject to Constraints” boxes.

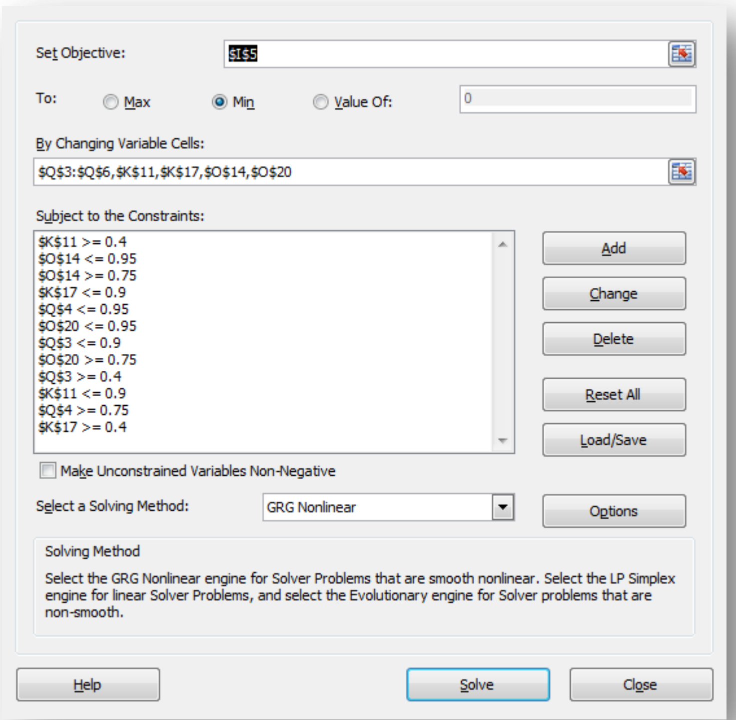
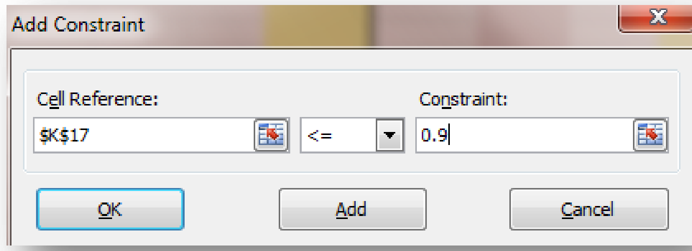


Figure 5. Solver window (same as Figure 2).

To add new references in the “By Changing Variable Cells” box, click on the selection icon at the right of the box and select the reference cells in the new pop up window. Each new reference string needs to be separated by a comma.

To add new references in the “Subject to the Constraints” box, click the ‘Add’ button to the right of the box. A new pop-up window appears in which you can select any new cell references. For example, if you want cell **I17** (a juvenile survival estimate) to be less than or equal to 0.9, click on ‘Add’ and in the new pop up window, click on the cell reference icon, and select **I17**. Choose the ‘less-than-or-equal-to’ sign in the middle drop down menu and type in 0.9 in the constraint field. Click ‘OK’.



Unfortunately, you have to add each constraint for each reference cell (or group of adjacent cells), so adding constraints can be somewhat time-consuming. Double check when you're finished that all constraints and signs are correct, particularly if you're doing this for the first time.

In the CJ, CA model, the "Changing variable cells" and "Subject to the Constraints" boxes should simply reference the 4 green optimization cells at the top of the spreadsheet in the Parameters box (Q3:Q6).

In the TSJ, CA model, "Changing variable cells" field should reference the cells in the Parameters box as well as all model estimates of Juvenile Survival (all years of Col K). To add new years of data in the Solver popup window, simply click on "Change" (instead of 'Add') and update the reference cells to include any new year of data.

In general, we suggest the CJ,CA model should have tighter constraints than the other models because this model assumes juvenile and adult survival is the same every year, so estimates should be near average. Other models (SCJ,SCA TSJ,CA and TSJ,CA,MSC) have time-specific estimates – because they vary from year to year, they could encompass a larger range of values across years.

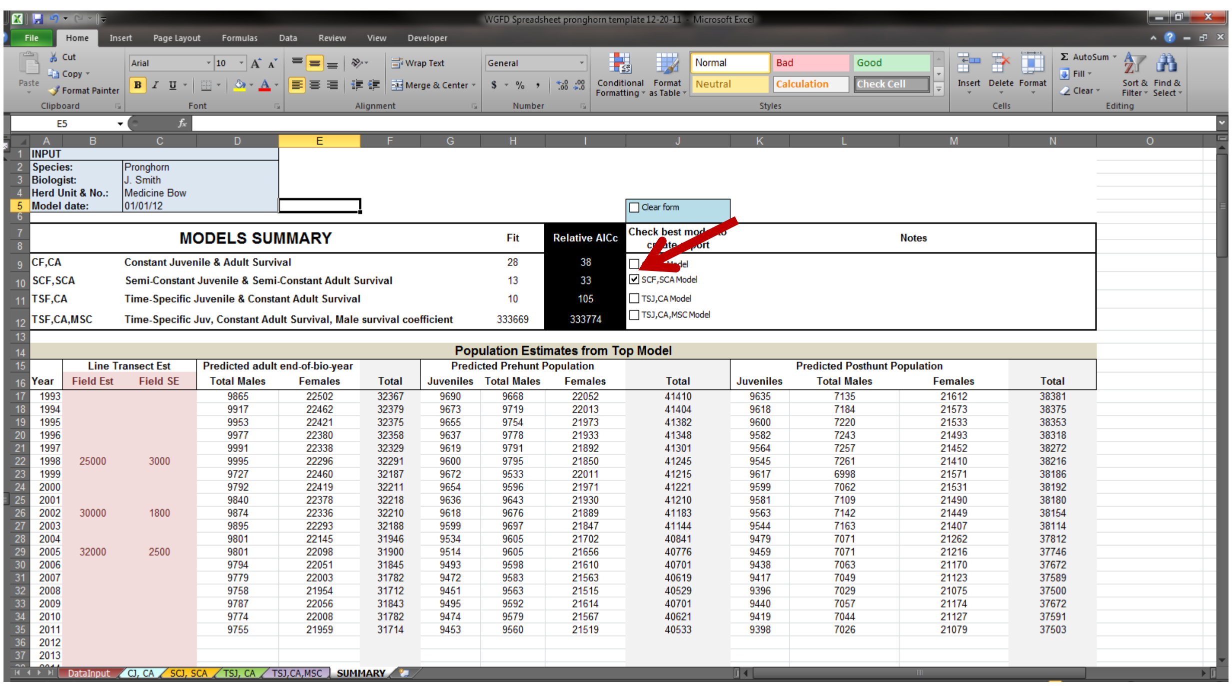
- Click on "Solve" at the bottom of the pop-up window. In a second or two, it should change the values in blue and black in the spreadsheet. Solver will not always find 'the best' values on its first try, so run Solver several more times until the values do not change.
- Look at the deviances and penalties – does one (or a few) years account for most of the deviance/penalties? If so, it suggests the field and model estimates are very different for that year(s). If this happens there are several options. First, you can consider either constraining that particular parameter differently or moving onto to a more complex model in which that particular parameter varies independently.
- Look at the new estimated values. If the Predicted Posthunt Population Size and the Juvenile and Adult survival model estimates seem reasonable, go to the next model and repeat the steps from the previous sections. If the new values do not seem reasonable…. something went wrong (see *Troubleshooting* Section 8, for suggestions).

COMPARING MODEL RESULTS (SUMMARY TAB):

- If your population estimates from the three models are drastically different, it suggests you shouldn't have much confidence in the models or in their estimates. Only compare the models you feel are reasonable and provide relatively consistent population estimates.
- Once you're satisfied your models given reasonable results, go to the Summary tab. The 'best' model is the one with the lowest Relative AICc value. AIC values are compiled in the Summary tab in the Model Results box. AICc is a measure of parsimony which basically says "how likely are your estimated model values, given the sample size (*n*) and the number of parameters in the model (*k*)?"
- If you are not confident in the population estimates from the 'best' model, consider modifying some of your parameter constraints and running Solver again on this model or all models. In the SCJ, SCA model you can also modify which years you allow to vary and which ones you do not.

CREATING SUMMARY REPORT:

Click on the 'check-box' to the right of the model with the lowest AICc value (**Fig 5**). This will populate the summary page below with the results from the best model. These Summary pages can be printed so all results fit within several landscape-oriented pages. The check boxes use a custom macro to copy and paste all results from the selected model into the report form on the Summary page. Note you must allow macros to be enabled to use this function. Above the model check boxes is a check box that clears all model results from the report form. Click this box to clear out previous results.



INPUT	
Species:	Pronghorn
Biologist:	J. Smith
Herd Unit & No.:	Medicine Bow
Model date:	01/01/12

Clear form

MODELS SUMMARY		Fit	Relative AICc	Check best model to create report	Notes
CF,CA	Constant Juvenile & Adult Survival	28	38	CJ,CA Model	
SCF,SCA	Semi-Constant Juvenile & Semi-Constant Adult Survival	13	33	SCF,SCA Model	
TSF,CA	Time-Specific Juvenile & Constant Adult Survival	10	105	TSJ,CA Model	
TSF,CA,MSC	Time-Specific Juv, Constant Adult Survival, Male survival coefficient	333669	333774	TSJ,CA,MSC Model	

Population Estimates from Top Model

	Line Transect Est		Predicted adult end-of-bio-year			Predicted Prehunt Population				Predicted Posthunt Population			
Year	Field Est	Field SE	Total Males	Females	Total	Juveniles	Total Males	Females	Total	Juveniles	Total Males	Females	Total
1993			9865	22502	32367	9690	9668	22052	41410	9635	7135	21612	38381
1994			9917	22462	32379	9673	9719	22013	41404	9618	7184	21573	38375
1995			9953	22421	32375	9655	9754	21973	41382	9600	7220	21533	38353
1996			9977	22380	32358	9637	9778	21933	41348	9582	7243	21493	38318
1997			9991	22338	32329	9619	9791	21892	41301	9564	7257	21452	38272
1998	25000	3000	9995	22296	32291	9600	9795	21850	41245	9545	7261	21410	38216
1999			9727	22460	32187	9672	9533	22011	41215	9617	6998	21571	38186
2000			9792	22419	32211	9654	9596	21971	41221	9599	7062	21531	38192
2001			9840	22378	32218	9636	9643	21930	41210	9581	7109	21490	38180
2002	30000	1800	9874	22336	32210	9618	9676	21889	41183	9563	7142	21449	38154
2003			9895	22293	32188	9599	9697	21847	41144	9544	7163	21407	38114
2004			9801	22145	31946	9534	9605	21702	40841	9479	7071	21262	37812
2005	32000	2500	9801	22098	31900	9514	9605	21656	40776	9459	7071	21216	37746
2006			9794	22051	31845	9493	9598	21610	40701	9438	7063	21170	37672
2007			9779	22003	31782	9472	9583	21563	40619	9417	7049	21123	37589
2008			9758	21954	31712	9451	9563	21515	40529	9396	7029	21075	37500
2009			9787	22056	31843	9495	9592	21614	40701	9440	7057	21174	37672
2010			9774	22008	31782	9474	9579	21567	40621	9419	7044	21127	37591
2011			9755	21959	31714	9453	9560	21519	40533	9398	7026	21079	37503
2012													
2013													

Figure 6. *Example of how to collate data from the best model into a report. Simply click on the check box beside the top model (lowest AICc).*

MODELING INTO THE FUTURE

A. Model 2-years into the future
   - Year 1
      - Expected harvest based on the number of licenses being issued and past harvest history
      - 5-year average calf/cow ratio and bull/cow ratio
   - Year 2
      - Average population segment harvest rates
      - Average sex and age ratio data

B. If the population is rapidly increasing or decreasing, modeling only one year into the future is recommended.

GENERAL TIPS ABOUT USING THE SPREADSHEET MODEL:

- Models tend to produce the best estimates when harvest rates are high. This is because harvest is assumed to be a 'known' value. If harvest accounts for a substantial portion of the mortality in the population (relative to adult survival), there is less uncertainty in the model estimates of survival and male:female ratios.
- After particularly catastrophic years (e.g., years with abnormally high overwinter mortality), you may need to reset the model and only use a few of the most recent years of data, based on suggestions from the Colorado Parks and Wildlife folks. If you reset a model, please justify this decision in your JCR narrative.
- We do not recommend moving, deleting or inserting cells, columns or rows. Changes may produce reference errors and may cause issues with the Macro used in creating the report, as the Macro only copies specific cells. However, if changes are necessary, new rows or years of data should be added from the bottom-most row, rather than middle or top row. If new years of data are added in the DataInput page, you may have to modify each model page separately by copying and pasting the formulas from preceding rows. Again, the top (in bold italics) row of data should never be changed.

8. TROUBLE-SHOOTING

Model estimates do not seem reasonable. While the Spreadsheet Model is relatively inflexible, there are a few ways to adjust the model to produce different estimates. These options should only be used if you can justify why particular values should be different than those produced by the model. The risk in trying to 'tweak' the model is unreasonable estimates may simply be symptomatic of bad (or inadequate) data, so tweaking the model may mislead you into thinking the model is a good one, when in truth the model should be excluded from the list of candidate models. Nonetheless, there are several ways to tweak the model:

1. Check that the model is set up correctly. In particular, check that Solver is optimizing the correct cells (everything in green), that Solver is minimizing cell I5 and the constraints are referencing the correct cells (again, everything in green). When new data is added, it is easy to forget to modify Solver to include the new years' of data (particularly in the **TSJ, CA** and **TSJ, CA, MSC** models).
2. Try changing the initial starting values in the green optimization boxes (cells **Q3:Q6**) or the initial year-specific survival values in the **SCJ, SCA**, **TSJ, CA** and **TSJ, CA, MSC** models and running Solver again.
3. If the population estimates are negative, your initial Total Male and Female population sizes (**Q5:Q6**) may be starting too low. We have noticed the "Generic Initial Population" values for males and females are not always very good and the model can be pretty sensitive to these values, especially when they are too low. Try increasing the Initial Total Male and Female population sizes by an order of magnitude (e.g. from 0.1 to 1.0) and running Solver again. After running Solver, look at the resulting optimized estimate. If it the estimate is still very low, try adding a constraint that forces the initial population values to be above some lower limit – for example, try adding the constraint Q5:Q6 >= 0.02, meaning the male and female population sizes in the first year of the model must be above 200 individuals each.
4. Try modifying other constraints in the Solver pop-up window. For example, if you know one year had a particularly severe winter, you might constraint juvenile survival in the **TSJ,CA** model to be lower than other years, perhaps constraining the potential estimates between 0.65 and 0.25. This change may be enough to help the model fit better.
5. Look at the penalties and deviances. Is there a particular year with particularly high deviance or penalty? If so, it suggests a poor model fit. This problem occurs most often in models where you have field estimates of survival but the model estimates are being held constant (for example, if you have field estimates of juvenile survival but are trying to use the **CJ,CA** model, it will likely produce a single estimate across all years). Most likely, you will have to use the **SCJ,SCA** model rather than the **CJ,CA** model.
6. In catastrophic weather years (e.g. extremely severe winters), you may be justified in fixing a model estimated juvenile survival rate to be a specific value. In other words, you prevent the model from optimizing on it, and make the BIG assumption you know the value exactly. This may be necessary to get the model to fit properly, particularly if the

other years are relatively invariable in comparison. This procedure is only justified in models with constant juvenile survival rates (**CJ,CA** and **SCJ,SCA**) where the particular year is likely much lower than other years. For example, in the **CJ,CA** model if the model estimates suggest juvenile survival is 0.7 across all years, but you think one bad weather year should have much lower survival (say, 0.35), then you would delete the reference to Q3 (Juv survival) for that year and replace it with 0.35. When you use Solver again, the model will ignore that year and may push the other survival estimates higher.

7. If none of these solutions seem to change the results, you may simply be stuck with bad data, a bad model (particularly **CJ, CA**), or both, in which case you should discard this model from your list of potential candidate models.

Population estimates are negative. Obviously something isn't right... see above for potential solutions. If no solutions work, you may have to discard this model from your list of potential candidate models.

Model estimates equal the constraint values. If one or more Model Estimates fall at the upper or lower constraints, it suggests the estimates are not particularly good. In our experience, in the time-varying models you will likely have some years hit parameter boundaries, which is okay. If this occurs in the 'constant' models, if suggests you should not trust the model. For example, if juvenile survival is 0.4 every year in the **CJ,CA** model, this is biologically unrealistic. Try changing the constraint in the Solver pop-up window (lowering or raising it by 0.05 or 0.1 at a time) and running Solver again. If this process continues to produce estimates on the parameter constraint boundaries, you will likely have to disregard the model from the list of potential candidate models.

"#DIV/0" error. This means the equation in the cell is being forced to divide by 0, which is not allowed mathematically. This will screw up other values in the spreadsheet and will prevent Solver from working properly. This problem needs to be fixed before you can proceed. Click on the cell with the error and try to determine which other cell it is using in the denominator of the calculation.

- One of the most common causes of this problem is when a field estimate has a SE value of 0.0. Ratio data, survival estimates and population estimates all need to have SE's greater than 0.0.
- If a SE estimate is not available, use an average value from past years or (for ratio data) simply input 1.0.
- If the error occurs in the row following the final year in which field data were collected, there is probably a problem with the formulas. The best solution to this problem is to copy the entire row of data from the previous year and paste onto the row with the error – this should restore the formulas to have their correct references.
- If nothing else works, try copying all rows from the original template file and pasting them into the tab you're working in. This should restore all formulas.

<u>"#Value" error.</u> This means the formula in the cell is missing a reference to another cell. This error occurs most often after you have deleted something the cell containing the error uses in its formula. Often, if one cell has this error, many cells will have the error.

- If you have recently deleted something, use the 'Undo' option and see if the error goes away.
- If not, look at another model to see if it also has the same errors – if it does, it suggest there is something wrong in the DataInput tab, such as years with missing classification data (do all Derived Ratios have values?).
- If the other models do *not* have similar errors, it means the initial model that you were using contains the source of the problem.
- If the error occurs in the row following the final year in which field data were collected, there is probably a problem with the formulas. The best solution to this problem is to copy the entire row of data from the previous year and paste onto the row with the error – this should restore the formulas to have their correct references.
- If nothing else works, try copying all rows from the original template file and pasting them into the tab you're working in. This should restore all formulas.

9. MODEL INPUT DATA – RANGE OF ACCEPTABLE VALUES

DATA INPUT PAGE

A. Sex Ratio at Birth
   - 50:50 for all species
   - Never use sex ratio at birth to improve fit

B. Wounding Loss (as a percentage of harvest)
   - Elk (based on work by Freddy in Colorado – Grand Mesa and Gunnison)
     - Calves – 10-15%
     - Bulls – 10-15%
     - Cows – 15-25%
   - Deer and Pronghorn
     - 10% for all
   - Never use wounding loss as an optimization variable

C. Pronghorn Oversummer Survival
   - Colorado assumes 1.0, but given some individuals likely die, starting value should be 0.98. If you believe over-summer survival to be lower, do not adjust lower than 0.95 without supplemental survival data.

D. Total Bull Adjustment Factor (only used in elk models)
   - This bias factor accounts for differences in sightability between cows and bulls during composition surveys. The bias factor adjusts the posthunt bull:cow ratio in the calculation of bull:cow deviance, i.e. [(((observed bull-cow ratio/bias factor) – predicted bull-cow ratio)/SE)^2]. **<u>A range of bias factors from 60 – 100 should be evaluated by trial and error, with a starting value of 75%</u>** (i.e. 75% of males are seen relative to females). Note that in heavily forested areas where sightability is poor (e.g. Eastern Washington State), male bias factors can be extremely low, ~35% (McCorquodale 2001). The user should keep in mind bias is less for yearling bulls than branch antlered due to adult segregation during the classification period. Thus, the population's age structure (if know) should influence the choice of the bias factor: if there are many yearlings and few adult bulls, the bias factor should be nearer 100%. Generally, changing bias factors has less effect on population estimates than on predicted sex ratios and survival rates (e.g., lower bias factors usually increased predicted sex ratios and decreased calf survival rates). **<u>However, the bias factor can strongly influence population estimates and it should therefore not be used to force population estimates towards a particular value.</u>**

MODEL SURVIVAL CONSTANTS

A. Elk
   - Elk Starting Values in Parameter Box (again, based on data collected by Freddy in CO)
     i. Juveniles – 0.86

ii. Adult - 0.96
iii. Male survival coefficient – 1.0 (less than 1.0 means $Surv_{Male} < Surv_{Female}$)

- Elk Range of Values in Solver Box
  i. Juvenile – 0.5 – 0.95, in general these values should range between 0.80 and 0.95 unless a particularly severe event requires a lower value.
  ii. Adult – 0.80 – 0.98, in general these values should range between 0.85 and 0.98
  iii. Elk herds with a full suite of large carnivores (grizzly bears, black bears, mountain lions and wolves) in NW WY may need to use a lower set of values. These values should be justified based on the results of collar data.

B. Deer and Pronghorn

- Deer and Pronghorn Starting Values in Parameter Box (based on 8-12 years of data from 5 study areas in CO)
  i. Juveniles – 0.65
  ii. Adults – 0.80
- Deer and Pronghorn Range of Values in Solver Box
  i. Adults – 0.7 – 0.95
  ii. Juvenile – 0.4 – 0.90
- Individual years may be constrained below the minimums, particularly juvenile survival, if high mortality is suspected

10. MODEL EVALUATION CRITERIA

A. Each model should be evaluated using the following criteria.

B. General

- Excellent – Model has 15-20 years of data; ratio data available for all years; juvenile and adult survival estimates with standard errors available at least 5 out of 10 years; at least two sample-based population estimates with standard error available; model aligns well; results biologically defensible.
- Good – Model has 15-20 years of data; ratio data available for all years in model; either juvenile and adult survival estimate with standard errors available at least 2 out of 10 years and at least one sample-based population estimate with standard error available OR these data obtained from adjacent or other similar herds; model aligns fairly well; results biologically defensible.
- Fair – Model has 15-20 years of data; ratio data available for all years in model; juvenile and adult survival estimates, if any, are obtained from adjacent or similar herds; model fit fair; results biologically defensible.
- Poor or No Model – Model may be ranked Poor or have No Model for a variety of reasons including: little data available; ratio data, if available, considered highly biased because of poor sample sizes or an inability to survey the entire

area; herd unit closure issues apparent; model does not run; results not biologically defensible.

C. Elk – Additional Thoughts
   - Using AIC values to pick best model is not recommended in herds with no sample-based population estimates or survival estimates because of inadequate observed data.
   - Look for low adjusted observed vs predicted bull:cow ratio deviance, particularly in recent years.
   - The population curve shows a relationship to male harvest, especially in General License areas (based on CO data, model based post-hunt population estimate between 10 – 12 times male harvest).
   - Survival rates are realistic based on available data.

D. Deer and Pronghorn Additional Thoughts
   - Look for low relative AIC value for comparable, optimized models with sample-based population estimates and survival estimates or in herds where these data are derived from adjacent or similar herds.
   - Look for low adjusted observed vs predicted buck:doe ratio deviance, particularly in recent years. Predicted buck:doe ratios that increasingly diverge higher or lower than observed ratios indicate the population is being overestimated or underestimated, respectively.
   - The population curve shows a relationship to male harvest, especially in General License areas (based on CO data, model based post-hunt population estimate between 10 – 12 times male harvest).
   - Survival rates are realistic based on available data.

11. DIFFERENCES BETWEEN ELK, DEER AND PRONGHORN MODELS

The elk, deer and pronghorn were purposefully created to be nearly identical in form and function. However, because of differences in how monitoring data are collected and in visibility differences among species, each model is slightly different.

Elk Model. The elk model includes a value called the "Total Bulls Adjustment Factor" which adjusts the Total-Male/100 Females (i.e. the bull:cow ratio) to account for differences in detectability between bulls (who often stay in small groups in wooded areas) and cows (who often stay in larger groups out in the open) during classification counts. The TBAF value is entered or modified at the top of the DataInput page but links to the Total-Male/100 Female deviance calculation in each of the model pages.

INPUT	
	Elk
	J. Smith
.:	Targhee 101
	1/1/2012

MODEL ASSUMPTIONS	
Sex Ratio (% Males) =	50%
Wounding Loss (total males) =	10%
Wounding Loss (females) =	
Wounding Loss (juveniles) =	
Total Bulls Adjustment Factor	80%

COLOR KEY	
example	Raw field data
example	Derived estimates
example	Years containing averaged ratio values

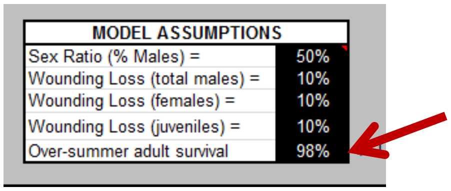
Classification Counts						Harvest					
Juv	Yrl Male	2+ Male	Total Males	Females	Total Classified	Juv	Yrl males	2+ Males	Total Males	Females	Total Harvest
5350	2000	2000	4000	12000	21350	50	3	2300	2303	400	2753
5350	2000	2000	4000	12000	21350	50	4	2300	2304	400	2754

Note there is only a single TBAF value for all models and for all years. Note also on the model pages, after running Solver the Derived Total-Male/100 Female estimates will be higher than the Field Estimates because the derived estimates are modified by the TBAF. If TBAF is 100% (i.e. bulls and cows have identical detectability) then derived and field estimates will be the same.

nt Juvenile / 100 Female		Posthunt Total Male / 100 Female				Total Classified	Winter Juvenile Survival		
Est	SE	Derived Est	Field Est	SE	Dev	N	Model Est	Field Est	SE
6	***0.73***	***41.9***	***33.3***	***0.61***	***0.1***	***21350***	***0.900***		
6	0.73	41.4	33.3	0.61	0.1	21350	0.900		
6	0.73	41.6	33.3	0.61	0.0	21350	0.850		
6	0.73	41.7	33.3	0.61	0.0	21350	0.827		
6	0.73	41.9	33.3	0.61	0.2	21350	0.900		
6	0.73	43.4	33.3	0.61	7.9	21350	0.450	0.65	0.08
6	0.73	39.3	33.3	0.61	15.3	21350	0.900		
6	0.73	41.2	33.3	0.61	0.5		0.766		
6	0.73	41.7	33.3	0.61	0.0	21350	0.709		
6	0.73	41.6	33.3	0.61	0.0		0.707		
6	0.73	41.6	33.3	0.61	0.0	21350	0.693		
6	0.73	41.6	33.3	0.61	0.0	21350	0.696		
6	0.73	41.7	33.3	0.61	0.0	21350	0.680		
6	0.73	41.6	33.3	0.61	0.0	21350	0.672		
6	0.73	41.6	33.3	0.61	0.0	21350	0.671		
6	0.73	41.6	33.3	0.61	0.0	21350	0.673		
6	0.73	41.6	33.3	0.61	0.0	21350	0.661		
6	0.73	41.6	33.3	0.61	0.0	21350	0.660		
6	0.73	41.6	33.3	0.61	0.0	21350	0.400		

Pronghorn Model. The pronghorn model calculates population size at the 'Beginning of the Bio-Year', as well as Preseason and Postseason. This is because Line-Transect abundance estimates are collected in June when pronghorn herds are relatively evenly dispersed. Thus, the pronghorn model has three additional columns – Total Males, Females and Total – relative to the elk and deer models. The population deviance is calculated based on this Beginning-Of-BioYear estimate. However, because classification data are collected later in the summer (~August), we must make an assumption about over-summer (June to August) adult survival in order to use the LT estimates in the model to calculate deviance. *The pronghorn model makes the critical assumption over-summer adult survival is a constant across all years.* In Colorado, biologists assume over-summer survival is 100%, so Preseason and Beginning-Of-BioYear adult population estimates are identical – this is a reasonable assumption in many areas given the life history of antelope. However, over-summer adult survival can be adjusted lower than 100%

(perhaps 99 or 98%) in the pronghorn model by modifying the ‘Oversummer Adult Survival’ cell in the DataInput tab.



MODEL ASSUMPTIONS	
Sex Ratio (% Males) =	50%
Wounding Loss (total males) =	10%
Wounding Loss (females) =	10%
Wounding Loss (juveniles) =	10%
Over-summer adult survival	98%

This value is linked to the model sheets and adjusts oversummer survival between the Beginning-Of-BioYear estimates and the Preseason estimates. Note the report form in the Summary tab has been modified to include the Beginning-Of-BioYear estimates.

Additionally, because classification data are collected in August (rather than December-January as for elk and deer), the pronghorn model also optimizes on the Prehunt Total Males and Females population estimates in the initial year, rather than on the post-hunt population estimates (which are used for elk and deer). Model estimated classification ratios (and deviances) are based on the Prehunt estimates as well. In deer and elk, these values are based on Posthunt estimates.

Deer Model. The deer model lacks a male visibility adjustment factor (elk) and a Beginning-Of-BioYear population estimates (Pronghorn), and is thus the most straightforward of the three species to model.

12. EQUATIONS

POSTHUNT POPULATION SIZES

The Posthunt population of Adult males and Females in Year *i* (Col **AE** and **AF**) are a function of the number of individuals the previous year (year *i*-1) multiplied by their survival rate, plus the number of surviving Juveniles of that sex minus the number of animals harvested that year (*i*):

Adult males$_i$ = ([Adult male$_{i-1}$] x [Model Est. Adult Survival$_{i-1}$]) +
([Juveniles$_{i-1}$] x [Model Est. Juvenile Survival$_{i-1}$] x [sex ratio of juveniles]) -
([Harvest Adult males$_i$] x [1 + wounding loss of adult males])

Females$_i$ = ([Females$_{i-1}$] x [Model Est. Adult Survival$_{i-1}$]) +
([Juveniles$_{i-1}$] x [Model Est. Juvenile Survival$_{i-1}$] x [sex ratio of juveniles]) -
([Harvest Females$_i$] x [1 + wounding loss of Females])

NOTE: "Juvenile survival" in the Spreadsheet model refers to survival of juveniles between 6-18 months of age. The juvenile:female ratios reflects the pregnancy rates combined with (ie multiplied by) the 0-6 month juvenile survival rates.

The number of Juveniles in Year *i* (Col **AD**) is simply the Juvenile to Female Ratio multiplied by the number of Females in the population divided by 100:

Juveniles$_i$ = [Field Est. Juvenile:Females$_i$] x [Females$_i$] x [1/100]

DEVIANCE AND PENALTY

Deviance = **Penalty** = $\left(\frac{FieldEst - ModelEst}{SE.FieldEst}\right)^2$

- This equation basically tells you how closely your field estimates are to your model estimates in a particular year. The farther your field estimate is from your model estimate, the larger the deviance or penalty. However, the difference between Field and Model estimates are weighted by the Standard Error (SE) of the Field Estimate. If your SE is large in a particular year, it means your Field Estimate is *im*precise which makes the relative size of the deviance or penalty small and less important when you use Solver to minimize the total deviance and penalty. Weighting by SE ensures precise estimates are more influential on the model than imprecise estimates.

- The whole goal of the model is to minimize the total deviance and penalties for all years, which are added together in cell **I5** (highlighted in yellow).

- There is no difference between deviance and penalty in how they are calculated. However, penalties are the weighted difference between Field Data and Optimized Model Estimates, whereas Deviance is the weighted difference between Field Data and Model values that are calculated based on both optimized model estimates (survival rates) as well as "known values" (harvest and juvenile ratios). For example, population deviance (Col **AH**) is the difference between field estimates and total population estimates (calculated from survival, harvest and juvenile ratios), while Juvenile and Adult Survival Pen (Col **L** and **P**) are the weighted differences between Optimized Model Estimate and Field Estimates.

STANDARD ERROR FOR CLASSIFICATION COUNTS

$$\textbf{SE} = \frac{a+100}{10}\sqrt{\frac{a}{n}}$$

…where

α = number of juveniles (or males) per 100 females
n = total number of classified juveniles (or males) + total number of females

This equation uses the individual animal as the sampling unit. It assumes classification sampling occurred *with* replacement, so animals could be counted more than once (*i.e.,* the total population size approaches infinity). More details on how to estimate SE in a finite population where sampling occurred *without* replacement, or how to calculate confidence intervals, or how to calculate an adequate sample size given a particular CI and population size, can be found in the following reference:

Czaplewski, R.L., Crowe, D.M. & McDonald, L.L. (1983) Sample sizes and confidence intervals for wildlife population ratios. *Wildlife Society Bulletin,* **11,** 123-128.

SURVIVAL ESTIMATES AND STANDARD ERROR

In the event that you have collared animals and would like to estimate survival rates and standard errors for the model, we recommend using the Kaplan-Meier method (Pollock et al. 1989), which involves using a staggered entry design that keeps tracks of mortalities and new collars over time. A spreadsheet is available from Colorado biologists to make this calculation easier. To convert the variance from this method into a standard error, simple take the square root of the variance.

Pollock, K.H., Winterstein, S.R., Bunck, C.M. & Curtis, P.D. (1989) Survival Analysis in Telemetry Studies: The Staggered Entry Design. *Journal of Wildlife Management,* **53,** 7-15.

ABUNDANCE AND STANDARD ERROR

Sightability-based estimates from the Idaho Fish and Game software and line-transect estimates (e.g. from program DISTANCE) provide abundance values that can be directly inputed into the spreadsheet model.

Standard Error from sightability models can be calculated using the total variance. To do this, add the Sampling, Sightability and Model variances together – provided at the bottom of the sightability output file—then take the square root of this variance to get the standard error.

13. COLUMN AND CELL DEFINITIONS

'DATAINPUT' CELLS

H3: Sex ratio is used for calculating the proportion of male and female juveniles that join the population of Adult males (Col **AE** in model spreadsheets) and Females (Col **AF** in model spreadsheets).

H4-H6: Wounding loss is the estimated proportion of each age class lost because of wounding. WGFD uses 10% though other states may use different rates. These values are added to the number of Harvested Adult males and Females in the calculations of total population of Adult males (Col **AE** in model spreadsheets) and Females (Col **AF** in model spreadsheets). Juveniles are currently not adjusted for wounding loss because juveniles are calculated from Post-hunt ratio data.

H7: *Differs by species*

- Deer: Sightability. This is currently not used in the model, but could be used to adjust abundance estimates if sightability has been estimated.
- Pronghorn: Over-summer adult survival. This is used to back-calculate the Beginning-Of-BioYear abundance from the Preseason Population estimates.
- Elk. Total Bulls Adjustment Factor. This value adjusts the derived estimates of Total Male/100 Females for detectability differences between male and female elk. If the value is 80%, it means males are detected 20% less detectable than females.

MODEL CELLS

(Found along the top of each model page)

A2:E6: These cells provide basic descriptions about the model, the species, the biologists doing the modeling, the herd unit and the date. These cells are linked to the SUMMARY tab so they should only be updated within the SUMMARY tab.

I3:I5: Model Fitting. These cells describe how well the model fits the data. Penalty and Deviance can be thought of as the degree to which the observed field estimates and the model estimates align or correspond with one another. Mathematically, deviance and penalty are exactly the same. However, Penalties are the weighted difference between Field Data and Optimized Model Estimates, whereas Deviance is the weighted difference between Field Data and Model values that are calculated based on both optimized model estimates (survival rates) as well as "known values" (harvest and juvenile ratios).

Sum of penalties sums up all columns that calculate penalties (Cols **E**, **N** and **R**).

Sum Lack of Fit sums up all columns that calculate a deviance (Col **I** and **AH**).

Total is the total penalties plus the total deviances. The Total is the cell (**I5**) Solver tries to minimize when it searches for the best parameter values. All model fitting values are calculated automatically so you do not need to change anything.

L3:L6 Model Selection. The values are used for selecting the best model among the different candidate models (See 'Selecting the best Model' for more details). Before comparing among models, you need to make sure to update *n* (the number of independent data points in the model, shaded in pink) and *k* (the number of Parameters being estimated in the model, shaded in green).

n = No. of years of Male/100 Female ratio field estimates (column **E**)
+ No. of Juvenile survival field estimates (column **J**)
+ No. of adult survival field estimates (column **N**)
+ No. of years of Harvest data (column **U**)
+ No. of years of Population size estimates (column **V**)

K = No. of green optimization cells

CJ, CA model: ***k*** = 4 (cells **Q3:Q6**)
SCJ, SCA model: ***k*** = 4 + sum of all juvenile (Col **J**) or adult (Col **N**) survival cells that vary by year
TSJ, CA model: ***k*** = 3 + sum of all juvenile survival cells (Col **J**)

DF = Degrees of Freedom = *n*-*k*

AICc → The Aikake Information Criterion corrected for small sample sizes. This value is used in the Summary Tab to identify the most parsimonious model and should not be modified by the modeler.

Q3:Q6: Parameters and Optimization cells. These values are optimized in Solver along with any time-specific survival rates. To do the optimization, Solver searches across all combinations of these values (subject to constraints) and identifies the best values as the ones that minimize the total deviance and penalty (cell **I5**). In the Constant Juvenile and Constant Adult Survival model (**CJ, CA**), Solver only has to search these four values. In other models, Solver also searches over time-varying survival values. The Initial Adult male and Initial Female populations (**Q5** and **Q6**) are divided by 10,000 to help Solver work more efficiently. Before running Solver, you must input rough guesses of these values.

V3:V6 Generic Initial Population. These values are meant as a convenience for inputting the starting values in the Parameters/Optimization box (**Q3-Q6**). They are not used directly in the model and they are simply rough estimates. Each time you run Solver, the values in the Parameters/Optimization box will change; having the initial starting values can be useful for future runs. To input these starting values into the Optim Cells, Copy/paste values in **V3:V6** into the Optim Cells in **Q3:Q6** (NOTE: you must 'paste special - values only' rather than simply copying and pasting). Starting values of Juvenile and adult survival are simply assumed to be averaged 5-year values for the species being modeled. The Starting initial Male and Female pop sizes are calculated by multiplying the initial posthunt population (**Y4**) by the male/100 female value.

Initial Male Pop Size (**V5**) = $Initial\ Pop\ Size \times \frac{M:F_{5yr\ ave}}{(M:F_{5yr\ ave}+J:F_{5yr\ ave}+100)} \times \frac{1}{10{,}000}$

Initial Female Pop Size (**V6**) **=** $Initial\ Pop\ Size \times \frac{100}{(M:F_{5yr\ ave}+J:F_{5yr\ ave}+100)} \times \frac{1}{10{,}000}$

Y4. Initial Posthunt Population size. This is a rough estimate of initial population size that assumes total population size is approximately the size of the male harvest multiplied by 12 (A. Holland, CDOW, *pers com*).

Initial Pop Size **(Y4) =** $\overline{Male\ Harvest}_{5yr\ ave} \times 12$

Y5-Y6. Initial Male/100 Female and Juvenile/ 100 Female. These values are calculated by taking the average of the first 5 years of field estimated ratio data. The Male:Female value is used in estimating the Initial Male and Female Population sizes (**V5:V6**).

Initial Male:Female **(Y5) =** $\overline{Male:Female}_{5yr\ ave}$

Initial Juvenile:Female **(Y6) =** $\overline{Juvenile:Female}_{5yr\ ave}$

MODEL COLUMNS

(Found in the middle of each model page. Note: columns may vary slightly among species)

A: Year

B. Field Estimate of Juvenile/100 Females Ratio. In the current deer model, Juvenile ratios are used to calculate the total Post hunt number of Juveniles each year (Column **AD**). Juvenile ratios are usually very precise (*i.e.,* they have a very small SE) and the WGFD assumes these ratios are 'known values,' similar to the Harvest data.

C. Standard Error (SE) of Field Estimates of Juvenile/100 Females. These values are not currently used in the model and could be deleted without effect. This column is included as a convenient way to store field data.

D. Derived Estimate of Adult Male/100 Female Ratio. This ratio is derived from dividing the total number of Post-hunt adult males (Col **AE**) by the total number of post-hunt females (Col **AF**). The model compares this ratio to the Field Estimated Total Male per 100 Female Ratio (Col **E**) in the estimate of Deviance (Col **G**).

E: Field Estimate of Post Hunt Total Male/100 Female ratio. Estimated from classification data. This value is compared to the derived estimate (Col **D**) to calculate deviance.

F: <u>SE of Total Male/100 Female ratio field estimate.</u> This value is used to calculate the weight (*i.e.,* relative importance) of the deviance for a particular year (see deviance calculation).

G. <u>deviance of Adult Male/100 Female Ratios</u>. This value represents the squared weighted difference between the Derived Adult Male/100 Female ratio (Col **D**) and the Field Estimated Adult Males per 100 Female Ratio (Col **H**). Without estimates of Juvenile or Adult Survival or Population Estimates, this is the only value that Solver uses to optimize on.

H. <u>Total Classified</u>. The total number of animals classified. Again, this column is included as simple way to store data.

I. <u>Model Estimates of Winter Juvenile Survival</u>. These estimates are used to calculate the proportion of juveniles that survive from the previous year to become females in Column **AF** (Predicted Posthunt Population of Females). When you run Solver, it searches for values of Juvenile Survival (as well as Adult survival) that minimize the deviance and penalty in cell **I5**. In the Constant Juvenile Survival Model (**CJ, CA**), these values are all constant and set equal to cell **Q3**; otherwise, they are semi or time varying and can be constrained within the Solver pop-up box in the "Subject to the Constraints" field.

J. <u>Field Estimates of Juvenile Survival</u>. These estimates are used to calculate a penalty (Col **N**) for Juvenile survival. The greater the penalty, the more importance it has influencing the optimal model estimated values. These values are not required each year to run the model, but they do help.

K. <u>Standard Error of Field Estimates of Juvenile Survival</u>. These estimates are used to calculate the weight of the penalty (Col **L**) for Juvenile survival. Again, the greater the penalty, the more importance it has in shaping the optimal model estimated values. These values are required each year you collect a Field Estimate of Juvenile Survival so the model can calculate the penalty.

L. <u>Penalty of Winter Juvenile Survival.</u> This calculates the penalty for the Winter Juvenile Survival rates. This column is summed and added to the SUM penalties in cell **I3.**

M. <u>Model Estimates of Adult Survival</u>. These estimates are used to calculate the proportion of adults (both Adult males and Females) that survive from the previous year in Columns **AE/AF** (Total number of Adult males and Total number of Females). When you run Solver, it searches for values of Adult Survival (as well as juvenile survival) that minimize the deviance and penalty in cell **I5**. In the Constant Adult Survival Models (**CJ, CA**, **TSJ, CA** and **TSJ, CA, MSC**), these values are all constant and set equal in cell **Q4**; otherwise, they are semi or time varying and can be constrained within the Solver pop-up box in the "Subject to the Constraints" field.

N. <u>Field Estimates of Adult Survival</u>. These estimates are used to calculate a penalty (Col **P**) for Adult Survival. The greater the penalty, the more importance it has in shaping the optimal model estimated values. These values are not required each year to run the model, but they do help. CDOW feels this is the single most important piece of additional data they collect.

O. Standard Error of Field Estimates of Adult Survival. These estimates are used to calculate the weight of the penalty (Col **P**) for adult survival. The greater the penalty, the more importance it has in shaping the optimal model estimated values. These values are required each year you collect a Field Estimate of Adult Survival so the model can calculate the penalty.

P. Penalty of Annual Adult Survival. This calculates the penalty for the Annual Adult Survival rates. This column is summed and added to the SUM penalties in cell **I3.**

Q. Year. Repeated to make scrolling easier.

R. Harvest of Juveniles. The number of juveniles harvested from the Harvest Survey Data. These are ***not*** used elsewhere in the model (the Predicted Posthunt Population of Juveniles (Col **AD**) is based on the Juvenile:Female ratio which is measured Post-hunt and already incorporates the hunted juveniles). Again, this column is included as a convenient way to store data.

S. Harvest of Total Males. The number of adult males (yearlings and 2+ males) harvested from Harvest Survey Data. These data are assumed to be 'known' values. These values are used to estimate the Predicted Posthunt Population of total males in Col **AE.**

T. Harvest of Females. The number of adult females harvested from Harvest Survey Data. These data are assumed to be 'known' values with no error. These values are used to estimate the Predicted Posthunt Population of Females in Col **AF.**

U. Total Harvest. Sum of juvenile, total male and female harvest.

V. Field Estimate of Post-hunt Population Size. These values come from systematic Line-Transect or Quadrat population surveys and are used to calculate Pop deviance (in Col **AH)**. Since trend counts do not correct for missed animals and do not allow for measures of standard error use of trend counts is inappropriate here.

W. Field Estimate of Standard Error of Post-hunt Population Size. These values come from systematic Line-Transect or Quadrat population surveys. They are used to weight the Pop deviance in Col **AH**. The greater the SE, the lower the weight (and thus importance) the Population Estimates have when you fit the model.

X:AA. Predicted Pre-Hunt Populations. These columns are simply the Predicted Post-Hunt Population sizes in year *i* plus the observed harvest. Harvest values are adjusted for wounding loss (Summary page, cells **D23:D25**).

AB-AC. Segment Harvest Rate. These columns calculate the percentage of the pre-hunt population harvested during the preceding harvest season. These are useful for analyzing harvest intensities over time but are not used elsewhere in the model.

AD. Predicted Post-Hunt Population of Juveniles. This value is derived by simply multiplying the Juvenile:Female ratio (Col **C**) and the Number of Females (Col **AF**).

AE. Predicted Post-Hunt Population of Adult males. This value is derived from many of the other model estimates (see Equation above). It essentially takes the previous year's Adult male population, multiplies it by survival, subtracts the number of Harvested Adult males, and adds the number of juvenile males that survived to adulthood from the previous year. The initial Adult male population (cell **Q5**) is one of the parameters estimated during the optimization step.

AF. Predicted Post-Hunt Population of Females. This value is derived from many of the other model estimates (see Equation above). It essentially takes the previous year's Female population, multiplies it by survival, subtracts the number of Harvested Females, and adds the number of surviving Female Juveniles from the previous year. The initial Female population (cell **Q6**) is one of the parameters estimated during the optimization step.

AG. Total Pop Size. This column is simply the sum of the Adult males, Females and Juveniles in Cols **AD**, **AE** and **AF** and represents the estimated population size each year. The model compares this value to any Field Estimated Population estimates (Col **V**) by calculating a population deviance (Col **AH**).

AH. Pop Deviance. This column allows the model to compare derived population estimates (Col **AG**) to Field Estimated Population estimates (Col **V**) by calculating a population deviance. By doing so, Field Estimates can provide additional information for estimating parameter values (Survival rates) during the optimization step.

14. MODEL ASSUMPTIONS

A model is only as good as its assumptions…

- Juvenile:Female ratios and all Harvest Data are assumed to be exact and contain no sampling or observer error.
- All mortality during the harvest season is a result of hunter harvest (i.e. there is no natural mortality during the hunting season).
- The Juvenile sex ratio is even: 50 male juveniles to 50 female juveniles.
- Wounding loss accounts for 10% of mortality
- With the exception of differences in harvest rate, the Adult Survival Rate (Col **O**) is equal for adult males and females except in the **TSJ, CA, MSC** model.
- Deviances and penalties for all types of data are weighted equally (i.e. by their Standard Error).
- The population is closed to immigration and emigration.

15. REFERENCES

Burnham, K.P., & Anderson D.R. (2002) *Model selection and multimodel inference.* Springer-Verlag,, New York, New York, USA.

Czaplewski, R.L., Crowe, D.M. & McDonald, L.L. (1983) Sample sizes and confidence intervals for wildlife population ratios. *Wildlife Society Bulletin,* **11,** 123-128.

Gaillard, J.M., Festa-Bianchet, M., Yoccoz, N.G., Loison, A. & Toigo, C. (2000) Temporal variation in fitness components and population dynamics of large herbivores. *Annual Review of Ecology and Systematics,* **31,** 367-393.

McCorquodale, S.M. (2001) Sex-specific bias in helicopter surveys of elk: Sightability and dispersion effects. *Journal of Wildlife Management,* **65,** 216-225.

Pollock, K.H., Winterstein, S.R., Bunck, C.M. & Curtis, P.D. (1989) Survival Analysis in Telemetry Studies: The Staggered Entry Design. *Journal of Wildlife Management,* **53,** 7-15

White, G.C. & Lubow, B.C. (2002) Fitting population models to multiple sources of observed data. *Journal of Wildlife Management*, **66,** 300-309.