gibbonsecr v1.0 manual

April 16, 2016

Contents

1 Overview

The `gibbonsecr` software package uses **Spatially Explicit Capture–Recapture (SECR)** methods to estimate the density of gibbon groups from acoustic survey data. This manual begins with a brief introduction to the theory behind SECR and then describes the main components of the user interface.

Click **here** to download a `.pdf` version of the manual.

---

2 Introduction to SECR

Over the past decade SECR has become an increasingly popular tool for wildlife population assessment and has been used to analyse survey data for a wide range of taxa. The main advantage of SECR over traditional capture-recapture techniques is that it allows direct estimation of population **density** rather than abundance. Traditional capture-recapture methods can only provide density estimates through the use of separate estimates or assumptions about the size of the sampled area. In SECR however, density is estimated directly from the survey data by using information contained in the pattern of the recaptures about the spatial location of animals. By extracting spatial information in this way, SECR does not require the exact locations of the detected animals to be known in advance.

2.1 Basic setup

The basic data collection setup for an SECR analysis consists of a spatial array of **detectors**. Detectors come in a variety of different forms, including traps which physically detain the animals, and **proximity detectors** which do not. The use of proximity detectors makes it possible for an animal to be detected at more than one detector (i.e. *recaptured*) during a single sampling occasion.

The plot below shows a hypothetical array of proximity detectors, with black squares representing detections of the same animal (or the same group in the case of gibbon surveys) and grey squares representing no detections.
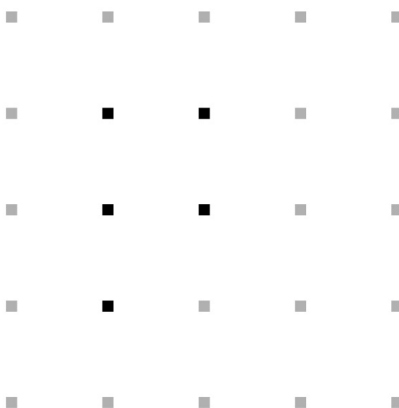

Figure 1: Hypothetical array

The pattern of the detections for this group (i.e. the pattern of the recapture data) contains information about its true location; an intuitive guess would be that the true location is somewhere near the cluster of
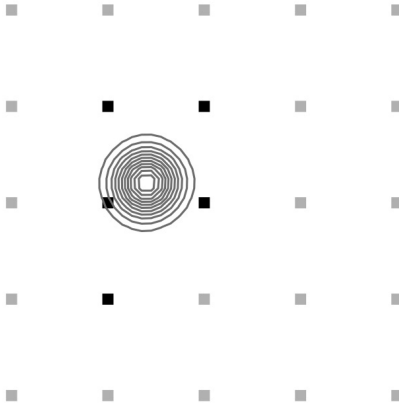
Figure 2: Location contours using recapture data

black detectors. The plot below shows a set of probability contours for this unknown location, given the recapture data.

In the case of acoustic gibbon surveys the listening posts can be treated as proximity detectors and the same logic can be applied to obtain information on the locations of the detected groups. However, the design shown in the figure above would obviously be impractical for gibbon surveys. The next figure shows probability contours for a more realistic array of listening posts where a group has been detected at two of the posts.
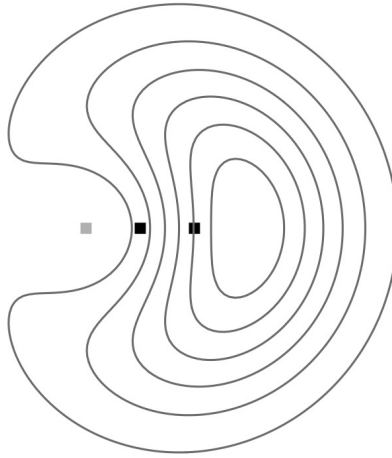
Figure 3: Location contours using recapture data

The main conclusion here is that using smaller arrays of detectors results in less information about the unknown locations.

2.2 Estimated bearings

SECR also allows **supplementary information** on group location to be included in the analysis in addition to the recapture data, for example in the form of **estimated bearings** to the detected groups. The next figure illustrates how taking account of information contained in the estimated bearings can provide better quality information on unknown locations.
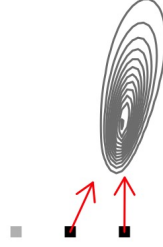
Figure 4: *Location contours using recapture data and estimated bearings*

Using estimated bearings in this way can lead to density estimates that are less biased and more precise than using recapture data alone (although the degree of improvement will depend on the precision of the bearing esitmates). Since the precision of bearing estimates is usually unknown, SECR methods need to estimate it from the data. This requires the choice of a **bearing error distribution**. The figure below shows two common choices of distribution for modelling bearing errors – the **von Mises** and the **wrapped Cauchy**. The colour of the lines in these plots indicate the value of the precision parameter (which needs to be estimated from the survey data). In both cases, the larger precision parameter, the greater the precision. The mean error is shown by the dotted vertical lines and is assumed to be zero (i.e. the bearing estimates are assumed to be unbiased).
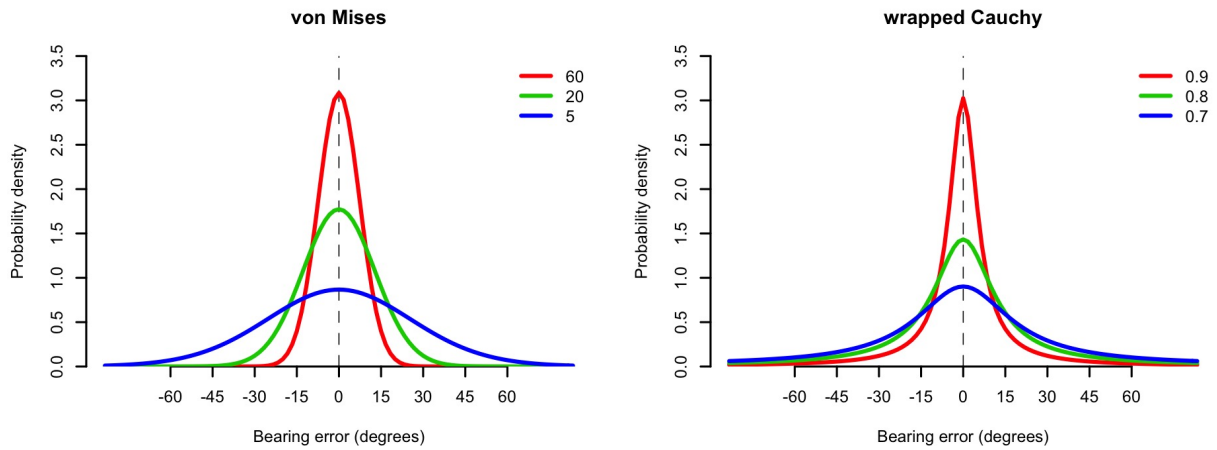
Figure 5: Bearing error distributions

2.3   Estimated distances

Estimated distances to detected groups can also be used to decrease the bias and improve the precision of the density estimate. Again, the degree of improvement will be subject to the precision of the distance estimates. The figure below shows probability contours for the true location of a hypothetical detection when estimated distances are used (but not estiated bearings).
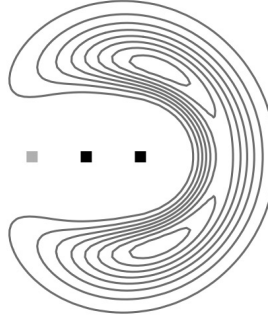
Figure 6: Location contours using recapture data and estimated distances

The figure below shows probability contours for the true location for the same detection using both estimated distances and estimated bearings.
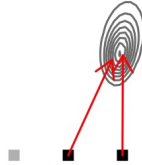
Figure 7: Location contours using recapture data, estimated bearings and estimated distances

Two choices of distribution are available in `gibbonsecr` for modelling estimated distances: **gamma** and **log-normal**. The plots below illustrate how the value of the parameter of these distributions (which needs to be estimated from the survey data) relates to precision of the distance estimates. The dashed vertical lines indicate the true distance.

Note that, for a given parameter value, the precision of the distance estimates decreases as the true distance increases.

2.4   Detection functions

Another key feature of SECR is that the probability of detecting a (calling) gibbon group at a given location is modelled as a function of distance between the group and the listening post. This function – referred to as the **detection function** – is typically assumed to belong to one of two main types of function: the **half normal** or the **hazard rate**. The specific shape of the detection function depends on the value of its parameters, which need to be estimated from the survey data. The half normal has two parameters: g0 and

Figure 8: Estimated distances distributions

`sigma`: the `g0` parameter gives the detection probability when the distance between group and detector is zero, and the `sigma` parameter controls the width of the function. The hazard rate has three parameters: `g0`, `sigma` and `z`. The `g0` and `sigma` parameters have the same interpretation as for the half normal, and the additional parameter `z` controls th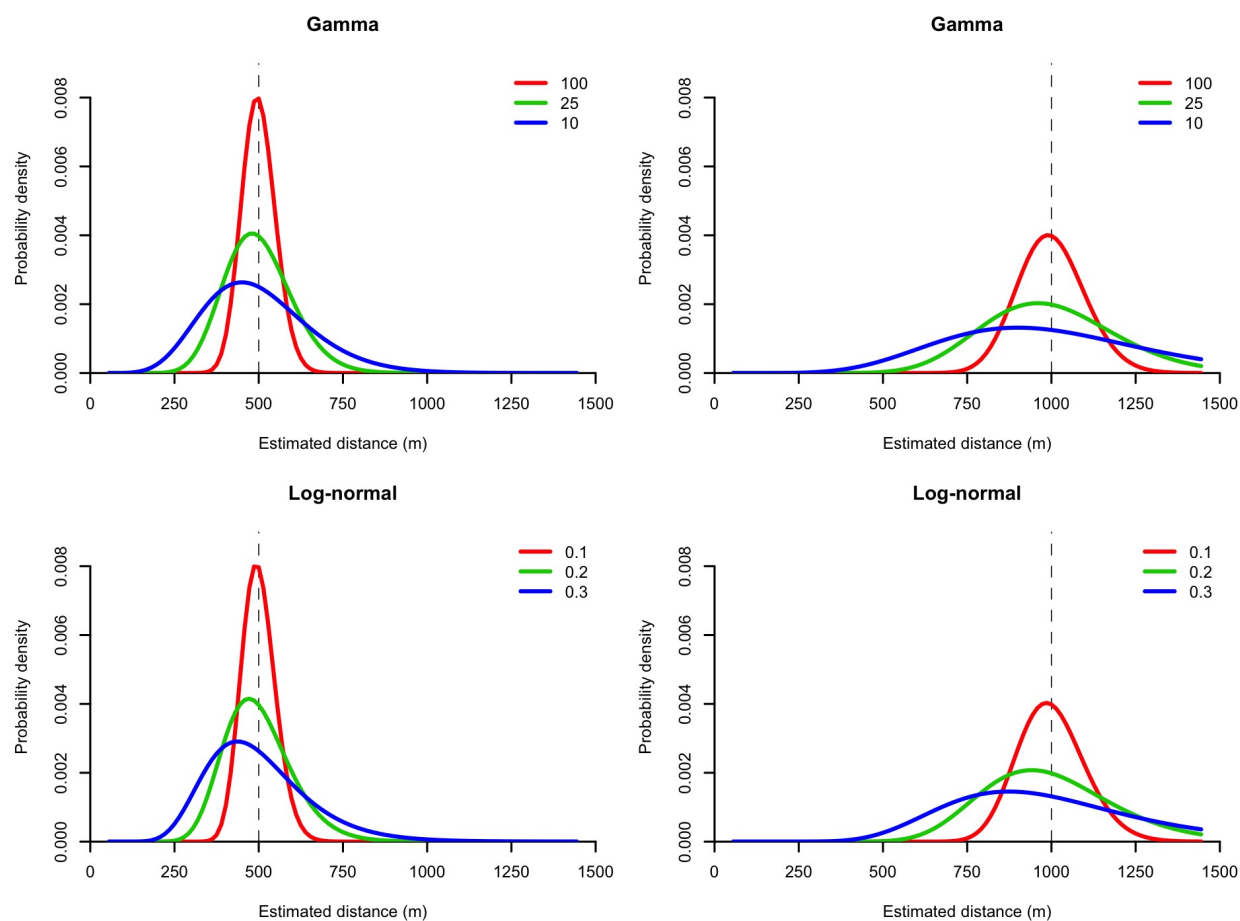e shape of the 'shoulder' and allows a greater degree of flexibility. The figure below illustrates the shape of these detection functions for a range of parameter values.
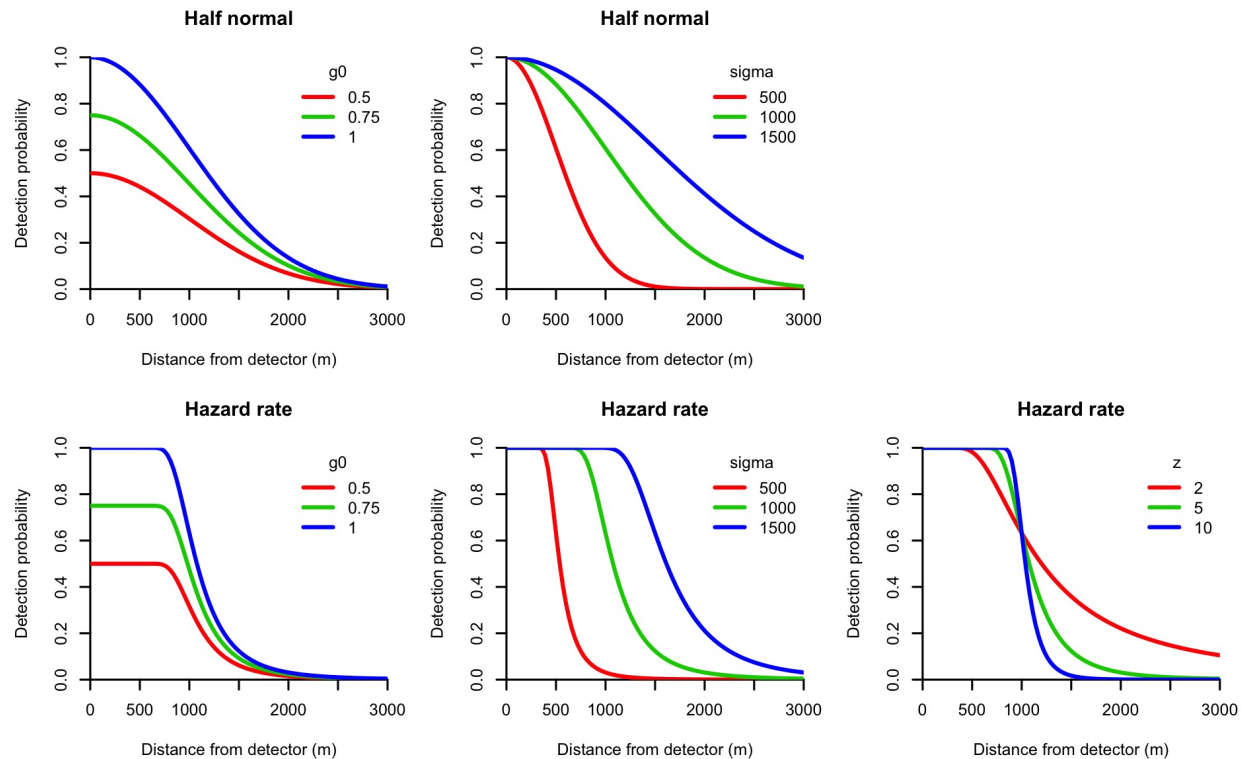
Figure 9: Detection functions

2.5   Detection surface

Associating a detection function with each listening post allows us to calculate the *overall* probability of detection – i.e. the probability of being detected at *at least one* listening post – for any given location. The figure below illustrates this concept using a heat map of a **detection surface** for a single array of listening posts, with color indicating the overall detection probability.

The region near the centre of this surface is close to the listening post array and has the highest detection probability. In this case, an group located close to the detectors will almost certainly be detected. The detection probability decreases as distance from the detectors increases.

2.6   Effective sampling area

The shape of the detection surface is related to the size of the **effective sampling area**. Since the region close to the detectors has a very high detection probability, most groups within this region will be detected and it will therefore be sampled almost perfectly. However, regions where the detection probability is less than 1 will not be completely sampled as some groups in these areas will be missed. The figure below illustrates this idea for a series of arbitrary detection surfaces.
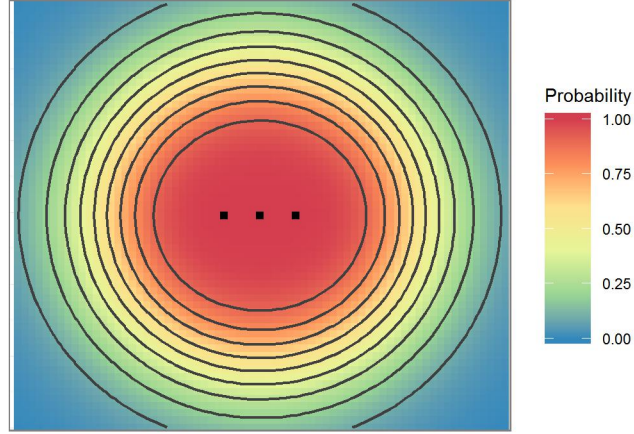
Figure 10: Dectction surface
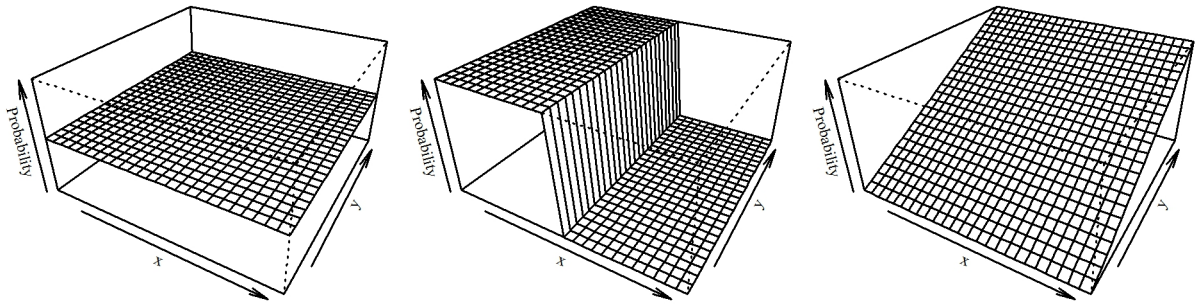
Figure 11: Arbitrary detection surfaces with the same effective sampling area

The first plot in this figure shows a flat surface where the detection probability is 0.5 everywhere. In this scenario every group has a 50% chance of being detected. If the area covered by the surface was 10km^2, then the *effective* sampling area would be 10km^2 x 0.5 = 5km^2. Using this detection process we would expect to detect the same number of groups as we would if we had perfectly sampled an area of 5km^2. In the second plot, half of the area is sampled perfectly and the other half is not sampled at all, so this has the same effective sampling area as the first plot. The third plot has a detection gradient and isn't as intuitive to interpret. However, the general to calculate the effective survey area is to calculate the *volume* under the detection surface. The third plot has the same volume as the other two, so it has the same effective area.

---

3   Loading the software

TODO

3.1   Launch from R

3.1.1   Install R

Make sure you have the latest version of **R** installed.

Download R for Windows

Download R for Mac

Optionally you can also install **RStudio** which is a more user-friendly interface to R and (e.g. it supports syntax highlighting and auto-completion).

Download RStudio

3.1.2   Install prerequisites

TODO

3.1.3   Install gibbonsecr

TODO

3.1.4   Launch the interface

Once everything is installed you can launch the user interface by opening R (or RStudio) and typing the following lines into the console.

```
library(gibbonsecr)
gui()
```

3.2   Launch from a desktop icon

TODO

3.2.1 Download the files

TODO

3.2.2 Make a shortcut icon

TODO

---

4 Data import

The first step in conducting an analysis is to import your survey data. This is done via the **Data** tab, shown in the screenshot below.
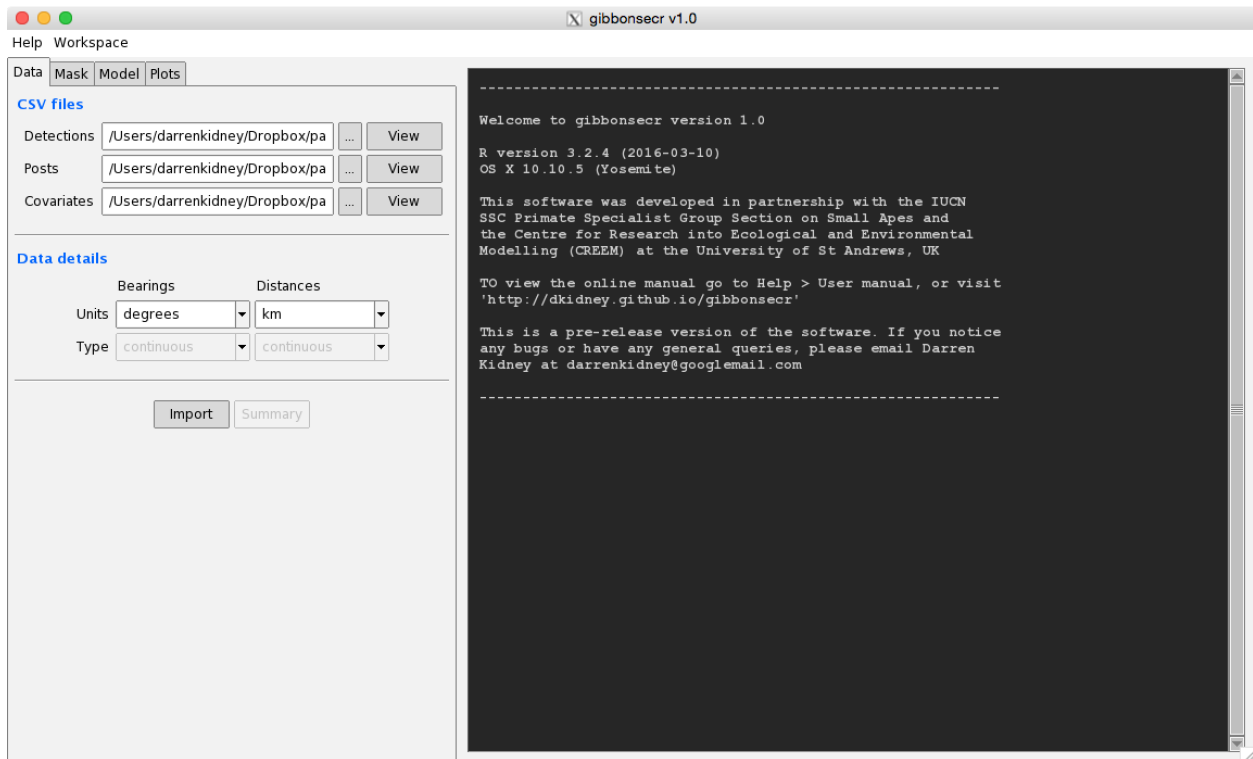
Figure 12: Data tab

4.1 CSV files

As a minimum requirement you need to prepare a **detections** file and a **posts** file, both of which must be in .csv format. You can also include an optional **covariates** file (which also needs to be in .csv format). Advice on how to structure these files is given below. The file paths to your data files can be entered manually into the text entry boxes in the *CSV files* section, or you can navigate to the file path using the . . . button.

4.1.1 Detections

The detections file needs to contain a record of each detection, with one row per detection. For example, if group 1 was recorded at listening posts A and B then this would count as two detections. This file needs to have the following columns:

- **array** – Id for the array
- **occasion** – Id for day of the survey (typically an integer from 1 to 4)
- **post** – Id for the listening post
- **group** – Id for the detected group
- **bearing** – Estimated bearing to the detected group
- **distance** – Estimated distance to the deteted group (optional)

The table below shows example entries for detections made at an individual array during a one-day (i.e. single-occasion) survey.

Table 1: Example detections data

array	occasion	post	group	bearing
6	1	A	6_A	170
6	1	B	6_B	192
6	1	B	6_A	180
6	1	B	6_C	40
6	1	C	6_B	220
6	1	C	6_C	36

4.1.2 Posts

The posts file contains information on the location and usage of the listening posts. This file needs to have one row per listening post and should contain the following columns:

- **array** – Id for the array
- **post** – Id for the listening post
- **x** – Longitude coordinate (in metric units)
- **y** – Latitude coordinate (in metric units)
- **usage** – Indicator showing the sampling days on which the posts were manned E.g. if on a 3-day survey a particular post was manned on day 1 and day 3 but for some reason not on day 2, then `101` would be entered in the usage column for that post. The number of digits in the usage column should be equal to the number of sampling days for the array.

The table below shows example entries for posts at one array during a one-day survey.

Table 2: Example posts data

array	post	x	y	usage
6	A	690085	1557174	1
6	B	690570	1557163	1
6	C	691082	1557128	1

4.1.3 Covariates

The covariates file should contains values for any other variables associated with the survey data. This file needs to have one row per samping day for each listening post and should contain at least the following columns:

- **array** – Id for the array
- **post** – Id for the listening post
- **occasion** – Id for day of the survey (typically an integer from 1 to 4)

These columns can all be used as covariates themselves, but any additional covariates should be added using additional columns. Use underscores _ instead of full stops for the covariate names.

The table below shows example covariates at a single array for a one-day survey.

Table 3: Example covariates data

array	post	occasion	month	observer
6	A	1	June	Rod
6	B	1	June	Jane
6	C	1	June	Freddy

4.2 Data details

Once the paths to the `.csv` files have been entered, select the appropriate units from the *Data details* dropdown boxes for your estimated bearings data (and estimated distances data if it was collected). Note that the current version of the software only allows `Type` to be set to `continuous` since interval methods for bearings and distances have not yet been implemented.

4.3 Data buttons

Once the paths to your `.csv` files have been entered and the data units have been checked, click the `Import` button to import the data. If the import is successful then a summary print out should appear in the output window on the right hand side. You can re-print this summary at any time by pressing the `Summary` button.

Example output following a successful import is shown below.

```
Detections:
        array  2 3 4  5 6 7 10 11 12 13 15 16 17 total
       ngroups 9 1 4  7 3 2  9  6 10  7 10  5  4    77
    ndetections 16 3 8 12 6 3 14  9 14 12 14  6  6   123
         ndays 1 1 1  1 1 1  1  1  1  1  1  1  1    13
         nposts 3 3 3  3 3 3  3  3  3  3  3  3  3    39


Recaptures = 53.2 %


Auxilliary data:
          Units    Type        Details
 bearings degrees continuous range: 11 to 360


Covariates:
```

```
Level Name        Type
array 'array'     category
      'month'     category
post  'habitat'   category
      'observer'  category
Details
13 levels: 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 15,...
2 levels: July, June
2 levels: primary, secondary
3 levels: Freddy, Jane, Rod
```

---

5   Making a mask

The SECR model fitting procedure requires the use of a **mask**, which is a fine grid of latitude and longitude coordinates around each array of listening posts. When an SECR model is fitted, the mask is used to provide a set of plausible candidate locations for each detected group. It is important to select a suitable mask to avoid unreliable results.
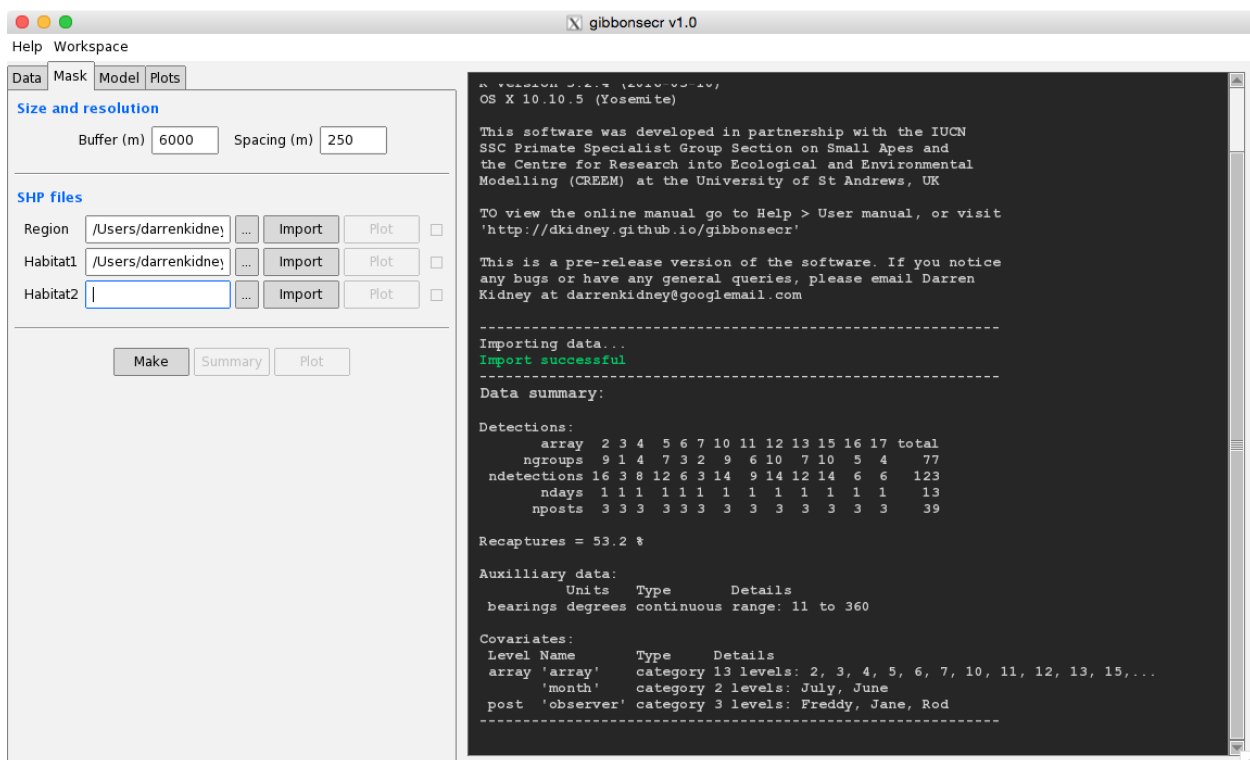

Figure 13: Mask tab

5.1   Size and resolution

There are two main settings you need to consider when defining a mask – the **buffer** and the **spacing** – which you can specify in the *Mask* tab.

5.1.1 Buffer

The buffer defines the maximum distance between each mask point and the closest listening post. It needs to be large enough so that the region it encompasses contains all plausible locations for the detected groups, but it shouldn't be unnecessarily large. Buffer distances that are too small will lead to underestimates of the effective sampling area and overestimates of density. However, increasing the buffer distance also increases the number of mask points, which means that the models will take longer to run, so the buffer also shouldn't be larger than it needs to be. **The ideal buffer distance is the distance at which the overall detection probability drops to zero**.

A good way to check whether the buffer distance is large enough is to look at the detection surface, which you can plot after fitting a model (see the section on plotting results). The detection surface plot produced by `gibbonsecr` is the same size as the mask, so the colour at the edge of the plot will show you the overall detection probability at the buffer distance. If the detection probability is greater than zero at the buffer distance then you should increase the buffer distance, re-fit the model and re-check the detection surface plot.

To illustrate this issue, the figure below shows a series of detection surfaces from models that were fitted using mask buffers of 1000m, 10000m and 5000m respectively.
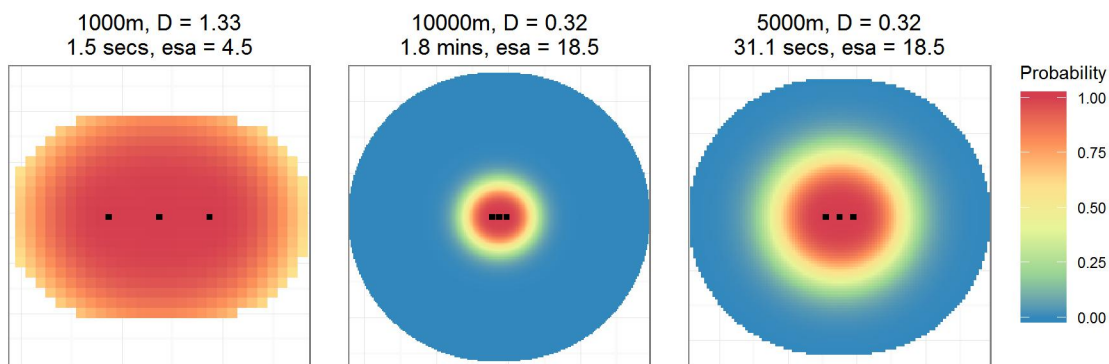

Figure 14: Choosing a mask buffer

The buffer in the first plot is too small as the detection probability at the edge of the mask is much greater than zero. In this case it is therefore extremely likely that the true locations of some of the detected groups will be outside the buffer zone. The buffer in the second plot is ten times greater than the buffer in the first and the detection probability at the edge of the mask is zero. We would expect the density estimate to be unbiased and in the density estimate is around 75% lower than the estimate in plot 1 (which suggests that the estimate in plot 1 is a severe overestimate). The buffer in the third plot is intermediate between the other two. The detection probability is still zero at the buffer distance, but the estimated density is very similar to plot 2, so it doesn't look to be biased. The computation time was also much quicker than for plot 2. In this case the mask in the third plot would be preferred.

5.1.2 Spacing

The buffer spacing is the distance between adjacent mask points. Decreasing the spacing will increase the resolution and increase the total number of mask points. Smaller spacings therefore provide a greater number of candidate locations and lead to more reliable results. However, increasing the number of mask points has a cost in terms of computing time and if the spacing is too small then models may take a very long time to run. As a general rule of thumb, try to use the smallest spacing that is practical given the speed of your computer, but try not to use spacings larger than 250m.

5.2 Shapefiles

TODO

5.2.1 Region

TODO

5.2.2 Habitat

TODO

5.3 Mask buttons

TODO

Example output following a successful import is shown below.

```
Summary stats:
 buffer     6000 m
 spacing    250 m
 npoints    2000 (average per array)
 area       125 sq km (average per array)

Covariates:
 Name       Type      Details
 'habitat'  category  2 levels: primary, secondary
```

---

6 Model fitting

Once you have made a mask you can start fitting some SECR models using the **Model** tab, shown below.

Model specification is split into two steps: (i) choosing what kind of detection function and bearing error distribution you want to use, and (ii) deciding whether to fix any parameter values or model them using the available covariates. These steps are described in more detail below.

6.1 Model options

The first section in the Model tab contains dropdown boxes where you can choose between different detection functions and different distributions for the estimated bearings and distances.

- **Detection function** – choose either the half normal (the default) or the hazard rate
- **Bearings distrbution** – choose from the von Mises (the default), the wrapped Cauchy, or none
- **Distances distribution** – choose from the gamma (the default), the log-normal, or none

Setting the bearings/distances distribution to `none` means that the bearings/distances data will be ignored in the analysis. Setting both bearings and distances distributions to `none` will result in a conventional SECR model being fitted using only the recapture data.
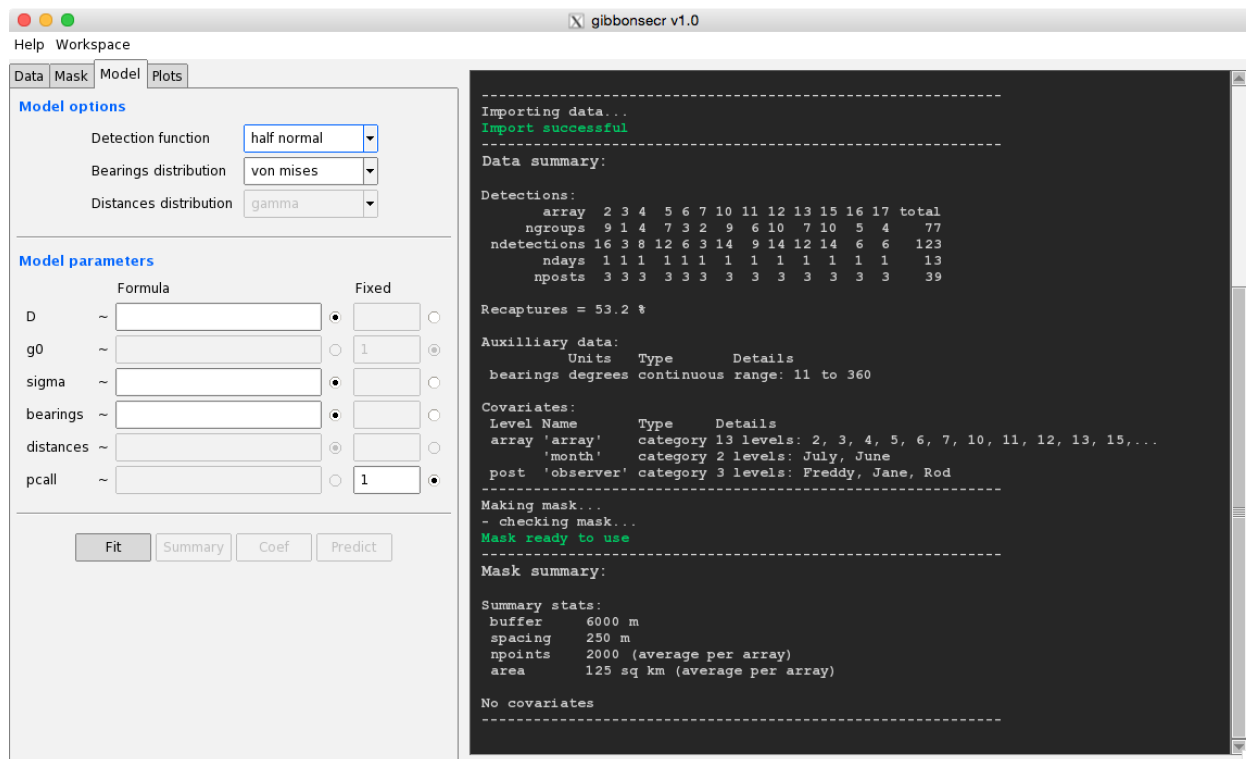
Figure 15: Model tab

6.2   Model parameters

The next section in the *Model* tab provides various options for refining your model. Each row in this section relates to a particular parameter in the SECR model.

- **D** – group density (number of groups per square kilometre)
- **g0** – detection function intercept parameter
- **sigma** – detection function scale parameter
- **z** – detection function shape parameter (hazard rate only)
- **bearings** – bearing estimates distribution parameter
- **distances** – distances estimates distribution parameter
- **pcall** – group calling probabiity (per day)

Hovering the cursor over the row labels on the user interface will open a temporary help box giving a reminder of these definitions.

6.2.1   Formulas

If you wish to estimate a particular parameter in your analysis then you need to make sure that the *Formula* entry box for that parameter is activated by clicking on the radio button on the right hand side of the entry box. If the Formula entry box is activated but left blank then a single coefficient for that parameter will be estimated (i.e. an intercept-only model). If you wish to specify a formula for a particular parameter you need to click the radio button on the right hand side of the formula entry box for that parameter, then type the names of the covariates you wish to use into the entry, separated by + signs. E.g. to model the `sigma` parameter using habitat and weather you would type the following into the Formula box for `sigma`:

17

```
habitat + weather
```

A note to experienced R-users: As well as `+` you can also the `*` and `:` operators to specify formulas. You can also use gam functions `s`, `te`, `ti` and `t2` (from the `mgcv` package) for numeric variables. However, the use of `as.factor` and `as.numeric` to coerce variables, and `-1` to change the model contrasts, is not supported.

6.2.2 Covariate levels

Covariates can be used when constructing formulas for the various model parameters using the model argument in the gfit function. However, some parameters can only be modelled with certain types of covariates.

Covariates can be divided into the following types, according to the level at which they vary:

- **array-level** - These are consistent within arrays. All listening posts and sampling occasions for a given array will have the same value for an array-level covariate, but values vary between arrays. For example, season is likely to be an array-level covariate, since it is highly unlikely that consecutive sampling occasions at a given array will overlap with more than one season.

- **trap-level** - These are consistent within listening posts for a given array. All sampling occasions for a given listening post will have the same value for a trap-level covariate. For example, habitat and elevation are likely to be trap-level covariates since they may vary between listening posts but will be consistent across occasions.)

- **occasion-level** - These are consistent within sampling occasions for a given array. All listening posts in an array will have the same value for an occasion-level covariate. For example, weather is likely to be an occasion-level covariate since it will probably be the same for all listening posts for a given occasion at a given array but may vary between occasions.

- **trap-occasion-level** - These can vary between listening posts and between sampling occasions. For example, observer might be a trap-occasion-level covariate since observers will (always) vary between listening posts for a given occasion at a given array, and they may also vary between occasions for a given listening post at a given array.

- **mask-level** - These are spatial covariates and can vary across mask points. Mask-level covariates from GIS polygon and point files can be imported and attached to mask objects using the import_shp and addCovariates functions.

The table below shows which covariates can be used to model which parameters.

Parameter	Covariate level
D	array, mask
g0	array, trap, occasion, trap-occasion
sigma	array, trap, occasion, trap-occasion
z	array, trap, occasion, trap-occasion
bearings	array, trap, occasion, trap-occasion
distances	array, trap, occasion, trap-occasion
pcall	array, occasion

6.2.3 Fixing parameter values

Sometimes you many not want or need to estimate a particular parameter, in which case you can fix its value. To do this, click on the radio button on the right hand side of the *Fixed* entry box and type the value of the

parameter in the box.

A general point to bear in mind when fixing any parameter is that it will generally lead to a more precise density estimate (i.e. one with narrower confidence intervals). If the fixed parameter is known with a high degree of certainty then this would be a desirable effect. However, if there is uncertainty over the true value of that parameter (e.g. you may have used an estimate from a previous study) then this will not be incorporated into the SECR results and the precision of the density estimate will be overestimated (i.e. the confidence intervals generated by the software will be too narrow).

6.2.4 Estimating g0

For one-day surveys the only option allowed by the software is to fix `g0` at 1. This is because when the listening post is zero distance from a calling group the probability of detecting it is extraordinarily unlikely to be anything other than 1. (Remember that the `g0` parameter gives the detection probability for a calling group at zero distance from the listening post).

For multi-day surveys however, the movement of groups between consecutive sampling days means that we have to redefine group 'location' as being the *average* location of the group. As a result, `g0` needs to be reinterpreted as the probability of detecting a calling group at zero distance from the *average* location. In this case it is much more likely that the detection probability for a calling group whose average location is zero distance from the listening post will be less than one. This is because a group is unlikely to always be at its average location during a multi-day survey (unless it happens not to move). For multi-day survey data it is therefore a good idea to estimate `g0`.

6.2.5 Estimating calling probability

For one-day surveys, `pcall` can't be estimated so the only option is to provide a fixed value. By default, `pcall` is fixed at 1 for one-day surveys, which means that the `D` parameter can be interpreted as the density of *calling* groups, rather than the density of groups. However, the software allows you to change this value (e.g. you may have prior knowledge of the calling probability for the study species) in which case the density parameter can be reinterpreted as the density of groups. For one-day surveys, changing the `pcall` value will result in a direct scaling of the density estimate. For example, if you had an estimated calling group density of 5, changing the fixed value for `pcall` to 0.5 and re-fitting the model would result in a group density estimate of 10.

For multi-day surveys there are three possible options for dealing with the `pcall` parameter when analysing your data.

1. **Estimate `pcall` from the survey data** – This requires that the survey data contain *temporal recaptures* – i.e. the group IDs should indicate which groups were detected on more than one survey day. The temporal recapture data needs to be reliable for this method to work. Note that the SECR software considers data from each array independently, so you only need to identify the temporal recaptures within arrays (and not between arrays).

2. **Fix `pcall` to a known value** – E.g. using data from a previous survey. Bear in mind that fixing `pcall` to a value that is too low will result in a overestimate of group density. Furthermore, remember that fixing the value of pcall is likely to lead to an overestimate in the *precision* of your group density estimate.

3. **Fix `pcall` to 1 and estimate calling group density** – This can be done by treat the data from each survey day independently. The only way to do this currently is to modify the `array` and `occasion` columns in your data files before you import the data. The values in the `array` columns should be edited so that the IDs are all unique – for example, array "1" for occasion 1 could be re-labelled "1_1", array "1" for occasion 2 could be re-labelled "1_2", etc. All entries in the `occasion` columns should be set to 1, and all entries in the `usage` column of the traps file should be set to "1". By doing this

the SECR software will treat each occasion independently and ignore any temporal recaptures. should be set so that the software treats the data for each array separately. Note that future versions of the software will hopefully automate this process.

6.3 Model buttons

There are four button at the botton of the *Model* tab:

1. **Fit** – Use this button to fit the model once you are happy with your model specification. Once your model is fitted a summary will appear in the output console.
2. **Summary** – Use this button to reprint the model summary
3. **Coef** – This prints a summary of the raw model coefficients
4. **Predict** – This allows you to predict the value of each model parameter. If you have used covariates in model formulas a pop-up box will appear allowing you to choose specific values of covariates (e.g. you may wish to predict density for all levels of a habitat covariates)

---

7 Plotting results

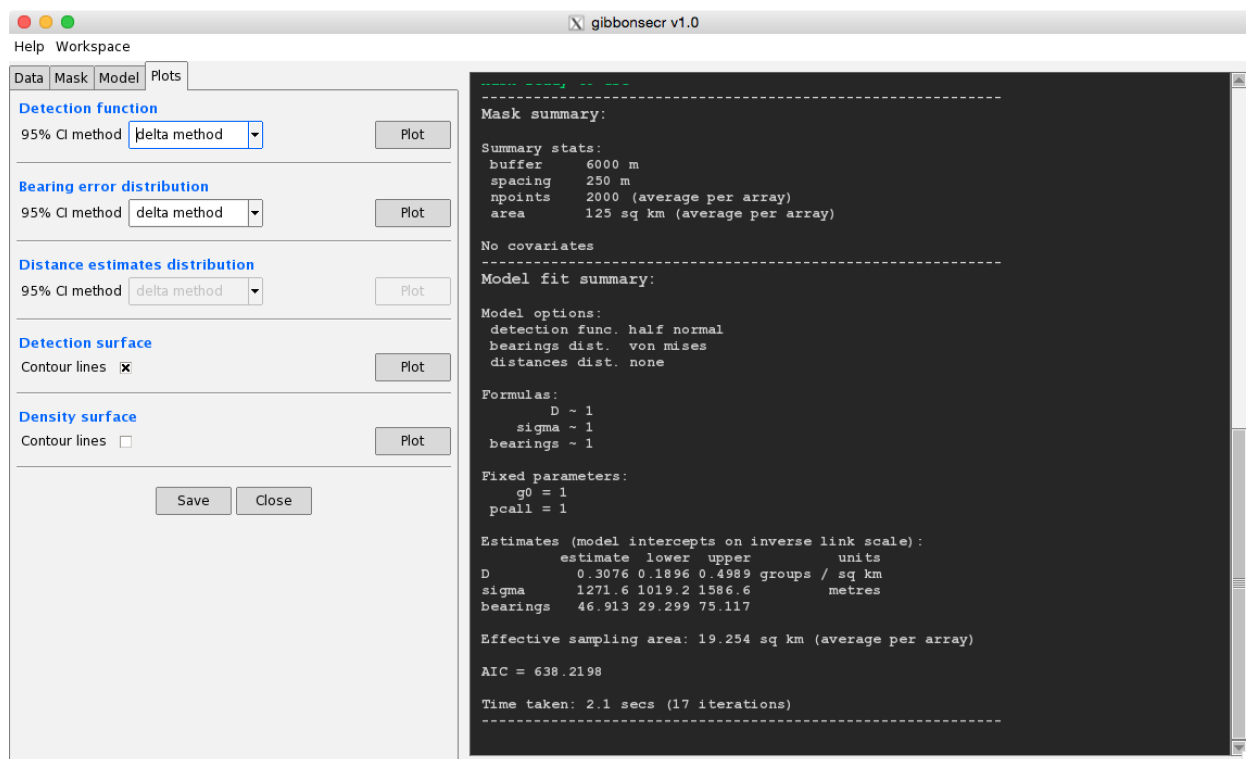Once a model has been fitted, the results can be plotted from the **Plots** tab, shown below.

Figure 16: Plots tab

This tab currently has five plotting options:

1. **Detection function** - always available

2. **Bearing error distribution** - only available if a bearings model has been fitted
3. **Distance estimates distribution** - only available if a distances model has been fitted
4. **Detection Surface** - always available
5. **Density surface** - always available

The first three plots have the following options for plotting confidence intervals:

- `none` - no intervals plotted
- `delta method` - confidence intervals calculated using a numerical approximation – this method is quicker to run than the bootstrap, but may be unreliable for more complex models
- `bootstrap` - confidence intervals caluated using a parametric bootstrap (resampling from the estimated sampling distributions of the model coefficients) using 999 resamples – this method may take a little longer to run but should be more reliable than the delta method.

The detection surface and density surface plots also require the selection of an array. The detection surface plot plots individual arrays, whereas the density surface plot needs to use a single set of array-level covariates.

7.1 Detection function

Example plots of fitted `half normal` detection functions are shown below, with and without confidence intervals.
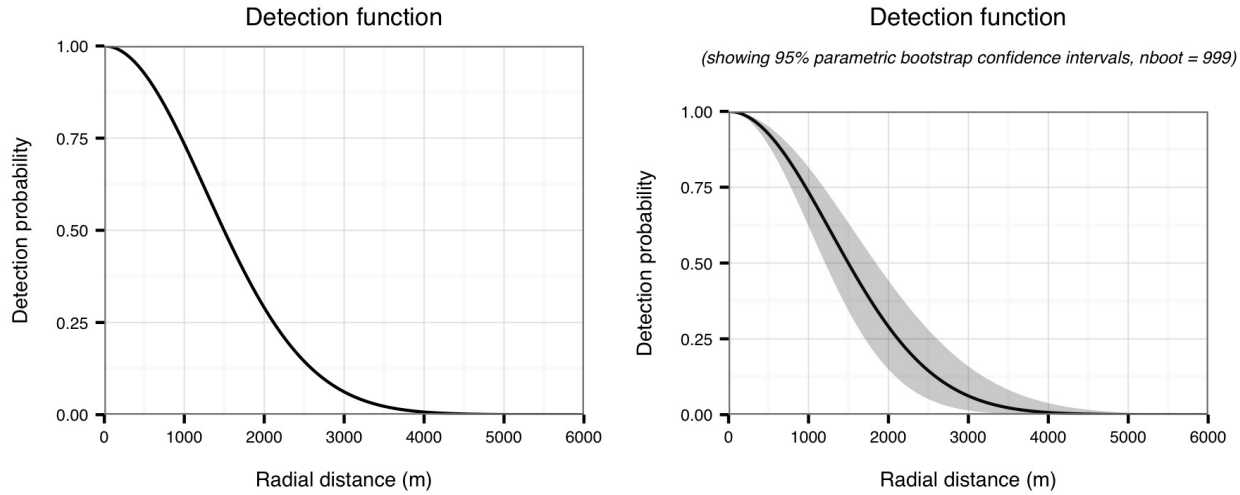

Figure 17: Example estimated detection function

7.2 Estimated bearings distribution

Example plots of estimated `von Mises` bearing error distributions are shown below, with and without confidence intervals.

7.3 Estimated distances distribution

Example plots of estimated `gamma` estimated distances distributions are shown below, with and without confidence intervals (using an example true distance of 500m).
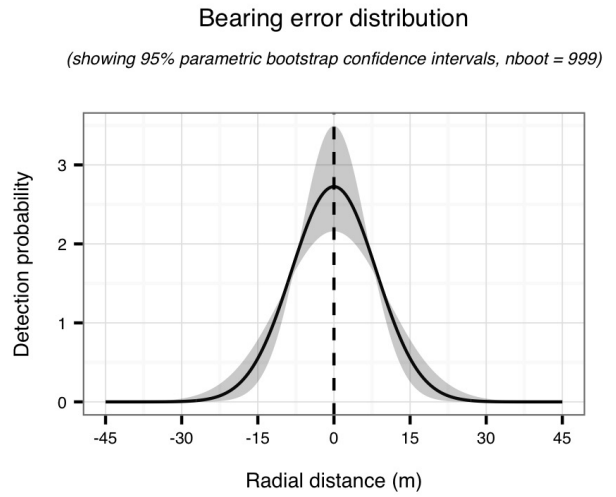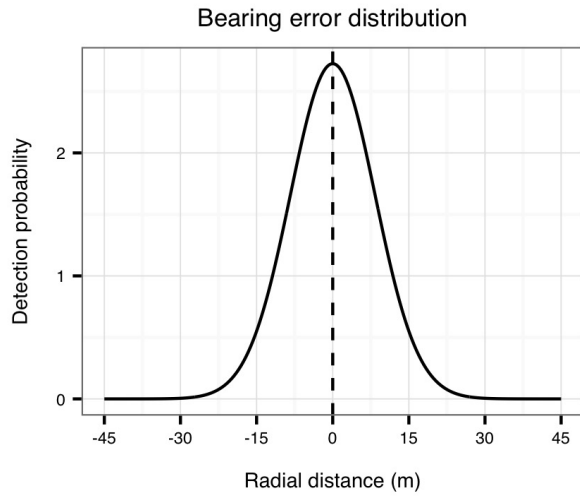
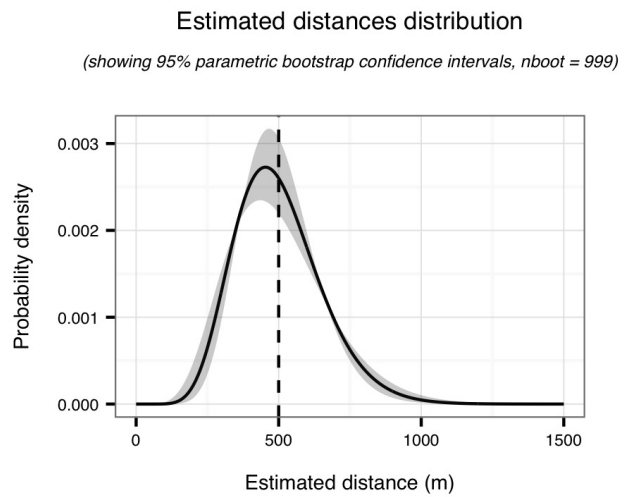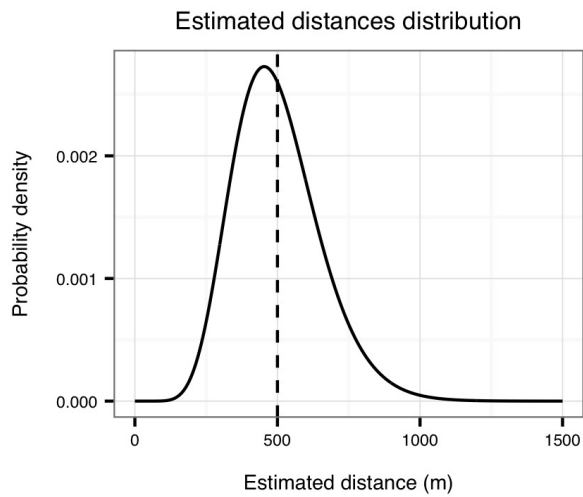Figure 18: Example estimated distribution for bearing errors

Figure 19: Example estimated distribution for estimated distances

7.4 Detection surface

An example plot of a estimated detecton surface using `half normal` detection function is shown below.
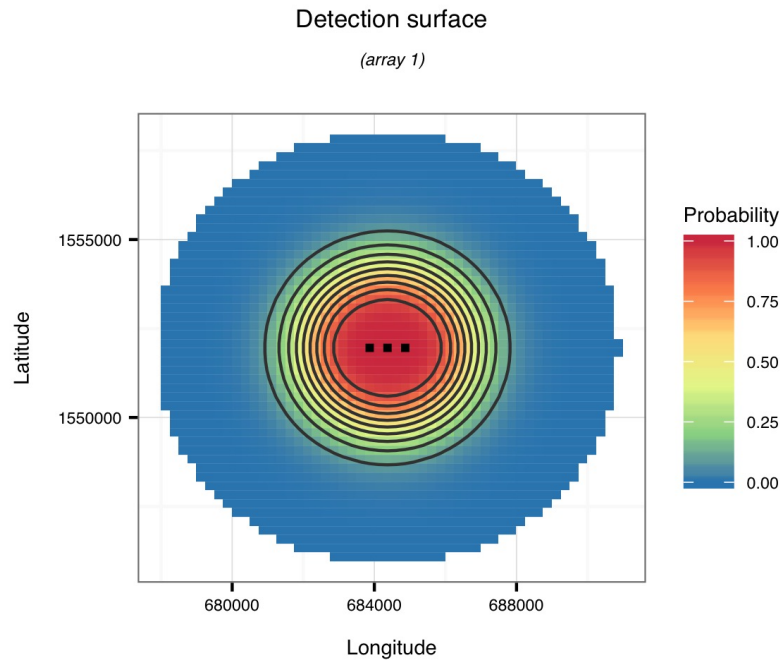

Figure 20: Example estimated detection surface

7.5 Density surface

An example plot of a estimated density surface using a cubic regression spline of lattitude and longitude – `D ~ s(x, y, k = 4)` – is shown below.

---

8 Model selection

An important element of statistical modelling is choosing a preferred model from a number of candidate models. For example, you may get a slightly different density estimate when using the hazard rate detection function instead of the half normal detection function. How do you decide which model, and therefore which density estimate, should be preferred?

8.1 AIC

A common way of choosing between competing models is to use something called the **AIC score**. This is a number that can be calculated from the fitted model which tries to measure how well the model balances having a good fit to the data whilst not being overly complex. The AIC score can be found at the bottom of the model summary printout (which is displayed after pressing the model *Summary* button at the bottom of the *Model* tab).

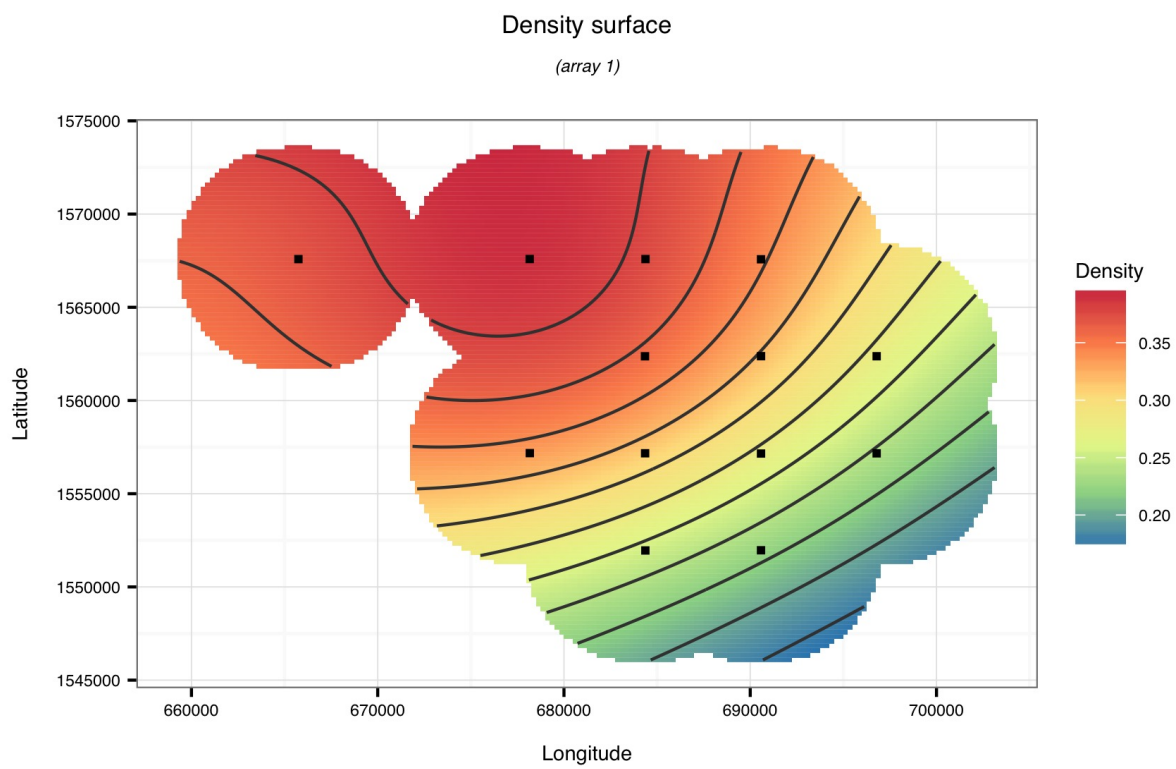When using AIC it is important to bear in mind the following points:

Figure 21: Example estimated density surface

- **Lower AIC scores are preferred.** For example, if model A has an AIC score of 10, and model B has a score of 100, then model A would be preferred to model B.

- **Negative scores are preferred to postitive scores.** For example, if model A has an AIC score of -10 and model B has a score of -100, then model B would be preferred to model A.

- **Diffences of 2 or more are meaningful.** This is a general rule-of-thumb for deciding between models with different number of coefficients. If the difference is less than this then you should be conservative and keep the model with the fewest coefficients. For example suppose you fitted model A with an intercept-only formula for density (i.e. no covariates) which had an AIC of 100. Then suppose you fitted a model B with habitat as a covariate for density (e.g. `D ~ habitat`) which had an AIC of 99. The difference is only -1 which isn't a big difference, and according to the rule of thumb the difference isn't large enough to prefer model B over model A. Now imagine that you fitted a third model, model C, using altitude as a covariate for density (e.g. `D ~ altitude`) which had an AIC of 95. The difference between model A and model C is 5, so according to the rule of thumb model C would be preferred.

- **Only models fitted to the same data can be compared using AIC.** For example, you could use AIC to help you decide whether or not to use the von Mises distribution or the Wrapped Cauchy distribution to model the bearing errors, because both models would have been fitted using the estimated bearings data. However you could *not* use AIC to compare two models where one used a bearing error distribution and one used no bearing error distribution, since the estimated bearings data would be used in the first model but ignored in the second models (the precision of the density estimate might be a better to make a decision in this case). For a similar reason, you also should *not* use AIC to compare models fitted using different masks.

- **The magnitude of the AIC score tells you nothing about how *good* a model is.** The difference in AIC between two competing models helps you decide which one is better, but they might both be poor models. Unfortunatley on of the disadvantages of SECR is that there is currently no means for calculating model goodness-of-fit statistics.

8.2   Model plausibility

Whilst AIC can be extremely useful it shouldn't be used blindly and you should also ensure that any preferred model is also *plausible*. For example, model A might have a lower AIC score than model B, but if model A looks entirely unrealistic (e.g. given your knowledge of the study system) then you should discard it. For example, a fitted bearing error distribution which implied that errors as large as 180 degrees were highly probable might be ignored if such an outcome is known to be very unlikely under normal field conditions.

---

9   Menu Options

9.1   The Help menu

9.1.1   Example Data

This contains a link to data from a single-day survey of *Nomascus annamensis*. Clicking the link should automatically load this data into the interface. It might be a good idea to explore this data first and have a quick practice at fitting models before importing you own data.

9.2 The Workspace menu

9.2.1 Setting the working directory

This option allows you to specify a home directory for your project. Once it has been set, this will be the default directory when browsing for files and saving workspaces.

9.2.2 Saving and loading workspaces

These options allows you to save the current session, including the mask and model, and re-load it for later use. E.g. it may take a while to make a high resolution mask (especially if you import gis shapefiles) or fit a model with several covariates and saving the workspace means you won't have to repeat those steps every time you re-open the software. Bear in mind that if you want to save a series of models you should save them in separate workspaces (since each workspace contains a maximum of one fitted model).

---

10 Reporting bugs

TODO