

Glossary

Speedgoat

2021-05-26

Contents

1	Introduction	5
2	Sightability	7
2.1	Sightability Data Entry	8
2.2	Running a Sightability Model	11
2.3	Step-By-Step Sightability Model Example	14
3	Survival	21
3.1	Running a Survival Model	23
3.2	Step-By-Step Survival Model Example	25
4	IPM	31
4.1	Running the IPM	32
4.2	Step-By-Step IPM Example	35
5	Glossary	45
5.1	General	45
5.2	Sightability	46
5.3	Survival	48
5.4	IPM	49
5.5	MCMC	49

Chapter 1

Introduction

This manual is intended to provide support and guidance for using the IDFG PopR Website to run population models. Models covered here include Sightability, Survival, and IPM.

This manual was created by Speedgoat and IDFG using bookdown.

Feedback and suggestions for improving the user manual are welcome at eric.newkirk@speedgoat.io.

To get started with PopR, visit the Speedgoat homepage and select Idaho from the login menu.



Chapter 2

Sightability

Sightability models use aerial survey data combined with detection probability models to estimate total abundance within a sampling unit.

The basic idea behind sightability is a fairly simple process. First the sightability of the group is estimated by taking the covariate values and previously estimated coefficients. For example, if the model had one covariate called *veg* and the value of *veg* for a given observation was .8 then we can estimate sightability with something like:

$$\text{logit}(\text{sight}) = 1.1 - 0.6 * \text{veg}$$

Here the average sightability is something like $\text{logit}(1.1)$ or 0.75 and the effect of *veg* is roughly -0.6, so if we plug in our observation of 0.8 we would get

$$\text{logit}(\text{sight}) = 1.1 - 0.6 * 0.8$$

The value of *sight* is then 0.65. Now let's say that the group had 10 animals that were observed from aircraft. The estimate for the group size is $10/0.65$ or about 15 animals. That is the adjustment applied to each group. To conduct a sightability survey a manager would first define the desired scope of inference, say a DAU. Within the DAU there are GMUs and subunits. The area is then stratified based on previous observations and snow cover. For the sake of an example let's say we have 100 subunits and we have decided that 60 are in the high stratum, 30 are in the medium and 10 are in low. Within each stratum we have to decide how many to sample. If we decide to sample all of the highs, half of the mediums, and a quarter of the lows then we will have 60 high, 15 medium and say 3 low subunits to fly.

With data in hand we can estimate sightability as outlined above and then we will extrapolate those numbers to all of the areas that were not sampled. For simplicity, say our sightability adjusted estimates for each stratum were 250, 75 and 10 for high, medium and low respectively.

The general idea behind extrapolating is straightforward. To know what area was we simply divide the number of subunits flown by the population of subunits within a stratum.

High: $60/60 = 1$

Medium: $15/30 = 0.5$

Low: $3/10 = 0.333$

Now to use this number we will divide the adjusted number of animals by the proportion of subunits flown

High: $250/1 = 250$ animals

Medium: $75/0.5 = 150$ animals

Low: $10/0.333 = 30$ animals

This example hopefully shows the importance of sampling and how critical it is to know what stratum an observation came from. Without the stratum we cannot know the proportion of the area sampled and cannot make an estimate relate back to the population of subunits.

The extrapolation step also demonstrates that the density of animals is assumed to be constant. In the case of the mediums the area sampled produced 75 animals and it is assumed that the unflown subunits with the medium stratum are the same, which is how we get away with dividing by the proportion of area sampled. If a biologist only flies subunits where they know there are lots of deer then that number will be extrapolated to the remainder of the GMU and DAU.

2.1 Sightability Data Entry

There are 3 separate tabs available for entering sightability data. The GPS tab allows you to upload waypoints from a GPS file, the metadata tab allows you to add information about the overall survey, and the observation tab allows you to enter the actual counts.

When entering sightability data keep in mind that the model won't work unless the number of subunits sampled and the number of subunits available in each strata are defined accurately. It's critical to enter this information correctly on the metadata and observation entry tabs, including entering observations of 0 animals for subunits that were flown but where no animals were found.

2.1.1 GPS Data Entry

The first step in entering sightability data is to upload GPS data for your survey (if you don't have a .csv file containing GPS data for your survey you can skip

this step). To get started click **Sightability** in the sidebar, then **GPS Data Entry**.

GPS data definition

To successfully upload your GPS data to PopR, the data must be in the following format:

Title	Date Created	Latitude	Longitude
1	2016-12-06 09:39:19 AM MST	44.93516298	-113.6274062
2	2016-12-06 10:05:37 AM MST	44.90122665	-113.6149363

- The first column is named Title (case sensitive) and contains a number that can be used to index the points
- A column representing Latitude, which must be named Latitude (case sensitive)
- A column representing Longitude, which must be named Longitude (case sensitive)

Any GPS data following these rules should work with the program. Note also that the data are collected in decimal degrees and assumed to be WGS84 with a negative Longitude value.

To upload GPS data:

1. Click Browse towards the upper left of your screen
2. Select the file containing the GPS data you wish to use
3. Click Open
4. Click the Write to DB button when satisfied that the data is accurate (if a file by the same name already exists the user will be prompted for a different name)

Once you have selected your file a table should appear at the right of the screen. This table can be copied to your clipboard, printed, downloaded and columns can be hidden for easier viewing. The table can also be sorted by using the arrows next to the column names. Filtering the table is also possible using the boxes at the top of each column. If you don't like the order of the columns try clicking on a column name and while holding the left mouse button dragging the column to the location you prefer. In addition to the table a map of the points is rendered. The map is intended as a quick check of the feasibility of the points. In order to help the user better understand where the locations were recorded there are three different potential map layers available via the button in the top right corner of the map. The map itself is also interactive. The user may navigate with their mouse by clicking and dragging. Zooming is accomplished

by the buttons in the top left of the map or by scrolling the mouse wheel. One tip when using the mouse wheel, if your mouse is over the map the map will zoom, but if your mouse is anywhere else the webpage will scroll.

Where are my files?

After clicking on the Browse button in PopR you should see a typical file explorer window. If you then click on My Computer your device should be visible. From this point you can navigate to the location of the files on your device and select the files one at a time. If you click on your device there are no files visible try changing the USB settings. To do this, on an Android device, drag your finger down from the top of the screen and select *Transferring images via USB*. Once selected you should see a popup that allows you to select one of several options. Toggle the options until you can browse the file structure of your device. For more help with this process or if your device is a Mac or iSomething try this link. Note that the MTP connection does not always work for me and so I often use the via PTP option.

On an Android device the typical location of files is something like: Computer > Galaxy Tab S2 > Tablet > Documents or Download

2.1.2 Metadata Entry

The next step is to enter the metadata for your survey. The Mmetadata entry tool exists to define the survey and prevents the user from having to so re-enter data that remains constant for a survey period. For example, the total number of subunits available to sample in each stratum remains constant over an entire survey period (generally a single year).

Click **Metadata Entry** in the sidebar under **Sightability**, then select the species and survey type. Enter the date the survey took place, the type of aircraft used, and the number of observations (rows) recorded on the datasheet. Select the region and analysis unit, and if you uploaded a file in the previous step, select it in the GPS Data File input. Enter the total number of subunits available in each density stratum (not the number flown), and finally enter any comments or notes about the survey. Once the form is complete click the Submit Response button to continue. A dialog will appear confirming that the data were saved.

If you made a mistake on the Metadata form you can correct it on the Past Surveys tab. This tab presents a spreadsheet with all of the survey metadata entered to date. You can delete rows by right-clicking or edit the information in any of the fields. If you make changes on the Past Surveys tab make sure to click the Save Edits button when you're finished.

2.1.3 Observation Entry

Once the Metadata is entered correctly click **Observation Entry** on the sidebar, then click the Load button to get started. Use the inputs in the dialog that appears to select the survey you entered in the previous step, then click Load Data. A spreadsheet will appear with the Species, Survey Type, Date, and Area populated for you. Fill in the additional columns (beginning with Stratum) with the information from your datasheet. You may need to use the horizontal scroll bar below the spreadsheet to access all the columns if they don't fit on your screen. The number of rows available in the spreadsheet will match the number of observations you entered in the Metadata form, but if you need to add additional rows you can do so by right-clicking in the last row and selecting "insert row below." You can view a summary of the data you've entered below the spreadsheet.

The observations entered here are used to determine which subunits were actually sampled, so it's critical to enter an observation for every subunit that was flown, even if the number of animals counted is zero.

Once all the data from your datasheet has been added to the spreadsheet click **Validate** to check for errors. If any errors are present a dialog will appear explaining what they are and how to fix them. Once any errors are corrected use the Save button to save the data to the server.

2.2 Running a Sightability Model

2.2.1 Loading Data

To run a sightability model select **Sightability** in the sidebar then click **Setup**. Select a species and DAU on the Load Data tab and click the Get Survey Data button to download data from the IDFG database. Once the download is complete a dialog will confirm that the data has been downloaded and alert you to any errors present in the data (see the errors section). Data will only be read for the species and DAU selected, so if you switch either selection later you'll have to repeat the process.

Once the data is downloaded you can review it in various formats in the pane to the right.

- Flight record: Summarizes the data downloaded by survey type and bio year.
- Sampling: Summarizes the data downloaded by stratum, showing the population of units available and how many of those were sampled. This is a great table to check against your expectations - how many high, medium and low subunits did you fly? Does the data match? This is also a good

place to make sure the stratum and unit were entered correctly for all of your observations - missing values in those columns are a common cause of errors.

- All Data: Shows all of the observation data downloaded for this species and DAU.
- Data to Model: Shows the observation data as it will be passed to the model for transparency.

If any errors appear in the data you can correct them through the **Metadata Entry** or **Observation Entry** pages in the sidebar. If you make any corrections be sure to load the data again when you return to the **Setup** page.

2.2.2 Model Options

After loading sightability data for a particular species and DAU move on to the Options tab to specify the parameters for your model. Select a year from those available in the data, then choose whether to analyze by area, unit, stratum, or sub-unit. It's important to note that your results cannot be used to extrapolate to a wider geographic area if you select unit, stratum, or subunit in this dropdown. Finally select a survey type and aircraft - the available selections for these may be limited based on the species selected.

Move on to the Sampling tab and review the values displayed. The population of subunits within each stratum (the total number) is shown to the user in the pop column. The number of subunits flown as part of the sample are displayed in the obs or observation column. Recall that the model will extrapolate the estimates based on the proportion of the total area that was sampled. The proportion is calculated by dividing the number of subunits where observations were made by the total number of subunits and is shown to the user in the last column, PropSampled. For example, the expectation is that high is close to 1, medium is close to 0.5 and low is something lower like 0.25. In the case that zero observations were made within a stratum a 1 will be inserted. This is because division by 0 is not possible. Extrapolation is performed by taking an estimate and dividing by the proportion sampled, so if we insert a 1 for either a census or a place where no observations were made the result is that no extrapolation is performed.

When you've selected the appropriate options and adjusted the sampling parameters (if necessary to achieve correct proportions sampled) move on to the Run tab and click Fit Model. If no errors are encountered dismiss the dialog and click **Tables** in the sidebar to view your results. If the model fails to run check the errors section.

2.2.3 Model Output

Once you've run a model you can view the results on the **Tables** page. The top half of the page shows the data used in the model, summarized by category:

- **Sample:** Displays how many subunits were available in each stratum, and how many of those were sampled.
- **Activity:** Displays the number of observations for each activity category.
- **Vegetation:** Displays the number of observations for each vegetation type.
- **Snow:** Displays the number of observations binned by snow cover.
- **Group Size:** Displays the number of observations binned by group size.

The bottom half of the page shows the model output:

- **Detection:** Displays the probability of detection (as a range) assigned to the input observations.
- **N:** Displays the abundance estimates produced by the model by demographic, including the variance for each estimate.

2.2.4 Adding Results to IPM

Use the **Admin** page to add the results of your sightability model to the input data for the IPM. Only add the results to an IPM dataset if you're sure you used the correct settings and data. If there is already sightability data in the IPM data file you select for the same year and unit, that data will be overwritten with the results of your current model. If you ran a model on composition survey data, sex and age ratios will be added; for sightability surveys abundance estimates will be added.

Select the IPM dataset to which you want to add your estimates, and verify that the estimates themselves look reasonable, then click Update DB to add the estimates to the database file.

2.2.5 Errors

Reporting errors This section of the documentation is a work in progress. We need your feedback to identify and document the errors users commonly experience running models with PopR - you can help us by sending an email to eric.newkirk@speedgoat.io any time you experience an error on the site. Please include as much information as you can about your settings, data, and the error message you received, including screenshots if possible. Thanks!

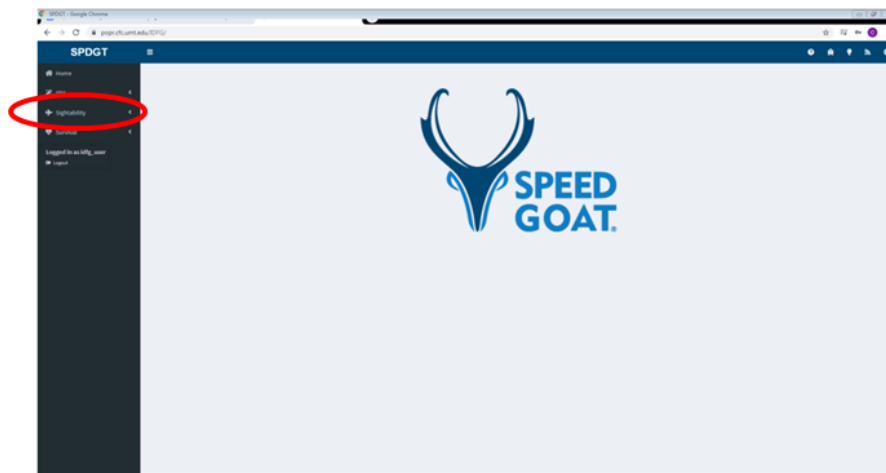
Missing values in the stratum or unit fields These errors should be apparent when reviewing the data in the **Sampling** tab of the **Setup** page. Missing values in these columns can cause data to be discarded or cause the model to fail. To fix them navigate to the **Observation Entry** page, load the data for the survey in question, and fill in the missing information. Don't forget to click the button to save your changes when all the values have been entered.

Empty rows at the bottom of the observation table Sometimes when the metadata and observations entered for a survey don't agree there can be empty rows at the bottom of the observation data entry table. If that happens load the survey data in the **Observation Entry** page and right-click on the empty row(s) to delete. As above, don't forget to save your changes.

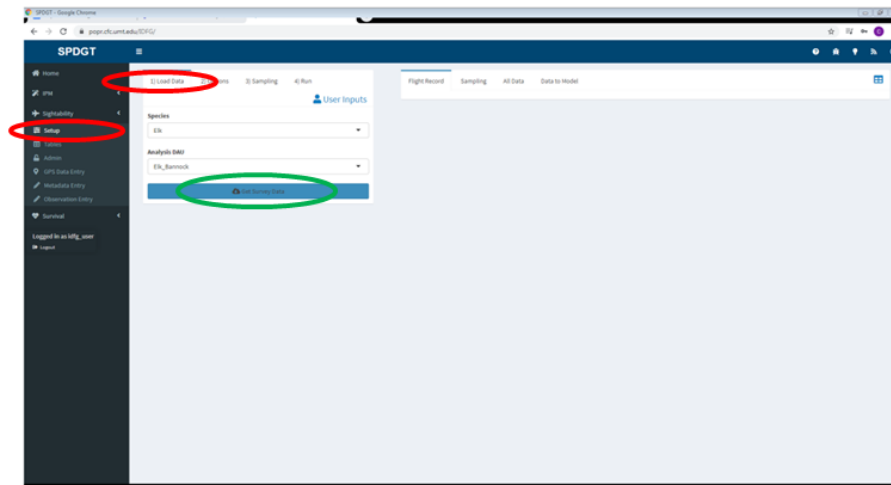
2.3 Step-By-Step Sightability Model Example

This example shows how to run a sightability model for mule deer in the Smokey-Boise DAU step by step, including screenshots of the website.

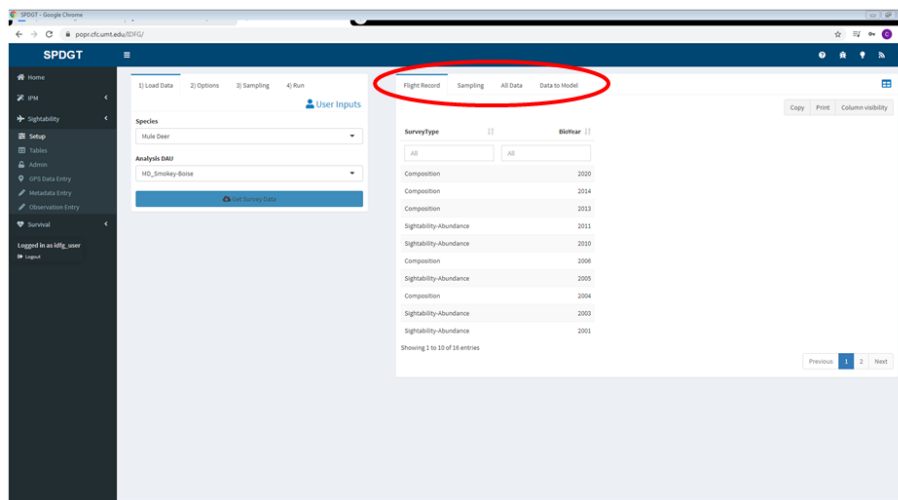
1. Start by clicking **Sightability** in the sidebar.



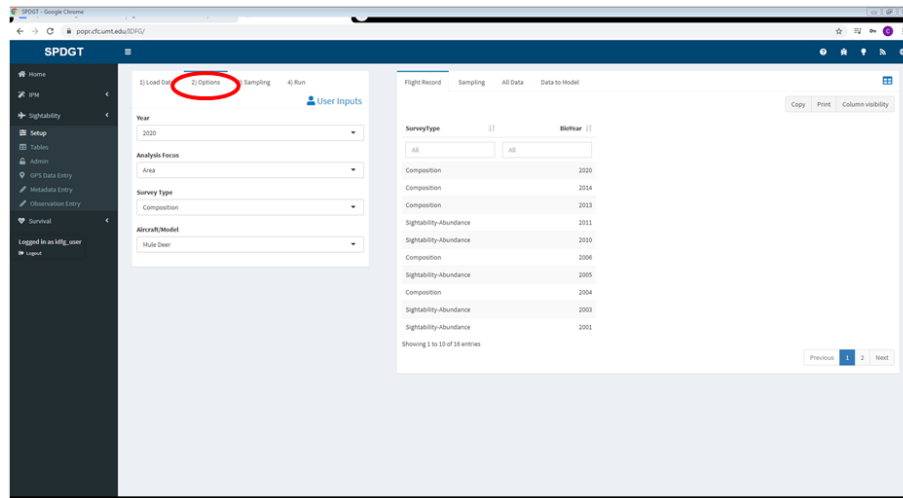
2. Select the **Setup** page in the sidebar, then select the species and DAU you want to model. In this example we use mule deer and MD_Smokey-Boise. Make sure to select the correct species and DAU before clicking the button to load sightability data.
3. Click the button labeled “Get Survey Data” to load the data for the species and DAU you selected.



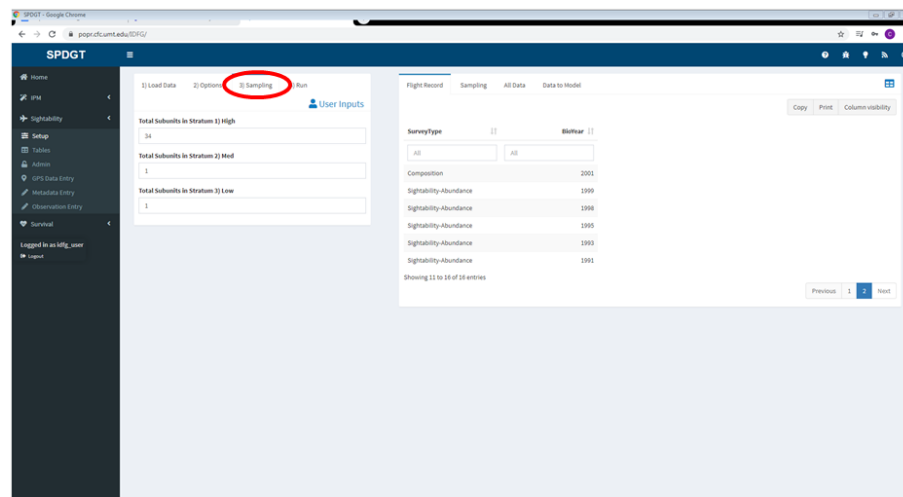
4. Once the data is loaded you should see a dialog indicating that the process was successful. Click Dismiss. If you receive an error see the errors section.
5. Review the data in the panes to the right (see Loading Data for details).



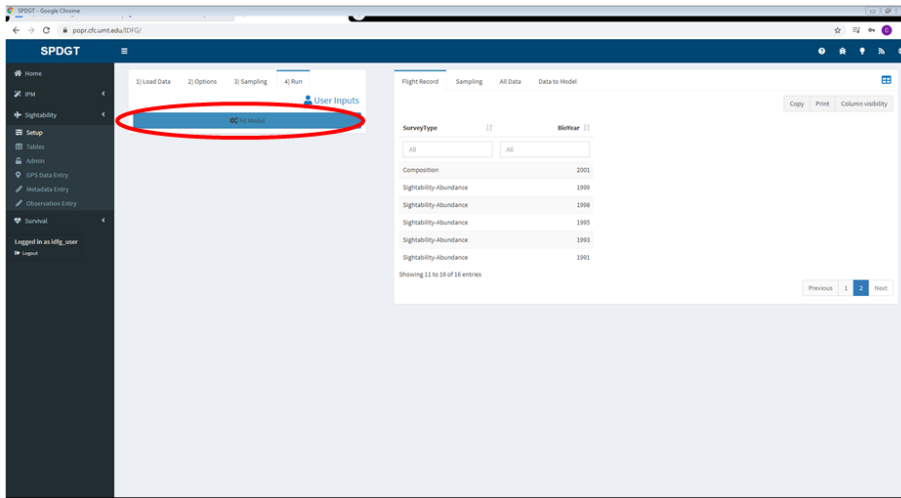
6. Move on to the **Options** tab in the **User Inputs** pane, and make the appropriate selections. In this example we run a model on 2020 composition data for mule deer, with area as the spatial focus of the analysis.



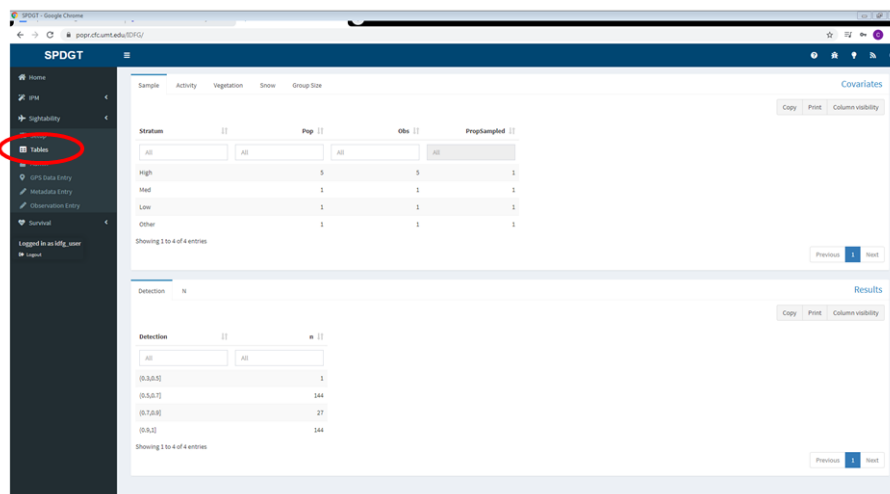
7. Move on to the **Sampling** tab and check that the data are accurate by selecting Sampling in the pane to the right. If the total population of subunits available in each stratum is incorrect you can fix it here.



8. Now that you've selected the appropriate settings and reviewed the input data, switch to the **Run** tab and click Fit Model.



9. Once model fitting is complete select the **Tables** page in the sidebar to view the output.



10. The covariates section on the top half of the page summarizes the observations used in the model and how they fall into the various strata, activity categories, etc. See Model Output for more details.
11. Select the **N** table on the bottom half of the screen to view abundance estimates.

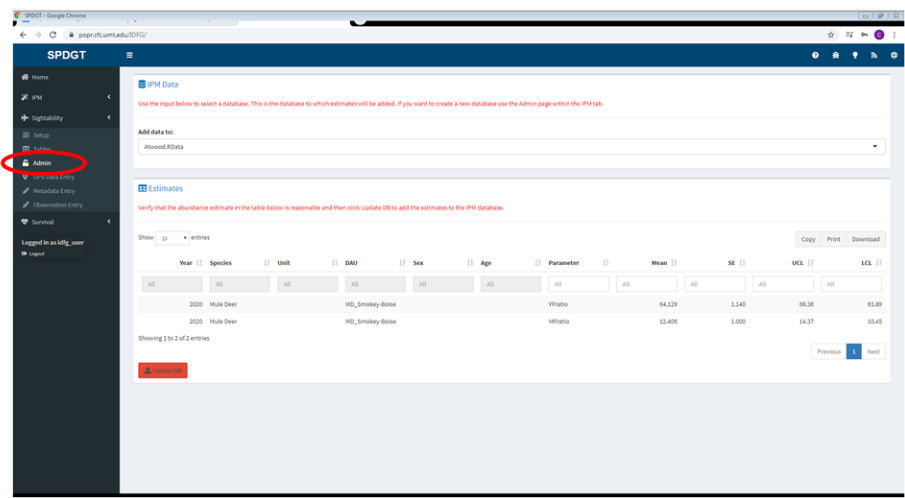
SurveyType	Area	Demographic	Year	Mean	Modifier	Sampler	SightRate	TotalRate
Composition	MD_Smokey-Bowle	Total	2020	2393.8	15952.653	0	5379.125	23351.778
Composition	MD_Smokey-Bowle	AdFemale	2020	1353.67	5250.052	0	1911.056	7342.009
Composition	MD_Smokey-Bowle	YOV	2020	886.1	2687.219	0	873.11	3580.329
Composition	MD_Smokey-Bowle	TwoPoints	2020	87.87	0	0	10.407	0
Composition	MD_Smokey-Bowle	ThreePoints	2020	42.13	1.048	0	11.825	52.872
Composition	MD_Smokey-Bowle	FourPoints	2020	37.83	0.309	0	3.87	4.178
Composition	MD_Smokey-Bowle	VF_M000	2020	0.64	0	0	0	0.633

12. Select the **Detection** table to view a summary of observations binned by probability of detection.

Detection	Area	TotalRate
All	All	1
(0.5,0.5]		144
(0.5,0.7]		27
(0.7,0.9]		144

13. If you're confident in the results and the settings used to run the model, you can add the sightability results to an IPM dataset using the **Admin** tab in the sidebar. Be sure to read the instructions and understand the difference between the various IPM databases before using the admin tab.

2.3. STEP-BY-STEP SIGHTABILITY MODEL EXAMPLE



Chapter 3

Survival

Estimating survival is a complicated process that reflects our inability to directly observe the quantity of interest. One way that we might start thinking about survival is by asking how many animals are alive at the end of some interval given how many started the interval. For example, we might collar 100 deer on the first of the month and then 30 days later we check on each animal and find that 98 deer are still alive. The expectation for survival then is $98/100$ or 0.98. This simple expectation only works because the fate of ALL animals was known and every single individual was captured at the same time. This is not what happens in practice. When capture occurs animals are collared over a period of time that can span a few months. If we applied the same simple math as above we would not be very accurate because animals are exposed to risk for different amounts of time. For example, if we collar one deer on January 1 and another on March 1 and we want to know how many live until December we can't use division because one animal was available to die in several more months. Another wrinkle that often comes into play is collar failure. This essentially causes the animal to leave the sample and complicates estimation.

Sample size is a funny topic in survival modeling because very little information is contained in knowing the animal is alive, but we learn a lot when the animal dies. We basically need to observe lots of death. For example, the estimator above suggested $98/100$, which is similar to the Kaplan-Meier estimator (KM). Many agencies use the KM estimator because it is intuitive and relatively simple to implement. However, there are several problems with this approach and they appear at the extremes of sample size. One extreme is the case where 0 animals die, in this case the variance is undefined. Another extreme occurs in almost every study, as time marches on our sample size shrinks and eventually we may only have 10 or 20 animals marked. Let's say we have a study and only 10 animals are still on the air. If during the winter 4 of the animals die then we arrive at a survival estimate of $6/10$ or 0.6. Is this correct? Not likely, not unless you really think that 40% of the population perished. To make this point a little

better say we only have 3 animals marked. What are the values of survival that we can estimate? Well, if 0 animals survive we can get 0, $\frac{1}{3}$, or $\frac{2}{3}$. What if the true value of survival was something like 0.85? Turns out you can't even get that number from a sample size of three. The same is true of coin flips, if you flip a coin three times you can't get 0.5 the true probability of getting a heads.

With all of the complexities of survival estimation we generally have to write somewhat complicated models to get at the quantities of interest while protecting against bias. The models in PopR use a discrete-time known-fate approach to survival estimation. The models allow the user to estimate by age, sex, space and subset to different time intervals. We have written the models in this way to users to explore a variety of questions, share information among data rich and data poor areas and easily incorporate covariates. A nice example of the models in action is Regional-scale models for predicting overwinter survival of mule deer (Hurley et al. 2017). This paper will also give nice descriptions of the math and covariates used in predictions. One point to keep in mind is the idea of sharing information via random effects. The models and results in IPM database `mule_deer_shared` come from running a statewide model where information is shared among DAUs within a month. This means that the model can use all of the data from the state to figure out what an average, say, March survival should be and then further refine the estimate for any given DAU by also using the DAU specific data. I like to say that the model returns the statewide average when little to no data exist, but gives the opportunity to be different from the statewide average when sufficient data are present. The same DAU specific information can be extracted from these models. For example, if you want to know survival in March of 2019 in the Beaverhead DAU that information is available. The last feature of survival models that I will discuss are the multi-state models. These models allow the user to estimate not only survival, but the probability of death by harvest or death by some other cause. One equation to keep in mind here is $\text{Survival} + \text{Mortality} = 1$. This is always true and must be true. If 30% of the population is harvested in a given year then the maximum value of survival is 0.7 if we assume that no animals died of any other cause. In fact, this equation is part of the reason why we care about abundance because the survival rate or harvest mortality rate can be applied to the abundance to determine the number of animals that died in a given year. Anyway, using the multi-state models the software will report Survival, Harvest Mortality and Other Mortality. Remember that if you are conditioned to seeing survival with harvest removed then survival is equal to $1 - \text{other mortality}$. For example, I expect adult female mule deer survival to be about 0.85 in the absence of harvest and I might get that value if I censor harvest in the known fate model. However, because $\text{survival} + \text{harvest mortality} + \text{other mortality} = 1$ the survival from the multi-state model will be lower. If we rearrange the above equation to solve for survival + harvest then we will probably get a number like 0.85 from $1 - \text{other mortality} = \text{survival} + \text{harvest mortality}$, which is essentially the probability of surviving non-harvest causes of death.

3.1 Running a Survival Model

3.1.1 Loading Data

To run a survival model select **Survival** in the sidebar then click **Setup**. Select a species and DAU on the Overview tab and click the Download Data button to download data from the IDFG database. Once the download is complete a dialog will confirm that the data has been downloaded and alert you to any errors present in the data (see the errors section). Data will only be read for the species and DAU selected, so if you switch either selection later you'll have to repeat the process.

Once the data is downloaded you can review it in various formats in the pane to the right. Data is pulled from SAMM, so if you find individuals that plot out in the wrong area or DAU, you'll need to fix them there.

- **Captures:** Shows a map of the capture locations of all the individuals used in the analysis. The table on the left shows any IDs that have bad or missing capture location information.
- **Mortalities:** Shows a map of the mortality locations for the DAU. The table on the left shows any IDs that have bad or missing mortality location information.
- **Collared Animals:** Shows a barplot of the number of collared individuals throughout the analysis years. You can hover over each bar with your mouse and at the top of the graph, it will show you the status of collared individuals by month. Categories include missing, other mortality, harvested, and alive.
- **Raw Data Summary:** Shows a summary of the raw data available to be for the selected DAU.

All the data available for the selected DAU can be viewed using the Captures, Monitoring, Mortalities, Encounter Histories, and Raw Data tabs in the bottom table. The Captures and Mortalities tables show information for *all* individuals, even if capture/mortality information is unavailable. You can sort using the arrows next to the headings and filter by entering names/values into the boxes under the headings.

You can also view the data loaded from SADD on the **Individuals** page, where you can view raw data, encounter histories, captures, and mortality records. You can also filter to a specific animal, which can be helpful in tracking down errors.

3.1.2 Model Options

After loading survival data for a particular species and DAU move on to the Subset tab to specify the parameters for your model. Select a sex (default is

all) and age class (default is adult) to model, then select the range of years to include in the analysis (default includes all data available).

Move on to the Structure tab and use the inputs to define the model that will be run on your data. First select a known fate (simple survival estimation) or multi-state (estimate mortality rates by cause - currently unavailable) model. If you're running a multi-state model use the Censor Fate input to exclude a particular source of mortality from the analysis. For example, if harvest is selected the estimated survival will answer the question "What would survival be in the absence of harvest?" Finally define how survival is allowed to vary with sex, time, and space. You must select more than 5 years for analysis to model survival as varying through time, and modeling survival as spatially varying is only available for statewide models.

If you plan on adding your results to an IPM dataset you should use the default settings:

- Sex: All
- Age: One at a time, A then J
- Analysis Years: 2000-Current Year
- Model Type: Known Fate
- Covariates: None
- Censor Fate(s): Harvest
- Model sexes as: Different
- Model time as: Monthly
- Model space as: Varying
- Thinning Rate: 1
- Burnin Length: 35,000
- MCMC Iterations: 25,000

When you've loaded your data, selected the appropriate options, and defined your model structure, move on to the Run tab, specify your MCMC settings, and click Fit Model to run the model. If no errors are encountered dismiss the dialog and click **Tables** in the sidebar to view your results. If the model fails to run check the errors section.

3.1.3 Model Output

Results are viewed under the **Tables** subheading on the left under **Survival**. Estimates can be viewed by year or by month. Mean estimates are accompanied by the standard deviation as well as the lower and upper confidence intervals. You can copy, print, or download the data using the buttons at the top right of the table. You can also sort and filter using the boxes below the column headings.

3.1.4 Adding Results to IPM

Use the **Admin** page to add the results of your survival model to the input data for the IPM. Only add the results to an IPM dataset if you're sure you used the correct settings and data. If there is already survival data in the IPM data file you select for the same species, year, and unit, that data will be overwritten with the results of your current model.

Select the IPM dataset to which you want to add your estimates, and verify that the estimates themselves look reasonable, then click Update DB to add the estimates to the database file. DO NOT add results to the IPM database if you didn't use the default settings outlined in the Model Options section.

3.1.5 Errors

Reporting errors This section of the documentation is a work in progress. We need your feedback to identify and document the errors users commonly experience running models with PopR - you can help us by sending an email to eric.newkirk@speedgoat.io any time you experience an error on the site. Please include as much information as you can about your settings, data, and the error message you received, including screenshots if possible. Thanks!

3.2 Step-By-Step Survival Model Example

This example shows how to run a survival model for mule deer in the Smokey-Boise DAU step by step, including screenshots of the website.

1. Start by clicking **Survival** in the sidebar.

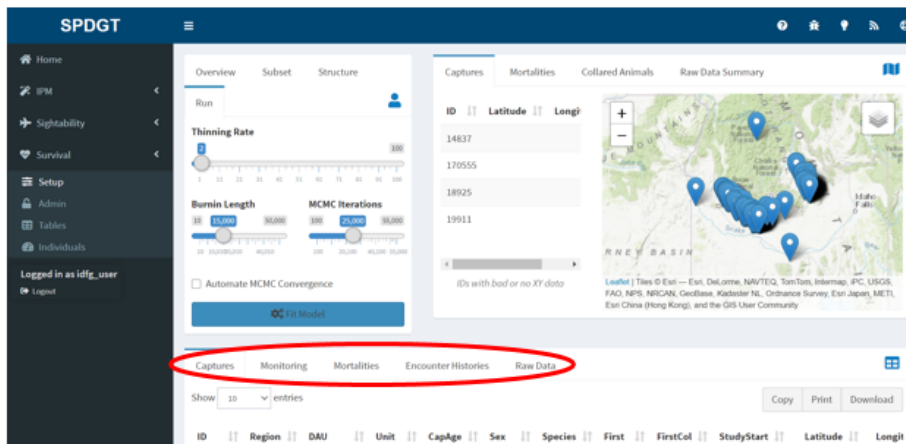
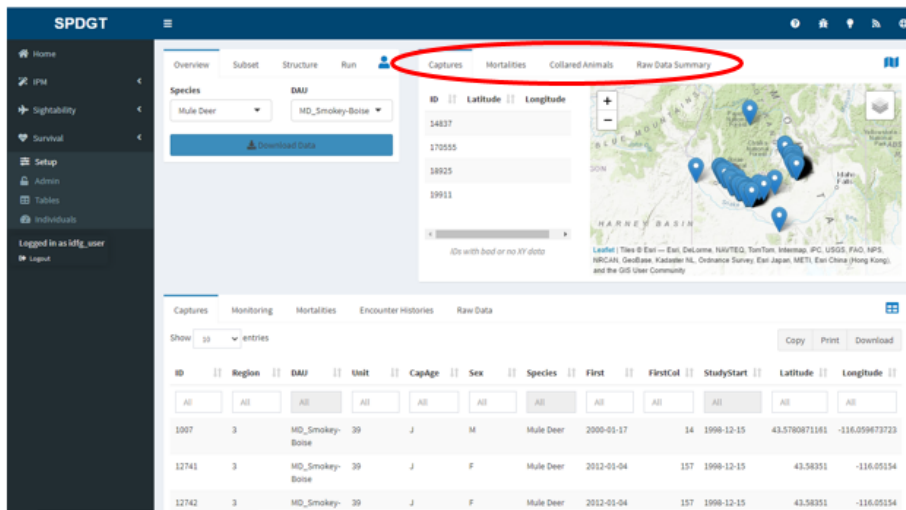


2. Select the **Setup** page in the sidebar, then select the species and DAU you want to model. In this example we use mule deer and MD_Smokey-Boise. Make sure to select the correct species and DAU before clicking the button to load sightability data.
3. Click the button labeled “Download Data” to load the data for the species and DAU you selected.

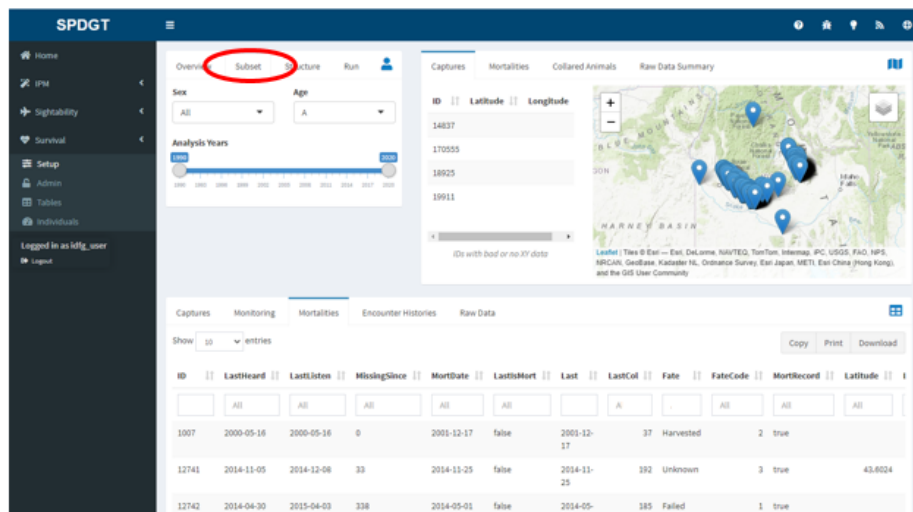
The screenshot shows the SPDRGT web application. The sidebar on the left has a 'Setup' button highlighted. The main panel has a 'Species' dropdown set to 'Mule Deer' and a 'DAU' dropdown set to 'MD_Smokey-Boise'. Below these dropdowns is a 'Download Data' button, which is circled in red. To the right of the dropdowns is a map showing the study area in the HARE BASIN. Below the map is a table with columns for ID, Region, DAU, Unit, CapAge, Sex, Species, First, FirstCol, StudyStart, Latitude, and Longitude. The table contains three rows of data for Mule Deer individuals.

ID	Region	DAU	Unit	CapAge	Sex	Species	First	FirstCol	StudyStart	Latitude	Longitude
1007	3	MD_Smokey-Boise	39	J	M	Mule Deer	2000-01-17	14	1998-12-15	43.5760871161	-116.099473723
12741	3	MD_Smokey-Boise	39	J	F	Mule Deer	2012-01-04	157	1998-12-15	43.58351	-116.05134
12742	3	MD_Smokey-Boise	39	J	F	Mule Deer	2012-01-04	157	1998-12-15	43.58351	-116.05134

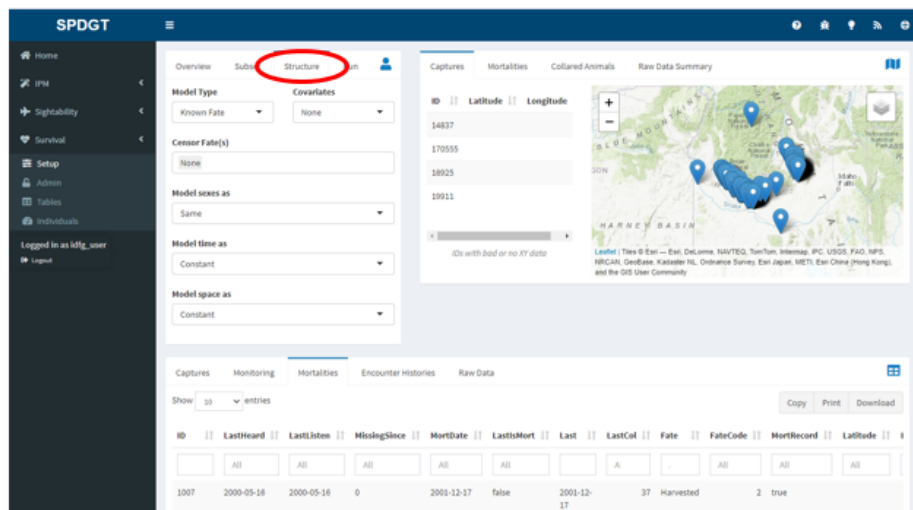
4. Once the data is loaded you should see a dialog indicating that the process was successful. Click Dismiss. If you receive an error see the errors section.
5. Review the data in the panes to the right and tables at the bottom (see Loading Data for details).



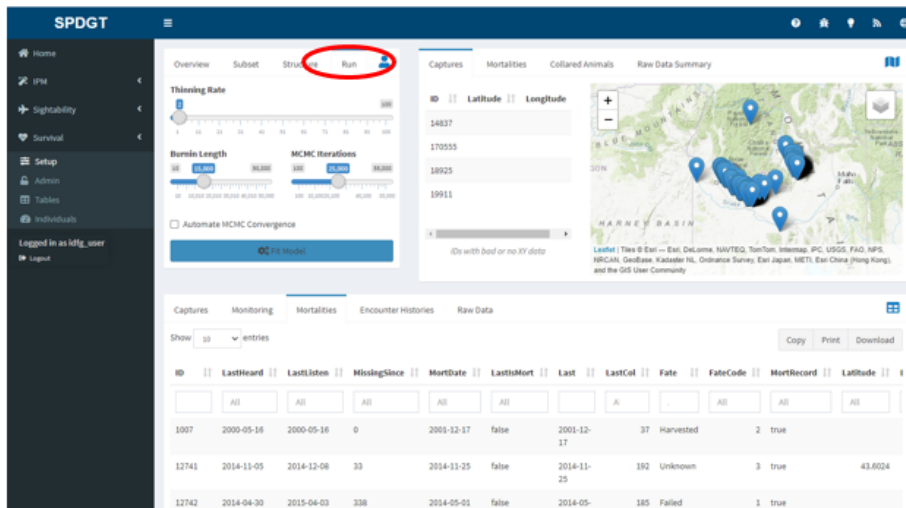
- Move on to the **Subset** tab in the **User Inputs** pane, and make the appropriate selections. In this example we ran a model for adults of both sexes using data from 2000 to 2020.



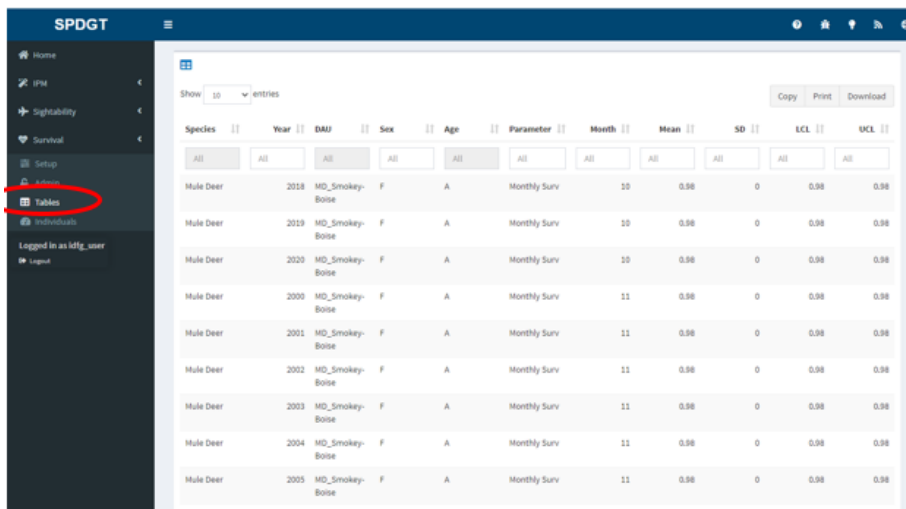
7. Move on to the **Structure** tab and define the model you want to run. Here we use a known fate model where survival does not vary with sex, space, or time.



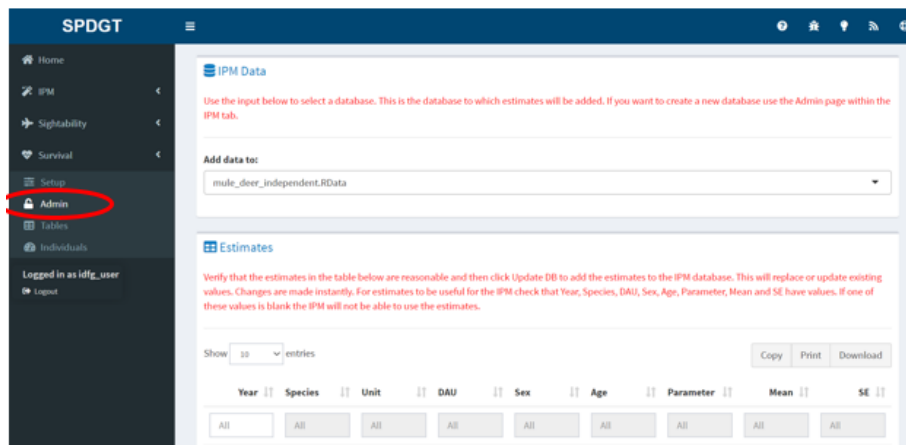
8. Now that you've selected the appropriate settings and reviewed the input data, switch to the **Run** tab and click Fit Model.



- Once model fitting is complete select the **Tables** page in the sidebar to view the output.



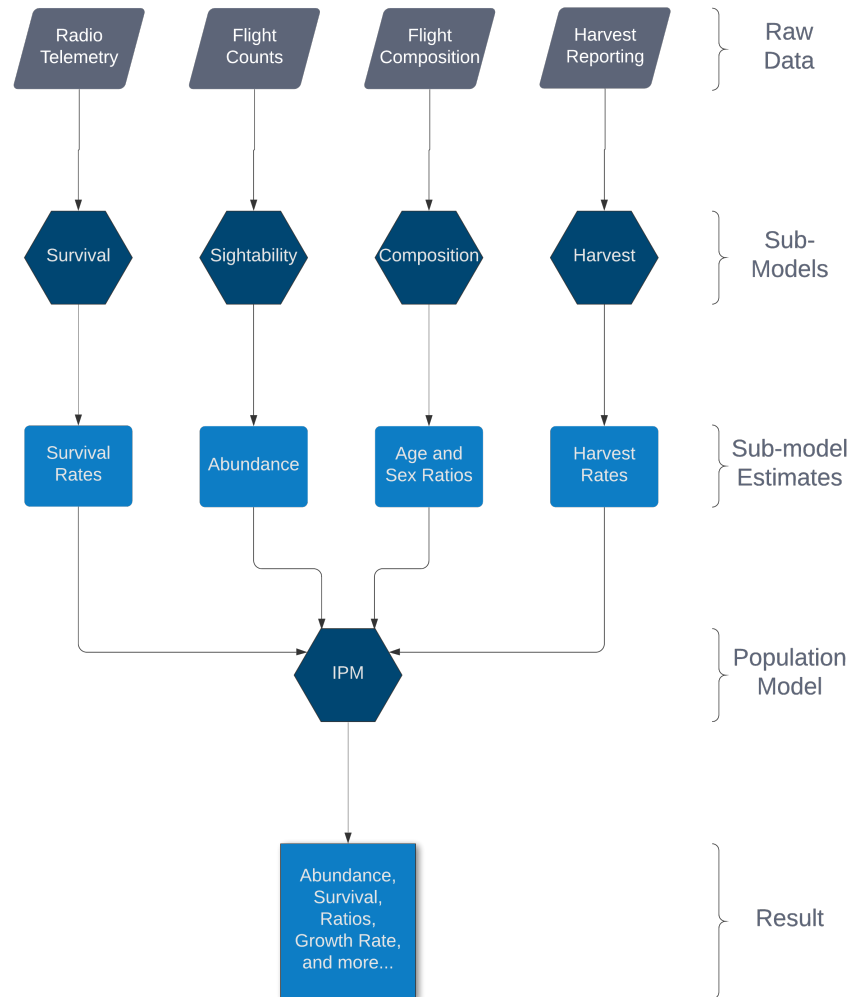
- If you're confident in the results and the settings used to run the model, you can add the survival results to an IPM dataset using the **Admin** tab in the sidebar. Be sure to read the instructions and understand the difference between the various IPM databases before using the admin tab.



Chapter 4

IPM

PopR allows you to combine data from various sources to model wildlife populations using an integrated population model (IPM), leading to more robust and defensible estimates. First the data from each source are modeled independently to estimate individual population parameters, such as sex ratio and survival. Then these results are combined in the IPM to find the best fit for all of the available data.



4.1 Running the IPM

4.1.1 Loading Data

To run an integrated population model select **IPM** in the sidebar then click **Setup**. Running an IPM requires 2 types of data: population estimates from

an IPM database and harvest data from the IDFG database. Start by selecting a species and DAU on the Overview tab, then select the IPM database you want to use and choose how the model incorporates future weather. At this point you can review the population estimates from previous models in the plots to the right and the tables along the bottom, but notice that there is no data available in the Harvest tabs. To fix that click the Fetch Harvest Data button to download the appropriate data from the IDFG database. Once the download is complete a dialog will confirm that the data has been downloaded and alert you to any errors present in the data (see the errors section). Now you can review the harvest data in the plots and tables in addition to the population estimates. Data will only be read for the species and DAU selected, so if you switch either selection later you'll have to download harvest data again.

Once you've loaded estimates from the IPM database and downloaded harvest data from SADD you can review everything in the plots to the right and tables below. All data available to the IPM for analysis is shown in these panes. If you cannot find data that should be used by the IPM for analysis in either of these areas then that data has not been made available to the model. This may indicate that the missing data needs to be passed to the IPM via the sightability or survival tabs.

4.1.2 Model Options

After loading data for a particular species and DAU move on to the Structure tab and use the inputs to define the model that will be run on your data. Use the inputs to define whether the model uses population reconstruction and whether reproduction and survival vary by year or remain constant. If you want to run the IPM with default settings for the current species click the Default Structure button below the inputs to load the default settings listed here:

- Use Population Reconstruction: No
- Reproduction: Time Varying
- Juvenile Survival: Time Varying
- Subadult Survival: Constant (elk only)
- Adult Survival: Constant

If the IPM doesn't run or produces poor results with these defaults you can try adjusting the settings to better match the available data, but keep in mind that at least one of the variables must be constant (rather than time varying) for the IPM to work.

The harvest tab allows you to predict future harvest rates and incorporate those predictions into the IPM. First use the slider at the top to select the range of years for which you want to estimate abundance. If your selected range extends beyond the available harvest data use the additional sliders to specify how many

animals you expect to be harvested by sex. You can use the **Options** button to set the sliders to the min, mean, median, or max from the existing harvest data, or to use the most recent data available (last). You can also click Reset to return to the default (mean) values. Be sure to click on the Harvest tab to the right to preview how your settings will be passed into the IPM.

When you've loaded your data, defined your model structure, and provided future harvest estimates (if needed), move on to the Run tab, specify your MCMC settings, and click Fit Model to run the IPM. If no errors are encountered dismiss the dialog and click **Tables** in the sidebar to view your results. If the model fails to run check the errors section.

4.1.3 Model Output

Results are viewed under the **Plots** and **Tables** subheadings on the left under **IPM**. On the tables tab you can view all of the abundance estimates by year, age class, and sex. The mean estimate, standard deviation, and confidence intervals are reported, along with the BGR diagnostic in the Rhat column. Values greater than 1.1 in this column indicate poor model fit, in which case you may want to adjust the model structure or MCMC settings. Values at or near 1, on the other hand, indicate the model fit the data well.

The **Plots** page shows the estimates plotted over time. Plots are grouped into 4 categories based on the population parameter displayed:

- Abundance
 - Total abundance
 - Female abundance by age class
 - Male abundance by age class
- Ratios
 - Age ratio (young per adult female)
 - Sex ratio (adult male per adult female)
- Survival
 - Female survival by age class
 - Male survival by age class
- Misc
 - Growth Rate ()
 - Maximum Sustained Yield (MSY - shows how growth rate, recruitment, and juvenile survival vary with total abundance)
 - Harvest Rate by sex
 - Harvest by sex

All plots have tools along the right side which allow the user to zoom in on different sections of the graph, hover over specific points to get detailed information, and save/download as an image file.

To generate a report click on **Build Report** in the sidebar, then click Download. When the download is finished click on the saved html file to open it. The report combines information about your settings, input data, and results in a single file. Especially important in the report is the Diagnostics section, which explains how well the model performed and provides guidance on any changes that might improve your results.

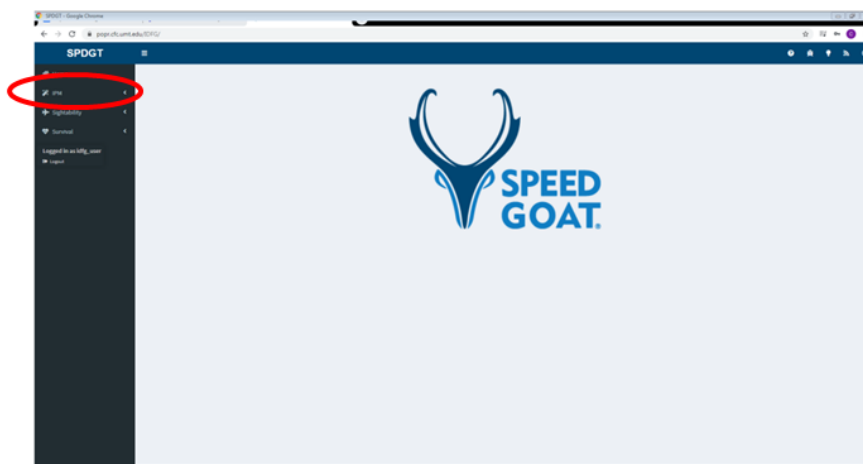
4.1.4 Errors

Reporting errors This section of the documentation is a work in progress. We need your feedback to identify and document the errors users commonly experience running models with PopR - you can help us by sending an email to eric.newkirk@speedgoat.io any time you experience an error on the site. Please include as much information as you can about your settings, data, and the error message you received, including screenshots if possible. Thanks!

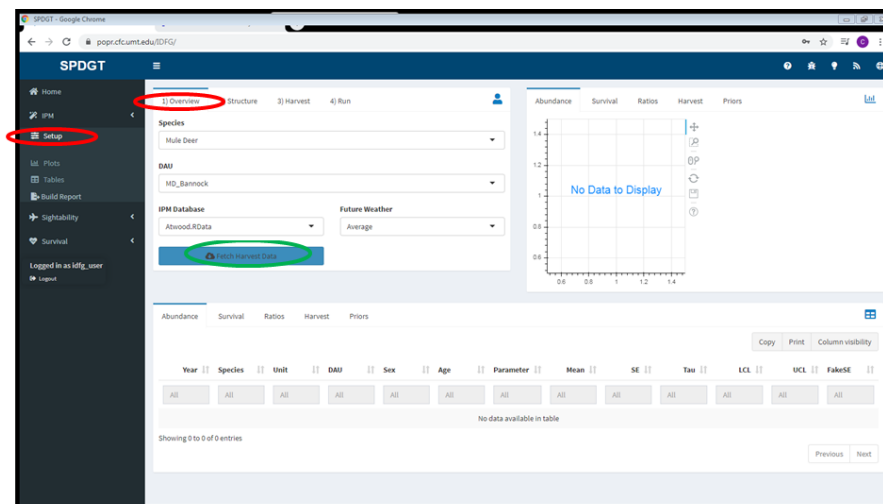
4.2 Step-By-Step IPM Example

This example shows how to run the IPM for mule deer in the Bannock DAU step by step, including screenshots of the website.

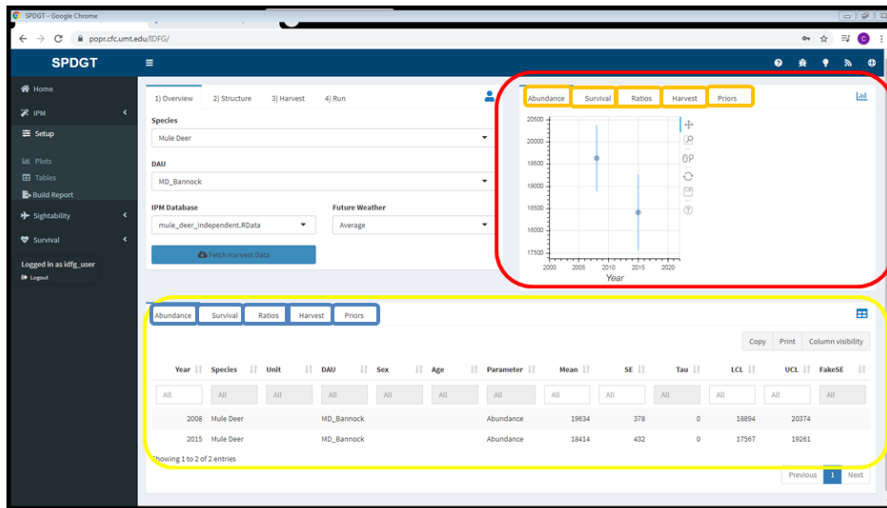
1. Start by clicking **IPM** in the sidebar.



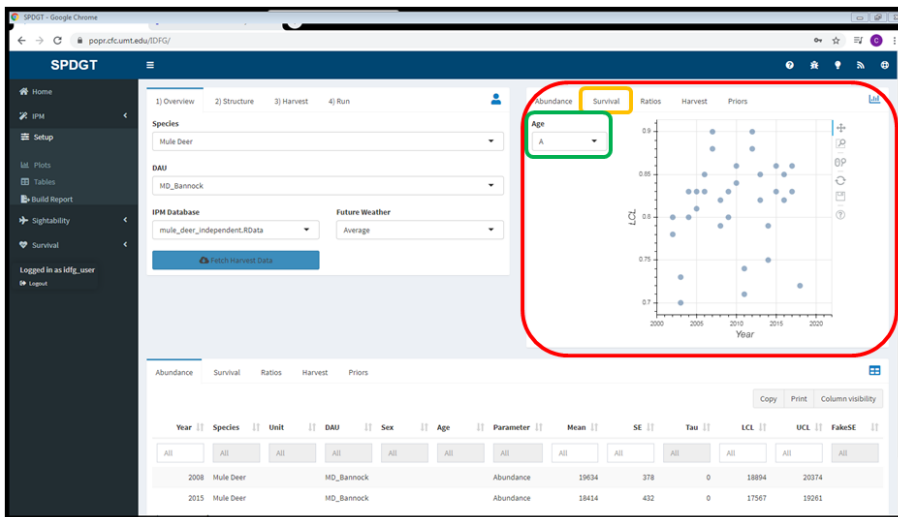
2. Select the **Setup** page in the sidebar, then select the species and DAU you want to model. Select the IPM database containing the estimates you want to use, and choose an option for incorporating [weather covariates](#glm-ipm-weather}. In this example we use mule deer and Bannock. Make sure to select the correct species and DAU before clicking the button to load harvest data.
3. Click the button labeled “Fetch Harvest Data” to load the data for the species and DAU you selected.



4. Once the data is loaded you should see a dialog indicating that the process was successful. Click Dismiss. If you receive an error see the errors section.
5. Review the data in the panes to the right and tables at the bottom (see Loading Data for details).



6. In this example, the abundance data that will be used by the IPM was collected in the Bannock DAU in 2008 and 2015 shown in the graph in the upper right (red rectangle) and in the table at the bottom of the page (yellow rectangle).



7. Move on to the **Structure** tab and click the button to apply the default settings. Your settings should match those in the screenshot below.

SPDGT - Google Chrome

popr.cfc.umt.edu/IDG/

SPDGT

1) Overview 2) Structure 3) Harvest 4) Run

Use Population Reconstruction?

FALSE

Reproduction

Time Varying

Juvenile Survival

Time Varying

Adult Survival

Constant

Default Structure

Abundance Survival Ratios Harvest Priors

Parameter

Recruitment 4

Age J

Sex F

Abundance Survival Ratios Harvest Priors

Copy Print Column visibility

Species	Sex	Age	Parameter	Mean	SE	LCL	UCL
Mule Deer	F	J	Survival	0.58	0.2	0.19	0.97
Mule Deer	F	A	Survival	0.87	0.07	0.73	1.01
Mule Deer	M	J	Survival	0.58	0.2	0.19	0.97

8. Continue to the **Harvest** tab to review the settings there. The defaults were used for this example, so there's no need to change anything.

SPDGT - Google Chrome

popr.cfc.umt.edu/IDG/

SPDGT

1) Overview 2) Structure 3) Harvest 4) Run

Analysis Years

1995 2025

Future Harvest

0

Options

Fit Model

Abundance Survival Ratios Harvest Priors

Parameter

Recruitment 4

Age J

Sex F

Abundance Survival Ratios Harvest Priors

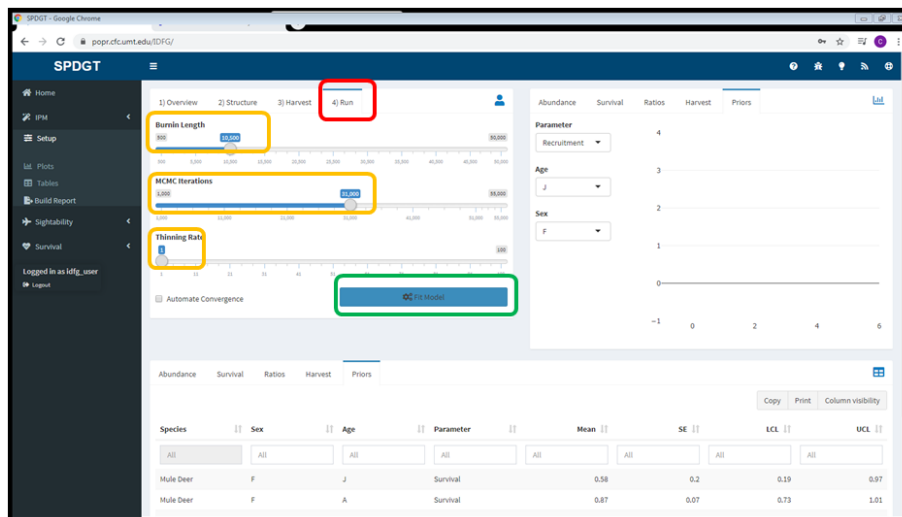
Copy Print Column visibility

Species	Sex	Age	Parameter	Mean	SE	LCL	UCL
Mule Deer	F	J	Survival	0.58	0.2	0.19	0.97
Mule Deer	F	A	Survival	0.87	0.07	0.73	1.01
Mule Deer	M	J	Survival	0.58	0.2	0.19	0.97

9. Now that you've selected the appropriate settings and reviewed the input data, switch to the **Run** tab. Change the burn in slider to 10,500, change the number of iterations to 41,000, and click Fit Model.

4.2. STEP-BY-STEP IPM EXAMPLE

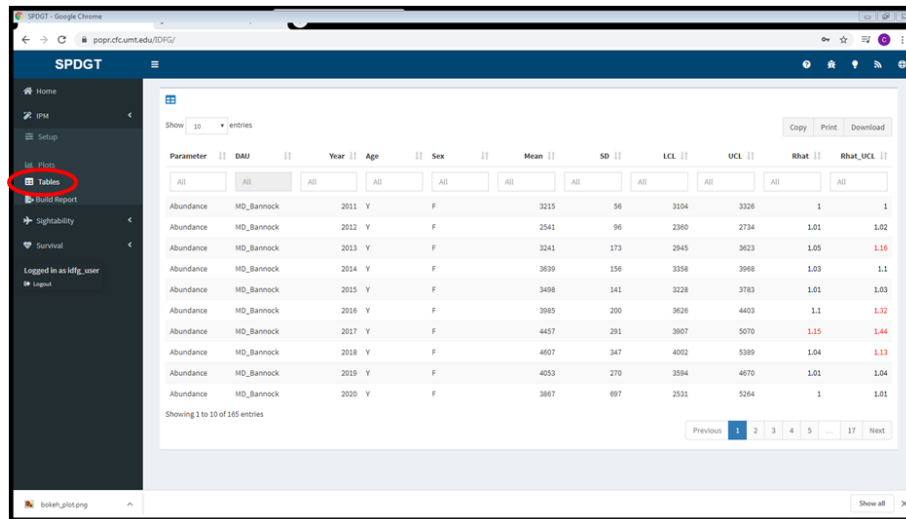
39



- Once model fitting is complete select the **Plots** page in the sidebar to view the output.



- Check the **Tables** page as well to see how the model performed. Note the Rhat values (> 1.1) that appear in red, indicating the model may not have converged.



SPDGT - Google Chrome

popr.cfc.umt.edu/IDG/

SPDGT

Home IPM Setup **Tables** Build Report Sightability Survival

Logged in as lbfj_user

Show 10 entries

Copy Print Download

Parameter	DAU	Year	Age	Sex	Mean	SD	LCL	UCL	Rhat	Rhat_UCL
Abundance	MD_Bannock	2011	Y	F	3215	56	3104	3326	1	1
Abundance	MD_Bannock	2012	Y	F	2541	96	2360	2734	1.01	1.02
Abundance	MD_Bannock	2013	Y	F	3241	173	2945	3623	1.05	1.18
Abundance	MD_Bannock	2014	Y	F	3639	156	3358	3968	1.03	1.1
Abundance	MD_Bannock	2015	Y	F	3498	141	3228	3783	1.01	1.03
Abundance	MD_Bannock	2016	Y	F	3985	200	3626	4403	1.1	1.32
Abundance	MD_Bannock	2017	Y	F	4457	291	3907	5070	1.15	1.44
Abundance	MD_Bannock	2018	Y	F	4607	347	4002	5389	1.04	1.13
Abundance	MD_Bannock	2019	Y	F	4053	270	3594	4670	1.01	1.04
Abundance	MD_Bannock	2020	Y	F	3867	697	2531	5264	1	1.01

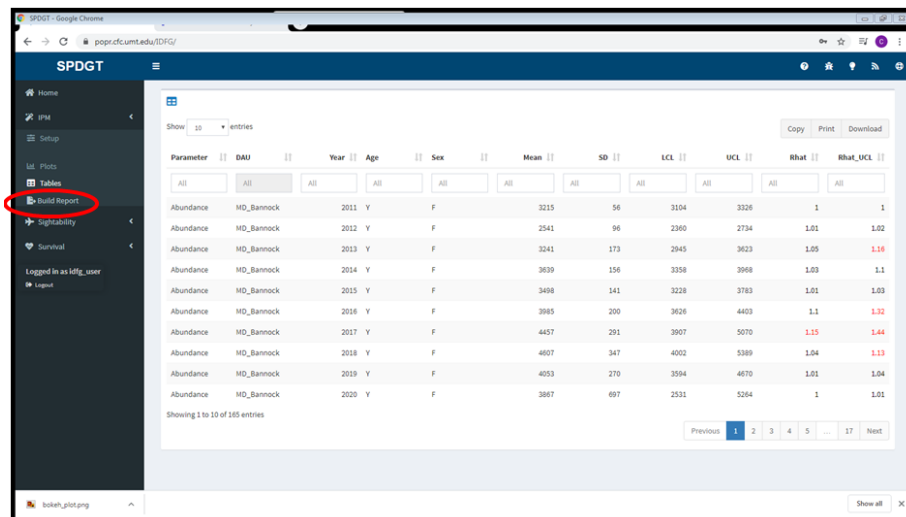
Showing 1 to 10 of 165 entries

Previous 1 2 3 4 5 ... 17 Next

bokeh_plot.png

Show all

12. Now build a report to see more diagnostic information. Click **Build Report** in the sidebar, then click the download button.



SPDGT - Google Chrome

popr.cfc.umt.edu/IDG/

SPDGT

Home IPM Setup **Build Report** Sightability Survival

Logged in as lbfj_user

Show 10 entries

Copy Print Download

Parameter	DAU	Year	Age	Sex	Mean	SD	LCL	UCL	Rhat	Rhat_UCL
Abundance	MD_Bannock	2011	Y	F	3215	56	3104	3326	1	1
Abundance	MD_Bannock	2012	Y	F	2541	96	2360	2734	1.01	1.02
Abundance	MD_Bannock	2013	Y	F	3241	173	2945	3623	1.05	1.18
Abundance	MD_Bannock	2014	Y	F	3639	156	3358	3968	1.03	1.1
Abundance	MD_Bannock	2015	Y	F	3498	141	3228	3783	1.01	1.03
Abundance	MD_Bannock	2016	Y	F	3985	200	3626	4403	1.1	1.32
Abundance	MD_Bannock	2017	Y	F	4457	291	3907	5070	1.15	1.44
Abundance	MD_Bannock	2018	Y	F	4607	347	4002	5389	1.04	1.13
Abundance	MD_Bannock	2019	Y	F	4053	270	3594	4670	1.01	1.04
Abundance	MD_Bannock	2020	Y	F	3867	697	2531	5264	1	1.01

Showing 1 to 10 of 165 entries

Previous 1 2 3 4 5 ... 17 Next

bokeh_plot.png

Show all

Download Report

Choose Report
Type
Analysis Report

Download

Dismiss

Parameter	DAO	Year	Age	SD	LCL	UCL	Rhat	Rhat_UCL
Abundance	WD_Bannock	2011	Y	56	3194	3326	1	1
Abundance	WD_Bannock	2012	Y	56	2360	2734	1.01	1.02
Abundance	WD_Bannock	2013	Y	56	2940	3623	1.05	1.16
Abundance	WD_Bannock	2014	Y	56	3338	3968	1.03	1.1
Abundance	WD_Bannock	2015	Y	56	3228	3783	1.01	1.03
Abundance	WD_Bannock	2016	Y	56	3626	4403	1.1	1.12
Abundance	WD_Bannock	2017	Y	56	3907	5070	1.13	1.44
Abundance	WD_Bannock	2018	Y	56	4002	5389	1.04	1.13
Abundance	WD_Bannock	2019	Y	56	3584	4670	1.01	1.04
Abundance	WD_Bannock	2020	Y	56	5264	5264	1	1.01

Showing 1 to 10 of 10 entries

13. Notice the diagnostic section in red, indicating that the model did not converge. In some cases this may indicate problems with the data or the model structure, but sometimes we just need to let the model run longer.

Idaho Fish & Game

IPM Analysis Report

August 10, 2020

Model Definition

- Species: Mule Deer
- DAO: WD_Bannock
- Years: 2011 - 2021
- Database: mule_deer_independent

Structure

- Adult Survival: Constant
- Juvenile Survival: Time Varying
- Reproduction: Time Varying

Model Fitting

- MCMC Burnin: 10,500
- MCMC Iterations: 41,000
- Thinning Rate: 1

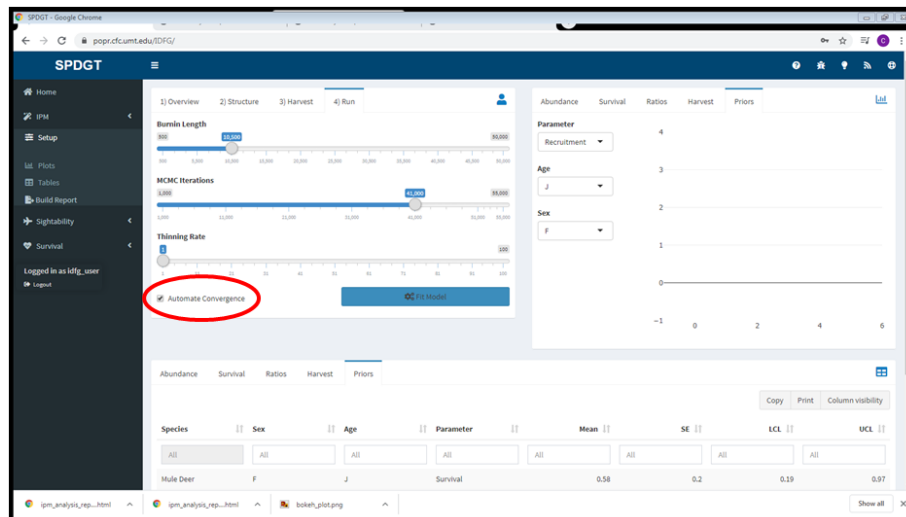
Diagnostics

The Brooks, Gelman and Rubin convergence diagnostic (Rhat) is a test used to determine whether the analysis has found a reliable solution. Values < 1.1 generally suggest that the model was run for a sufficient number of iterations while values > 1.1 suggest that the model should be run for more iterations. The maximum values of the point estimate and upper CI are reported below.

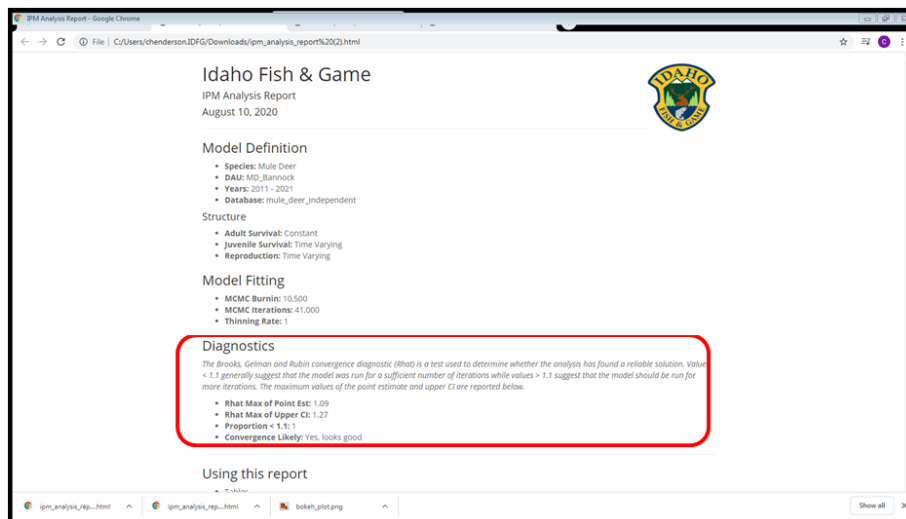
- Rhat Max of Point Est: 1.46
- Rhat Max of Upper CI: 2.18
- Proportion < 1.1: 0.71
- Convergence Likely: Absolutely Not!

Using this report

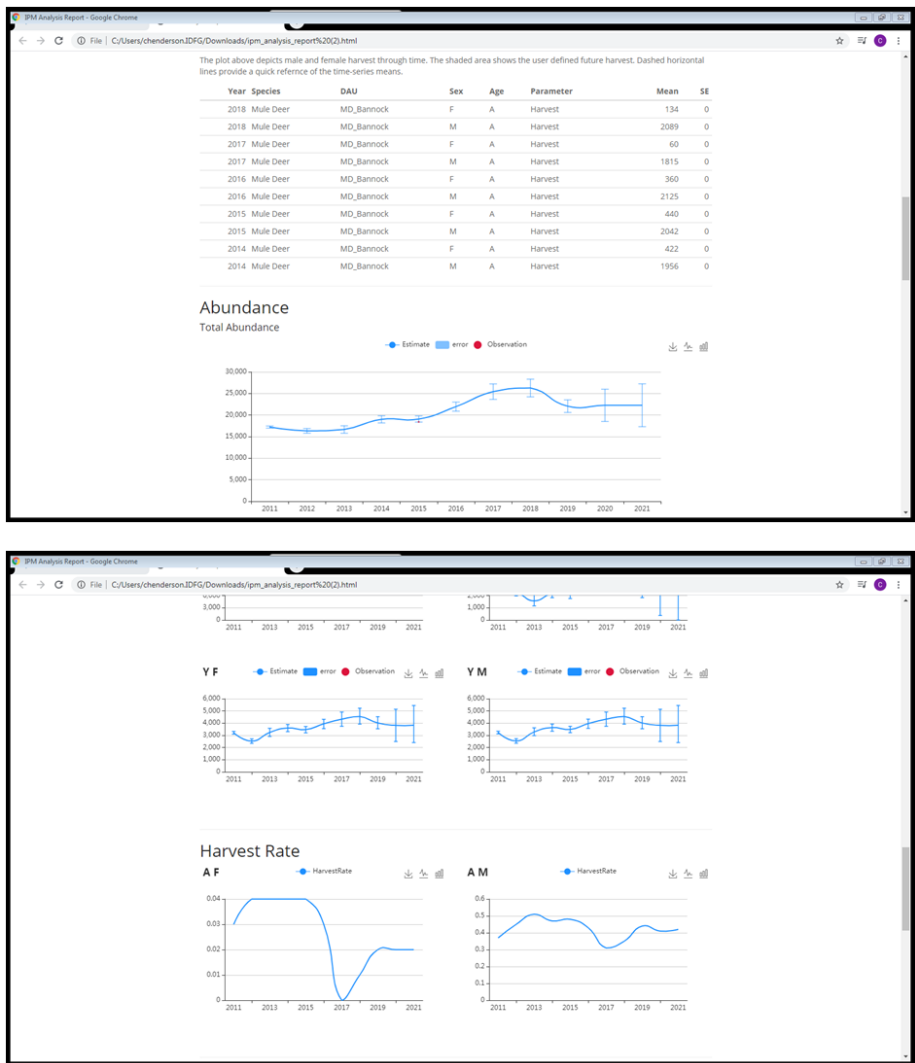
14. Return to the Run tab on the **setup** page and click the checkbox to automate convergence, then run the model again.



15. Build a new report and check the diagnostics section again. Allowing the model to run as long as needed resulted in much better performance.



16. Check out the tables and plots in the rest of the report to review your results.



Chapter 5

Glossary

5.1 General

Age Classes Age classes must be defined for each species to capture variability in parameters like survival and harvest rate, as well as to determine when animals begin to reproduce. Regardless of actual birth date, animals transition to the next age class on December 15, the model anniversary used in the IPM. Assuming calves and fawns are born approximately mid-June each year, the age classes translate to roughly the following:

- Mule Deer
 - A (Adult) - 1.5 years and up
 - J (Juvenile) - 6 months to 1.5 years
 - Y (Young of Year) - birth to 6 months
- Elk
 - A (Adult) - 2.5 years and up
 - S (SubAdult) - 1.5 to 2.5 years
 - J (Juvenile) - 6 months to 1.5 years
 - Y (Young of Year) - birth to 6 months

Bio Year Since winter surveys typically span a time period including two years the term BioYear consolidates the survey period into a single year. For example, surveys can start in December 2019 and run through January 2020 so the BioYear for this survey is 2020. The term was created for easier record keeping and does not have a biological definition. The biological year begins on December 15 and was intended to correctly align data collection with biological processes. Estimates from the models operate on this biological year.

DAU (Data Analysis Unit) Recent location data from mule deer fitted with GPS radio-collars provided a more informed understanding of how mule deer populations are distributed across Idaho. In light of this new information, IDFG revised the geographical boundaries of the data analysis units (DAU) that are used for mule deer population monitoring and management and incorporated these changes into the IPM. Some aerial surveys can no longer be viewed in the IPM because they don't include abundance for all of the GMUs in the new DAU. Past aerial surveys were incorporated into the site through a series of conversations with area biologists. Questions about the data can be directed to Paul Atwood or Josh Nowak.

SAMM (AKA SAD) SAMM, or Statewide Animal Monitoring Master, is a database hosted by IDFG for housing monitoring data. Data is read into PopR from SAMM via API, but SAMM data cannot be edited through PopR. In order to add or edit data the changes must be made in SAMM.

5.2 Sightability

Aircraft/Model There are separate elk sightability models depending on the aircraft used for the survey. The aircraft types include a Hiller 12E, Bell 47G, and Super Cub. Mule deer sightability models do not have separate models for different types of aircraft.

Area A generic term used to catch the species specific differences in management area naming. Instead of elk zones and deer pmu/dau we simply use area to describe the spatial extent we are focused on for a particular analysis.

Composition Survey This survey type is used to gather data related to age and sex ratios and typically samples a smaller portion of the population in an area of interest. See also: Sightability Survey

Demographic The demographic field in the sightability output table denotes which age and sex classes are described by the estimated values. Classes are based on sex and age class and vary by species and survey type.

Group size Refers to the number of individuals in a single observed group. The size of the group impacts the probability of detection and influences how the model estimates the true number of individuals. Group sizes include 1, 2, 3, 4, 5, 6, 7-15, 16-30, and >30.

Movement Categories The 3 movement categories are bedded, standing, and moving. This is based on the most active individual when initially observed. The movement categories impact the probability of detection and influence how the model estimates the true number of individuals.

Obs The number of observations in a given strata.

Pop The population of sampling units in a given strata. In other words, the total number of subunits in a strata.

PropSampled This is the proportion of units sampled in a given strata. This can be calculated by obs/pop , divide the number of subunits sampled in the strata by the total number of subunits in the strata.

Sightability Survey An intensive aerial survey methodology to estimate the entire population size in a given area. Data, in the form of raw counts, from this survey are analyzed using a sightability model to estimate true population size from direct counts. See also: Composition Survey

Snow The 3 snow cover categories are 0-20%, 21-79%, and >80%. The snow cover categories impact the probability of detection and influences how the model estimates the true number of individuals.

Stratum The landscape is divided into separate strata based on the probability of being occupied by the species being surveyed. Typically subunits are classified as high, medium, or low strata based on their suitability as habitat for a particular species. Other strata should only be used after consulting an expert and in special circumstances when you do not want that subunit to be used to extrapolate to subunits that are not flown.

Subunit Used to define the proportion of the unit that was sampled. For example, if a unit consists of 10 subunits and the survey samples 3 of the 10 subunits then $3/10$ or $.3$ of the unit was sampled. This is used to extrapolate sightability estimates such that if this same survey produced a count of 100 then the estimate is $100/.3$ or 333 animals in the unit.

Unit Synonymous with GMU.

Variance The variance components of the sightability output are of interest to the user because they provide clues as to how to improve their surveys.

- **ModelVar:** Model variance
- **SampleVar:** Sample variance
- **SightVar:** Sightability model variance
- **TotalVar:** Total variance is the sum of the model, sample, and sightability variances. The square root of the total variance is the standard deviation of the estimated value.

Vegetation There are 5 vegetation class categories including grass/open/agriculture, sagebrush, juniper/mtn. mahogany, mtn. brush/aspen, and conifer. Vegetation class represents the dominant vegetation within a 30m buffer around where an individual deer or group of deer is observed. The vegetation class categories impact the probability of detection and influences how the model estimates the true number of individuals. PopR has data columns for all habitat data collected on IDFG elk and mule deer sightability and composition survey forms. IDFG elk aerial survey data sheets contain only % screen and that is the only habitat variable used in the elk sightability model. IDFG mule deer aerial survey and composition data sheets contain data columns for % screen, screen type, and veg type and veg type is the only habitat variable currently used in the mule deer sightability model.

5.3 Survival

Known Fate Survival Model Survival models attempt to estimate the probability that an animal alive at this time will be alive at some future time. A known-fate model does this in part by assuming that the fate of every animal is known. This name was chosen early in the development process. As time has passed we have tweaked and developed the model to relax assumptions and better meet the needs of our collaborators. The full model description is beyond the scope of this Glossary, but know that the Known-Fate model selection means that the results of running a model will show the user estimates of survival and mortality. This is in contrast to the multi-state models where multiple sources of mortality can be estimated, such as death by harvest or death by predator. See also: Multi-state Survival Model

Multi-state Survival Model The multi-state survival model allows users to estimate survival, harvest mortality rate and other mortality rate. The different estimates sum to 1 and so to be comparable to the known-fate models users should take 1 - other mortality as equivalent to survival in the known-fate context. See also: Known Fate Survival Model

5.4 IPM

IPM Database IPM database files contain the estimates from survival and sightability models, and are used when running the IPM. As of September 2020 there are 3 IPM database files in use for IDFG:

1. Atwood.RData – DO NOT USE, this database is for development and research purposes only
2. mule_deer_independent.RData – This database will use survival data from the DAU you are modeling over multiple years when there is not any survival data for the DAU and the year in the database. Use this database if the DAU you are modeling has a large amount of data available and survival is relatively constant between years.
3. mule_deer_shared.RData - Strongly recommended. This database will use survival data from other DAU's in the state for the current year when there is not any survival data for the DAU and year in the database. Use this database if the DAU you are modeling has little to no data available or if survival is highly variable between years.

Future Weather Future weather covariates were developed using Mark Hurley's dissertation and the analyses contained therein.

- Average: Mean observed weather conditions
- Good: Upper 95% quantile of observed weather
- Bad: Lower 5% quantile of observed weather
- None: Do not use weather data covariates in the analysis

5.5 MCMC

Automate Convergence Users may choose to simply check the “Automate Convergence” box below the MCMC sliders menu in the PopR interface. This option will assure that an adequate Burn-in Length and number of MCMC Iterations have been used to produce a statistically sound estimate and error distribution.

Burnin Length “Burn-in” is a term for an initial process that gives the Markov Chain time to approach the solution to the problem by throwing away some less reasonable starting points at the beginning of a Markov Chain Monte Carlo run. In PopR, managers should simply use the default Burn-in Length setting when developing an estimate through the standard user interface.

Iterations PopR uses Markov Chain Monte Carlo (MCMC) methods to “fit” IPM population estimates to the available data. MCMC methods estimate parameters in complex models by systematically updating informed prior distributions with information gleaned from field data (e.g. observed harvest). Therefore, they allow us to describe each parameter in terms of a distribution and that distribution’s shape. Typically, 25,000-100,000 MCMC iterations will be required to fit an IPM. If the number of MCMC iterations is set too low the uncertainty about an estimate is likely to be misrepresented.

Brooks-Gelman-Rubin Statistic (\hat{R}) In PopR, we use the Brooks-Gelman-Rubin (BGR) statistic, often represented as \hat{R} or Rhat, as an assessment of convergence and the criteria used when the “Automate Convergence” option is used. The BGR statistic suggests convergence when estimates of \hat{R} are below 1.1 or more generally close to 1. This statistic is reported in the output tables for most models and highlighted in red when \hat{R} estimates are above 1.1. The default settings will produce results that are unlikely to change even if run longer, but users should increase the number of MCMC iterations if \hat{R} estimates are above 1.1 and computing time allows.

Thinning Rate Thinning tells the sampler to only retain every n th value from the chains. This technique is sometimes used to reduce autocorrelation in the chains, but comes at the cost of reduced efficiency of the sampler. A more reasonable use of thinning is when hardware limitations are being reached, which typically comes in the form of running out of memory. This will not be an issue in PopR and, therefore, the recommended setting for the Thinning slider is 1.