# Fit of data to Truncated Poisson Distribution

Press the "Next>" button at the bottom of the form to find the truncated Poisson mean sighting frequency distribution (crimson line with triangles) which best fits the observed sightings frequency distribution (blue bars). The best fit truncated Poisson mean (in this case 2.64) is used to estimate the maximum number of sightings allowable before an animal should be treated as "trap-happy".
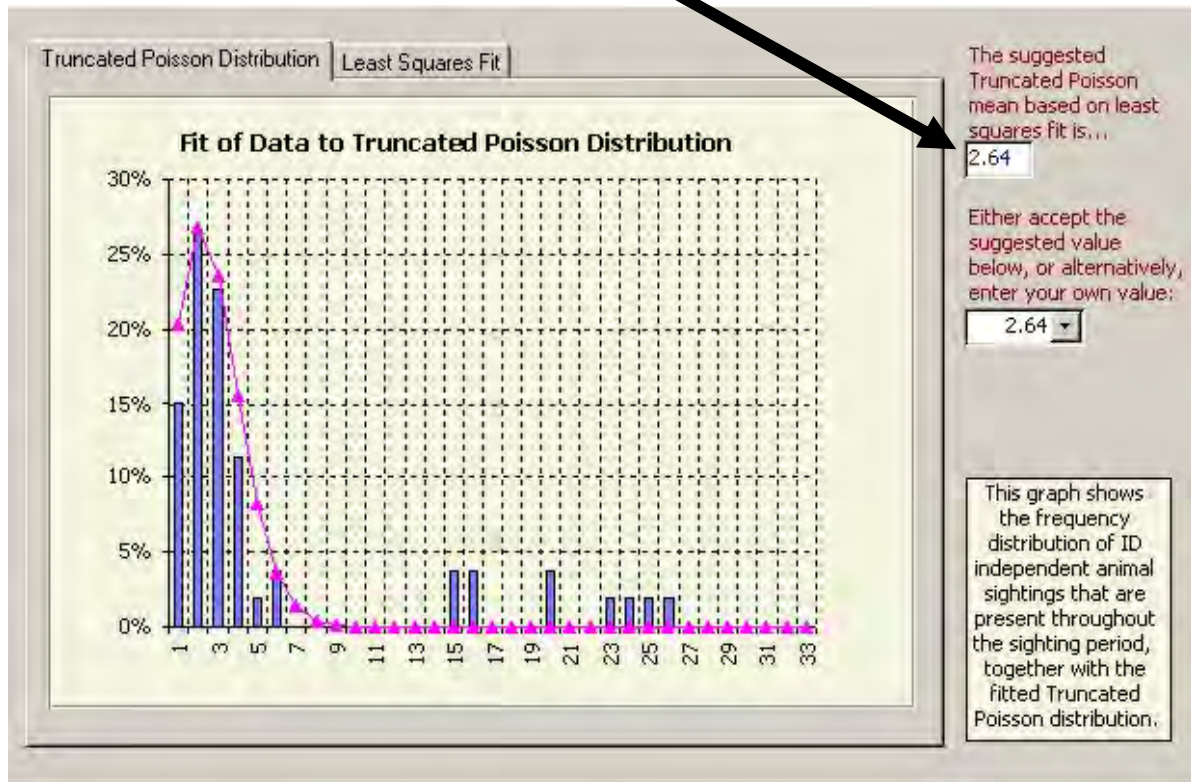
The following graphs illustrate the poorer fit of truncated Poisson means of 1.25 and 5.0.
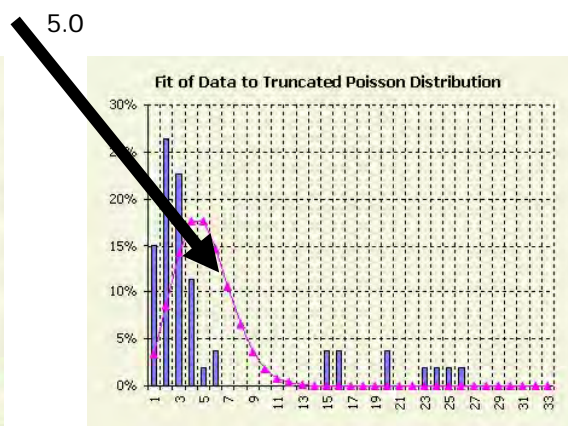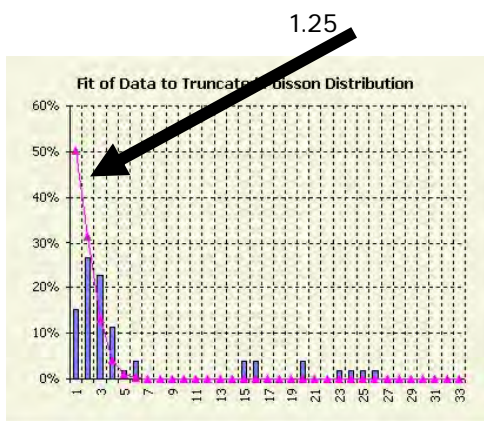
"TRAP-HAPPY" ANIMALS AND THEIR IMPACT ON THE RESULTS..

Trap-happy animals are seen very much more often than other animals and violate a key mark-recapture assumption that all animals have an approximately equal chance of being seen. In the case of rhino, such animals are really sighting-happy. However the term trap-happy is used as it is commonly used in the mark-recapture literature. The term was originally coined to refer to individual rodents that learned that live traps being used to catch and release animals during a study contained food. A few individuals kept coming back for more and were recaptured again and again- hence the term trap-happy! One can also have "trap-shy" animals.

If trap-happy animals are ignored during an analysis they will act to bias population estimates downwards. This is because you will have more re-sightings than normal. This higher than usual re-sighting rate, results in *RHINO* predicting there are fewer animals out there that have not been seen during the recording period than there actually are (hence biasing you estimate).

Best-fit truncated Poisson graph to observed sighting frequency data & RMS error graph.

TWO GRAPHS ARE ACCESSIBLE ON THIS FORM..

This form has two tabs at the top (named "Truncated Poisson Distribution" and "Least Squares Fit").

The tab at the front represents the currently displayed graph. To view the other graph, simply click on the other tab. Users can toggle backwards and forwards between the two graphs by clicking on the tabs. An example of the truncated Poisson graph is shown on the previous page.

Truncated Poisson distribution graph

TRUNCATED POISSON DISTRIBUTION GRAPH..

A Poisson distribution is known to predict the expected frequencies for relatively rare events well.

We know how many of the sighted rhino have been seen once, twice, three times etc. during a survey period. However, we do not know how many other ID animals are out there that we have not yet seen during the study period. In fact, this is what in essence we are trying to use Mark-Recapture methods to estimate.

The fact that we don't know how many animals were not seen is why the observed rhino sighting frequency distribution is compared against a truncated Poisson distribution rather than a Poisson distribution. The truncated Poisson distribution probabilities for a given mean sighting frequency are obtained by initially calculating the Poisson probabilities of seeing a rhino different number of times (including not being seen – i.e. a sighting frequency of 0); then dropping the Poisson probability of not seeing a rhino, and then normalising the remaining Poisson probabilities (of sighting animals once or more are) by multiplying them by a constant so they sum to 1. In this way the sum of the percentage probabilities for both the observed sighting frequencies and the predicted sighting frequencies (derived from the truncated Poisson distribution) both sum to 100%.

This graph shows the best fit truncated Poisson distribution (crimson line) superimposed onto the observed sighting frequency distribution of identifiable animals present throughout the whole data analysis period (blue bars).
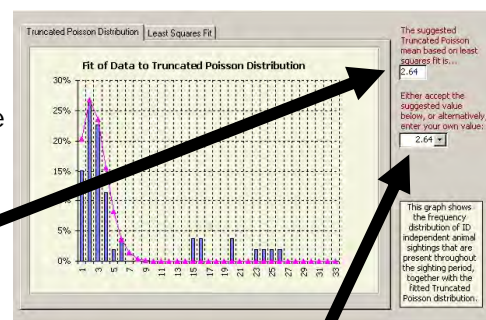
The best-fit mean sighting number (which minimises Root Mean Square Error) is given in blue in the top box on the right. The best-fit mean will initially be shown in black as the default value in the lower cell. If you want to plot a truncated Poisson distribution with a different mean sighting frequency simply edit the black number in the lower box.

The mean sighting frequency still refers to all animals, including those that have not been seen (i.e. had a zero sighting frequency during the reporting period).

If you change the number in this box, you need to make sure that before you click the "Next >" button to proceed, that the black number in the lower box has been updated and contains the mean sighting frequency you want used in the calculations. This value is used to determine how many sightings is enough to be treated as "trap-happy".

In most cases you will want to use the suggested default value shown in blue in the box above.

LEAST SQUARES FIT GRAPH..

If you click on this tab a Root Mean Square (RMS) error graph is displayed covering a range of sighting frequencies up to 10. Select this tab to see how the goodness of fit varies with mean sighting frequency. The best-fit mean sighting frequency is found where the RMS error is the lowest which in this example is 2.64.

If the graph shows a clear minimum, like the example opposite, you can have confidence that your sighting frequencies by and large fit the assumptions. However if there is a flatter plateau minimum you may want to analyse the data for a range of possible mean sighting frequencies with similar RMS errors.
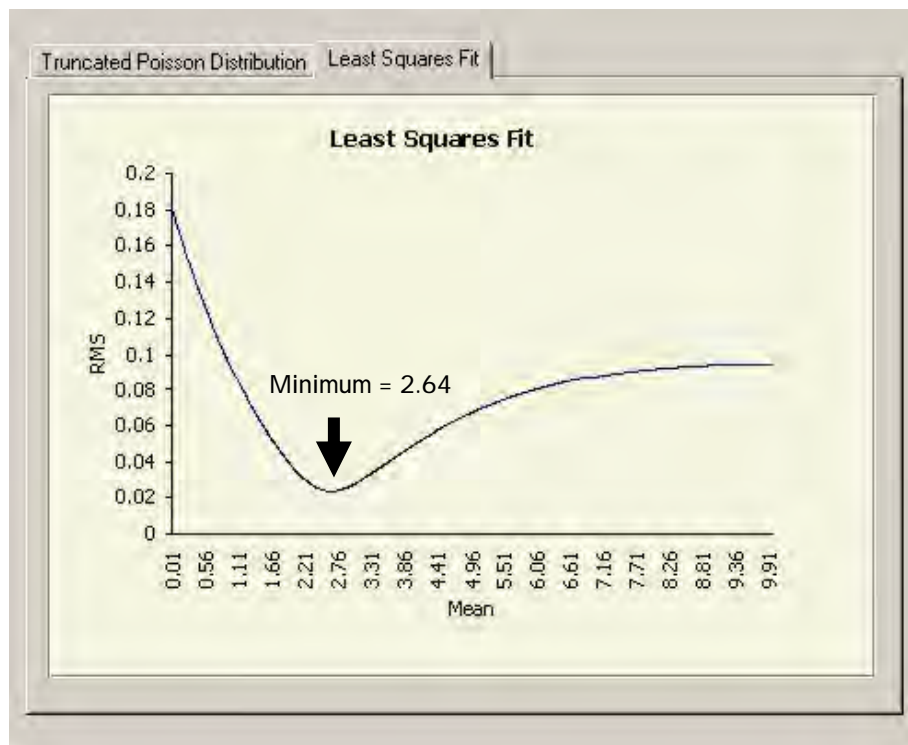
Use of graphs to identify sampling problems

Looking at these two graphs may help you identify possible weaknesses in your monitoring programme.

If most of your sightings are largely of a few very trap-happy animals, then your sampling effort has probably been concentrated in too small an area and you need to make an effort to get better ground coverage before you can even think about getting a reasonable estimate of total population size.

If you have a very flat long plateau in the RMS error graph, and your best fit Poisson mean is high enough to predict there should only be a few animals seen once, yet your sighting frequency histogram shows there are many animals seen only once, it is likely that yoursampling effort differs markedly in different areas of your Park or Section. In the one area being well sampled animals may be being seen regularly, whereas in the other area sampling effort may be so low that most animals are being seen only once. The poor truncated Poisson fit will be because in effect you have two different frequency distributions being superimposed upon each other. In such cases you should probably plot on a map where the sightings of animals seen at different frequencies have occurred. You could also investigate whether you could analyse the high and low intensity sampling areas separately (provided your mean sighting frequency in the low intensity sampling area is > 1.33. Alternatively you should try to increase sampling in the low intensity sampling area and so obtain a more even sampling effort.

The RMS error is calculated for the range of possible mean sighting frequencies up to a maximum of 10. If your non trap-happy rhinos are being seen more frequently than this, then you probably have seen all your ID rhino, and you won't need RHINO to produce a population estimate.

For a given mean frequency, the RMS error is calculated as square root of the average squared deviations between the probabilities for the observed sighting frequencies and the probabilities of expected sighting frequencies (based on the truncated Poisson distribution).

Identifying and dropping trap-happy animal sightings.

VIOLATION OF ASSUMPTION OF EQUAL SIGHTABILITY CAUSES BIAS..

One of the key mark-recapture assumptions is that all animals have an approximately equal chance of being seen. Animals that are seen very frequently compared to other animals are called "trap happy animals". In the case of rhinos such animals are really "sighting-happy", but the term trap-happy has been retained, as it is a commonly used term in mark-recapture methodology.

Trap-happy animals are far more likely to be seen than other animals, and therefore violate the assumption of equal sightability. Trap-happy animals need to be identified and their sightings dropped from the data-set prior to mark-recapture analysis to avoid violating this assumption.

If trap-happy sightings are not identified and dropped prior to mark-recapture analysis, you will tend to underestimate numbers (bias results). This is because with large numbers of trap-happy observations, you will have many more re-sightings, and hence fewer new (first) sightings than if all animals had an equal chance of being seen. As a result RHINO will wrongly conclude that there are fewer animals out there that you have not seen yet - leading to an underestimate of true numbers.

Dropping trap-happy sightings prior to mark-recapture analysis is therefore a bias correction measure, which deals with a specific form of violation of the equal sightability assumption.

To continue click on the "Next >" button. You will then be taken to a form where you can identify trap-happy animals and can select to drop trap-happy sightings prior to mark-recapture analysis.
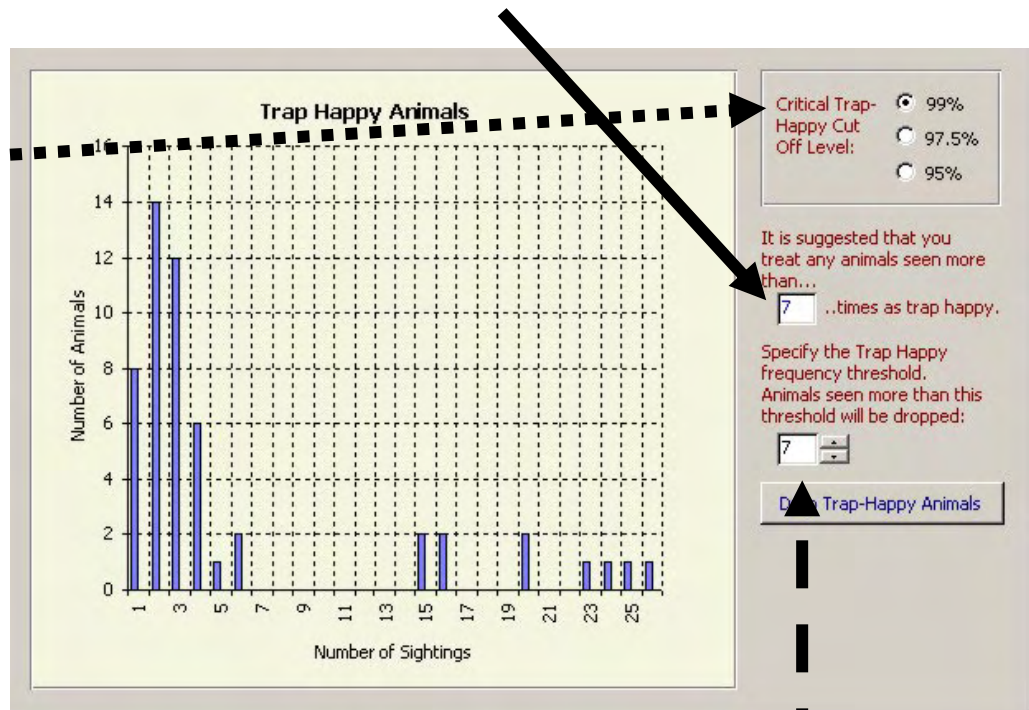
Critical trap-happy sighting threshold

ESTIMATION OF TRAP-HAPPY SIGHTINGS THRESHOLD VALUE..

RHINO uses the estimated mean sighting frequency (entered in the previous form) and the user specified critical trap-happy cut off level set on this form to suggest the maximum number of sightings you should consider normal (shown in blue in the cell below). It is strongly recommended that you use the default 99% critical trap-happy cut-off level.

However, If you want to select one of the other two critical trap-happy cut-off options click to select the radio-button next to the value you wish to select. Maroon text at the right of the form recommends that you should treat any animals seen more than specified threshold frequency (in blue) as treated as trap-happy. If you have used the default 99% critical trap-happy cut off level the chance of observing an animal more than the blue threshold value by chance is less than 1%.

SELECTION OF TRAP-HAPPY SIGHTINGS THRESHOLD VALUE TO USE..

In the second box (the one with a black number), you should enter the maximum number of sightings you are prepared to consider as normal. In most cases, you should use the default number RHINO suggests (which will be the same as the blue number in the box above).
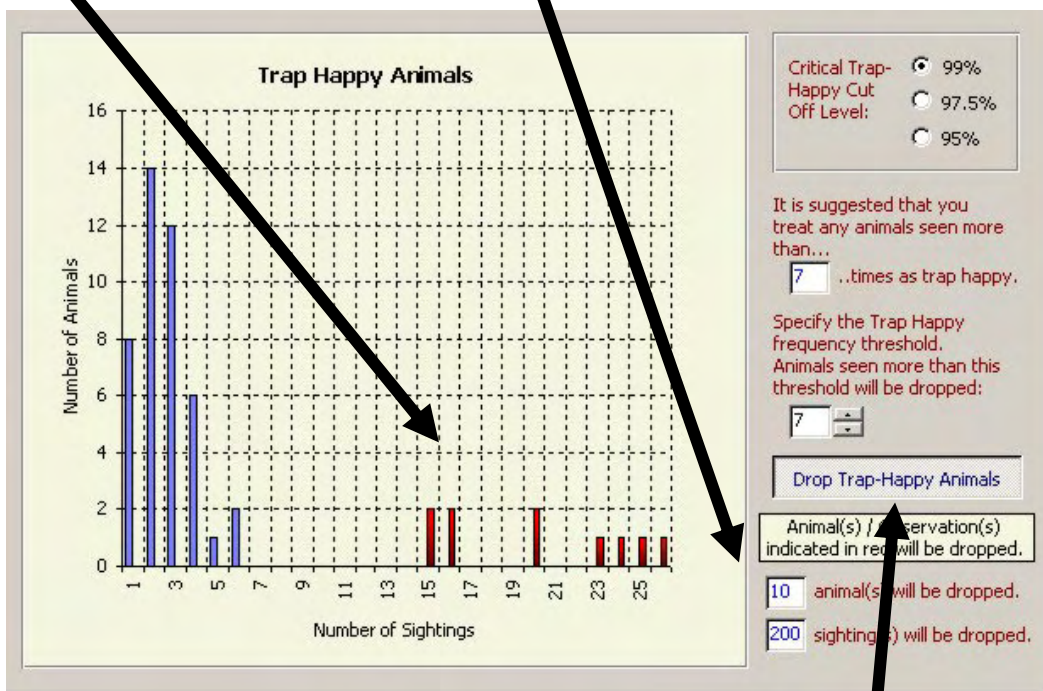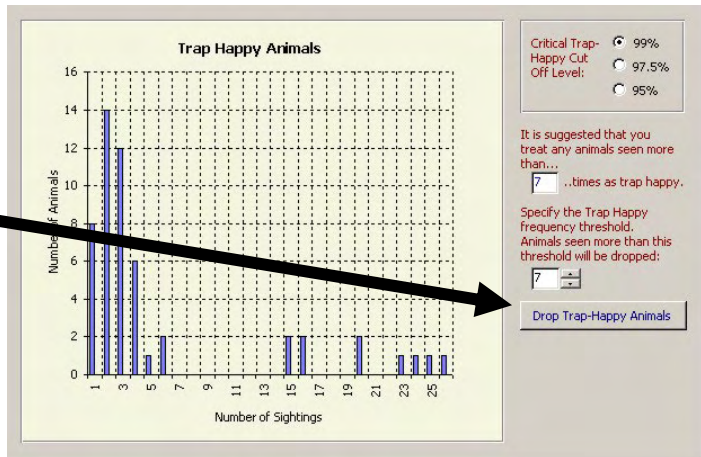
DROPPING TRAP-HAPPY ANIMALS...

If you want to correct for trap-happy animals (recommended), click on the button that says "Drop-Trap-Happy Animals".

RHINO will treat animals seen more often than the specified (black) threshold value as trap-happy, and drop sightings of these animals from the dataset prior to mark-recapture calculations. Two cells at the bottom right of the form will now indicated how many different trap-happy ID animals and the total number of trap-happy ID sightings that will be dropped from the dataset prior to mark-recapture calculations. The animals dropped should also now be shown in red on the graph.

If selected the "Drop trap happy animals" is shown depressed in a light grey.

Dealing with Trap-happy animals

If you change your mind, and decide against dropping trap-happy animals, simply click on the "Drop trap-happy animals" button again. This button allows you to toggle between dropping trap-happy animals and including them in your mark-recapture analysis.

When you are ready to proceed click the "Next >" button.

HOW RHINO DEALS WITH THE DROPPED TRAP-HAPPY ANIMALS..

After mark-recapture calculations have been completed, the number of dropped trap-happy animals is added back (to the number of non-trap-happy animals estimated by mark-recapture) to give a total estimate of the number of ID independent animals. RHINO does this by shifting the mark recapture posterior probability distribution up by the number of trap-happy identifiable animals dropped, and then obtaining the statistics from this trap-happy adjusted graph. In the example on the previous page sightings of 10 trap-happy animals would be dropped and therefore after mark-recapture calculations have been completed the X axis of the ID independent posterior distribution would be shifted up by 10.

FAQ – WHAT ABOUT TRAP HAPPY CLEAN ANIMALS ?

Users commonly ask how RHINO deals with the possibility of trap-happy clean animals. By virtue of these animals being clean, we don't know the sighting frequencies for individual clean animals and therefore we don't really know which clean animals are trap happy.

RHINO simply assumes that the proportion of clean trap-happy sightings that are of trap-happy clean animals is the same as it is for ID animals.

Identifiable and Clean Priors

RHINO – A BAYESIAN TECHNIQUE REQUIRING PRIORS…

RHINO is a Bayesian Mark-Recapture technique. It starts off by requiring the user to define a set of "prior" probabilities of the ID independent population being a range of possible sizes. These probabilities are used to start off the analysis, and they are updated on the basis of the evidence (e.g. sighting/re-sighting dataset and special code information) resulting in the production of the final "posterior" probability distribution at the end of an analysis.

Population estimates and the Bayesian equivalent of confidence levels (called credible posterior intervals) are then derived from these final "posterior" distributions.

Before calculating the statistics, the final ID independent distribution is adjusted up by the number of trap-happy ID animals that were dropped from the dataset prior to mark-recapture calculations. Thus the final posterior ID independent graph includes trap-happy independent animals.

Bayesian methods use information (evidence) to update prior probabilities, and so produce posterior probabilities (the answer). If you have knowledge about a population you may be able to to supply quite "informative" priors. If not, you can supply what are called "uninformative" priors (which only specify the possible minimum and maximum ID population segment value, and allocate and equal prior probability of sighting to each value of X within the specified prior range.

UNLESS YOU HAVE A LIMITED NUMBER OF SIGHTINGS THERE IS NOT MUCH TO BE GAINED BY SPECIFYING VERY INFORMATIVE PRIORS IN RHINO...

Unlike some other Bayesian techniques, experience has shown that once you have sufficient sighting data, the priors supplied to RHINO end up having little or no influence on the final results (unless of course wrong minimum prior and maximum prior values have been specified

when it will be impossible for the answer to converge on the true population size).

Informative priors will be of most use where you have recently started a monitoring programme and have fewer sightings per animal. This will especially be the case if your average sighting frequency per animal is below 1.5. However, as soon as your average sighting number/animal increases to around 2, the results obtained in RHINO, will be the same irrespective of whether informative or uninformative priors have been specified. For this reason, unless you have very limited data there is not much to be gained by supplying very informative priors.

PRIORS ARE REQUIRED FOR THE ID INDEPENDENT POPULATION SEGMENT..

It is important to realize that the priors you initially are asked to supply are for the "ID independent" (identifiable adult/sub-adult) segment of the population only, and NOT for the whole population. This is because RHINO uses a Bayesian mark-recapture technique to initially estimate only ID independent numbers. Numbers of clean independents and dependents (calves) are estimated separately by RHINO. For details of the calculations please refer to the technical manual.

ACTUAL ID POPULATION SIZE MUST FALL WITHIN RANGE OF PRIORS SUPPLIED..

You MUST ensure that the maximum prior you specify is greater than the actual number of ID independent rhinos, and that the minimum prior specified is less than the actual number of ID independent rhinos. This is because, if the true population size falls outside the range of priors you have specified, RHINO will never be able to correctly predict the true population size. If in doubt, it is highly recommended you set the minimum prior at 1, and that you always set the maximum prior quite a bit above the maximum guesstimate of ID independent number to be on the safe side.
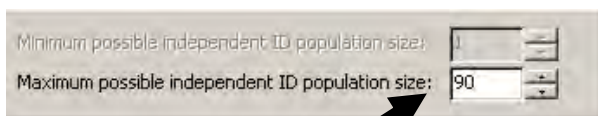
Specifying prior probabilities

If you specify a maximum prior lower than the actual ID independent rhino number, the population estimate will converge on the maximum prior value but no higher. RHINO willwarn you if it detects this may be the case
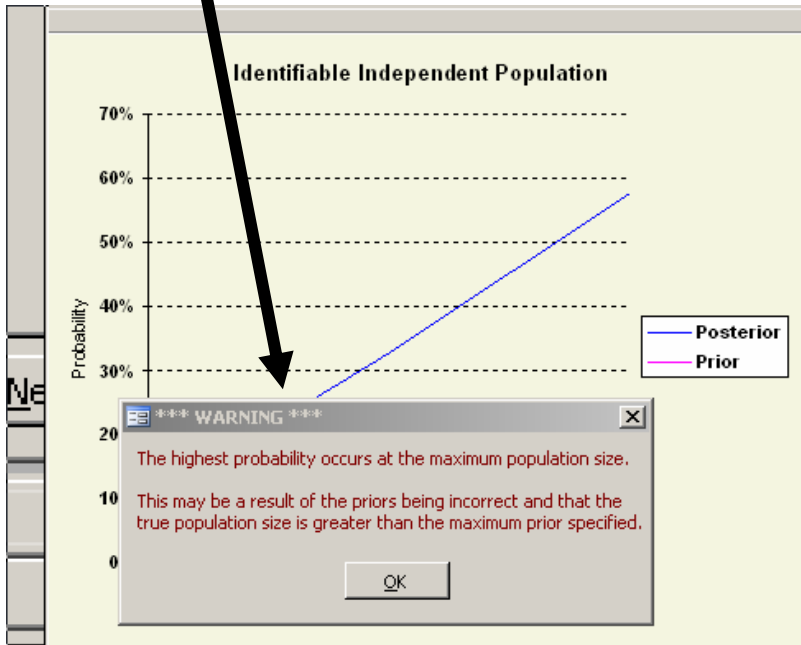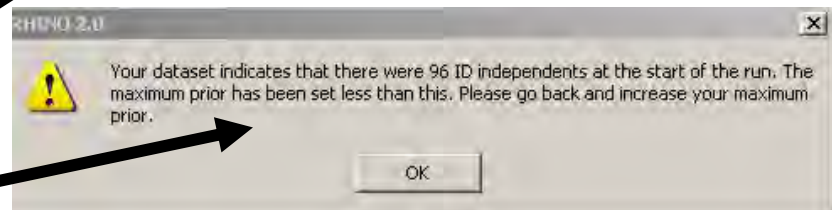

If your ID independent population estimate is almost the same as your maximum prior (which was the case in the example opposite) you should check to see if you may not have set the maximum prior too low.

To try to avoid this problem, by default RHINO sets the default maximum prior at a minimum of a 100 or 50% more than the number of different ID independent animals in the dataset.

Similarly if you specify a maximum prior less than the number of ID independents RHINO will detect this and warn you
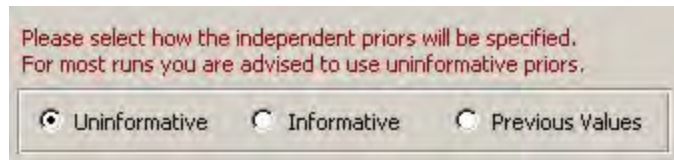
If you specify a minimum prior higher than the actual ID independent rhino number, the population estimate will converge on the minimum prior value but no lower. If your estimate is almost the same as your minimum prior, you should check to see if you might not have set the minimum prior too high.

ID Independent Prior Options

ID INDEPENDENT PRIOR OPTIONS..

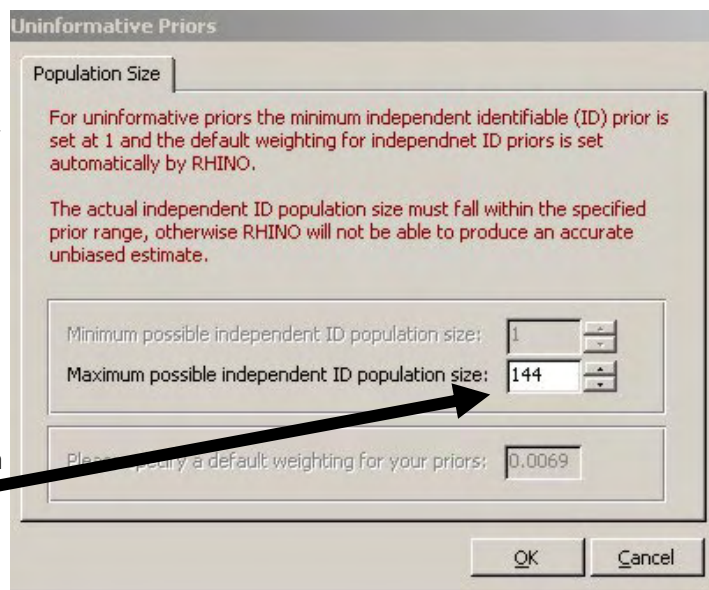Click on one of the radio-buttons at the top of the form to select one of the three following options..

Please select how the independent priors will be specified. For most runs you are advised to use uninformative priors.

- ⦿ Uninformative ○ Informative ○ Previous Values

"Uninformative" priors;

"Informative priors"

"Previous values"

The first two options allow the user to specify new priors that can be saved for future use, The third option allows users to re-use a set of priors that were entered and saved earlier.

UNINFORMATIVE PRIORS...

If you select the "Uninformative" option, a dialog box will appear which will prompt you for the maximum ID prior value. You will probably will select "Uninformative" priors for most runs.

To try to ensure that there is a good chance that the maximum prior is bigger than the actual ID independent popululation segment sise, the maximum possible ID independent Prior will be set by default to the larger of 100 or 50% more than the number of ID independents in the dataset which in this case = 144.
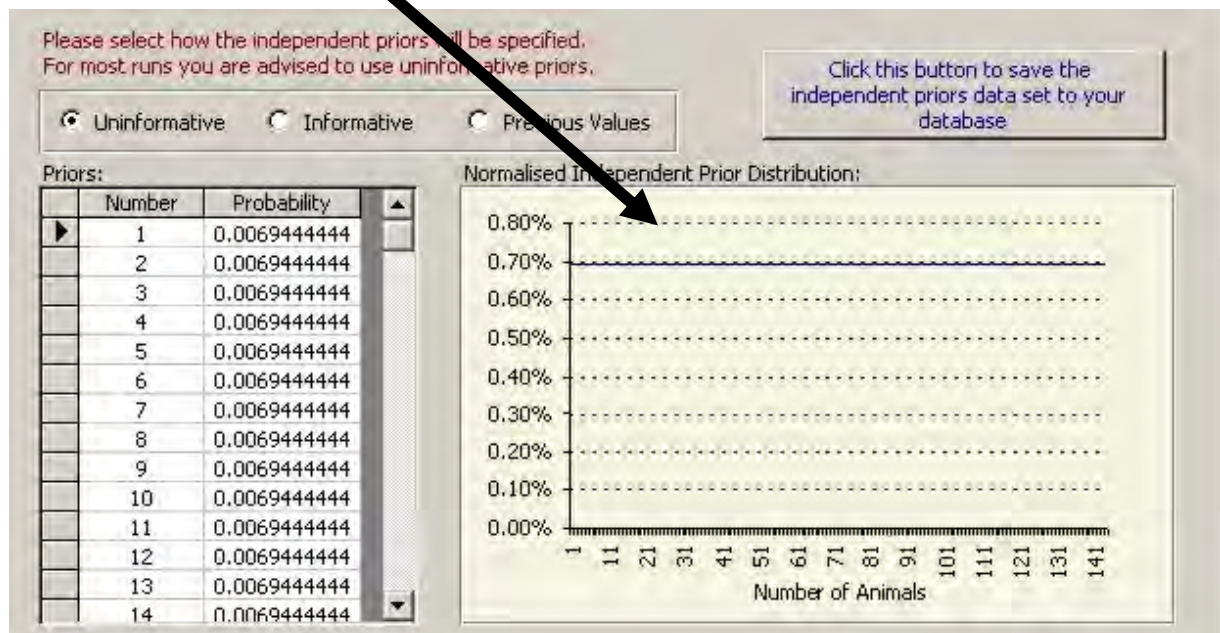
Uninformative Priors

Population Size

For uninformative priors the minimum independent identifiable (ID) prior is set at 1 and the default weighting for independnet ID priors is set automatically by RHINO.

The actual independent ID population size must fall within the specified prior range, otherwise RHINO will not be able to produce an accurate unbiased estimate.

Minimum possible independent ID population size: 1

Maximum possible independent ID population size: 144

Please specify a default weighting for your priors: 0.0069

OK   Cancel

Uninformative priors

In RHINO 2.0 the default minimum prior box is disabled when specifying uninformative priors and by default RHINO 2 sets the minimum prior at 1. RHINO 2.0 users now only have to specify the maximum ID independent prior.

Just make sure that the maximum prior is definitely bigger than the actual ID independent population size. The "default weighting" box is also disabled when specifying uninformative priors. RHINO simply allocates an equal prior probability to every possible population size in the specified prior range.

If you want to maximize prior entropy and ensure that the final result is most influenced by your sightings data and not your priors you should select to use "Uninformative" priors.

After you have specified the maximum uninformative prior value, click the "OK" button. RHINO then automatically generates and displays a graph of the specified uninformative prior distribution.

Informative priors

If you select a maximum uninformative prior of say 100, what you are doing is specifying that the ID independent population segment size is somewhere between from 1-100 at the start (before you analyse any data) but at the start there is an equal chance it could be any number in this range.

If you want to save these uninformative priors to use them again later, click on the "Save priors" button in the top right of the form, and follow the instructions. However, it is so quick to enter uninformative priors there is little point in saving them.

INFORMATIVE PRIORS ...

You might select this option if you know how many animals were put into a re-established population, and you know that given rhino biology, the population cannot have got bigger than X rhinos, and is most probably somewhere between Y and Z. Thus, you are in the position where you have the knowledge to specifying maximum and minimum priors, as well as variable starting probabilities of the population being different sizes. In this case you would select to use "Informative" priors.

If you select this option you will be asked to specify both minimum and maximum priors and you will also be asked to supply a default probability weighting value. This can be any number you want and can be an integer. This weighting value will be used as the default value to initially populate the prior weights table that will be displayed after you click on the "Next >" button.
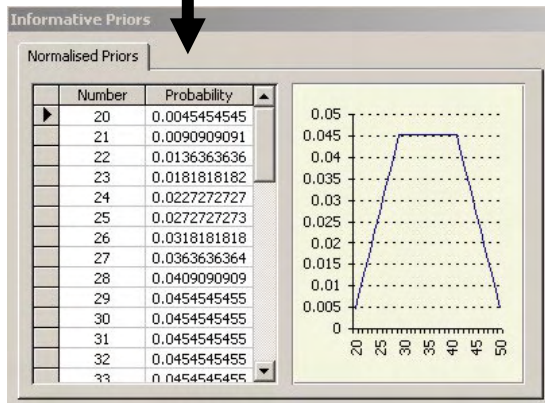
To change weights for any particular value of N in the "Define weighted priors" form, simply type in the new weights.

Note how the informative prior distribution changes shape as you input the weights

Informative Priors

Define Weighted Priors

|	Number	Weighting
▶	20	1
	21	2
	22	3
	23	4
	24	5
	25	6
	26	7
	27	8
	28	9
	29	10
	30	10
	31	10
	32	10
	33	10

< Back   Next >   Cancel
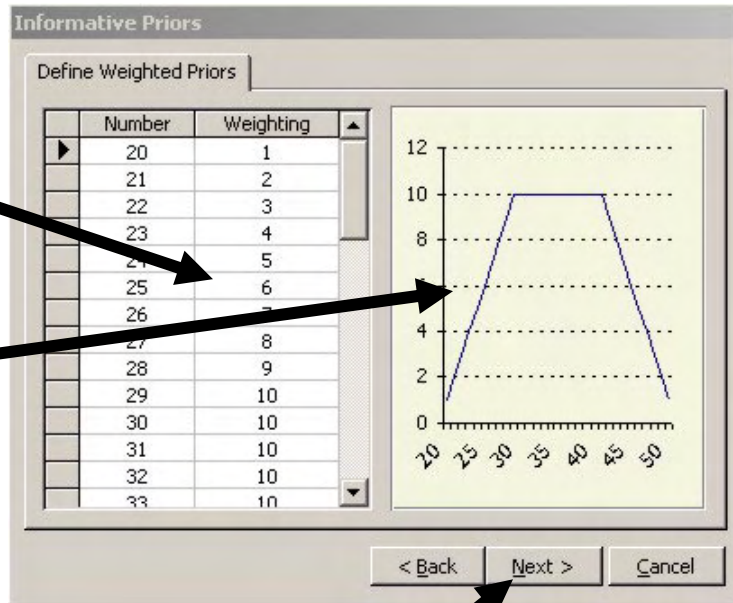
When you have finished entering in all the weights click on the "Next >" button and a normalised prior probability distribution will be produced (the weights having been scaled so that they sum to 1).

Informative Priors

Normalised Priors

	Number	Probability
▶	20	0.0045454545
	21	0.0090909091
	22	0.0136363636
	23	0.0181818182
	24	0.0227272727
	25	0.0272727273
	26	0.0318181818
	27	0.0363636364
	28	0.0409090909
	29	0.0454545455
	30	0.0454545455
	31	0.0454545455
	32	0.0454545455
	33	0.0454545455

When you are ready press the "Finished" button on the normalized priors form.

Please select how the independent priors will be specified. For most runs you are advised to use uninformative priors.

○ Uninformative   ● Informative   ○ Previous Values

Click this button to save the independent priors data set to your database

Priors:

	Number	Probability
▶	20	0.0045454545
	21	0.0090909091
	22	0.0136363636
	23	0.0181818182
	24	0.0227272727
	25	0.0272727273
	26	0.0318181818
	27	0.0363636364
	28	0.0409090909
	29	0.0454545455
	30	0.0454545455
	31	0.0454545455
	32	0.0454545455
	33	0.0454545455

Normalised Independent Prior Distribution:

To save informative priors for future use, click on the "Save priors button" and follow instructions.
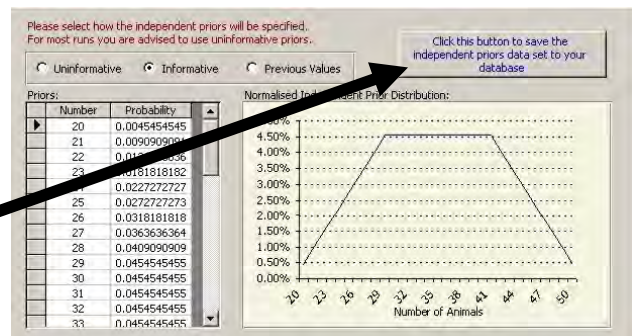
If you wish to display partial statistics later (calculated after each fifth of the data has been analysed) in addition to the final posterior posterior ID independent probability distribution, you must set the minimum prior value to 1.

USING PREVIOUSLY SAVED PRIOR VALUES …

If you select this option, a "Prior management" form will appear with a menu of the different saved priors.

As you click on a prior file name, a small graph is displayed in the right box showing the shape of the selected prior distribution.
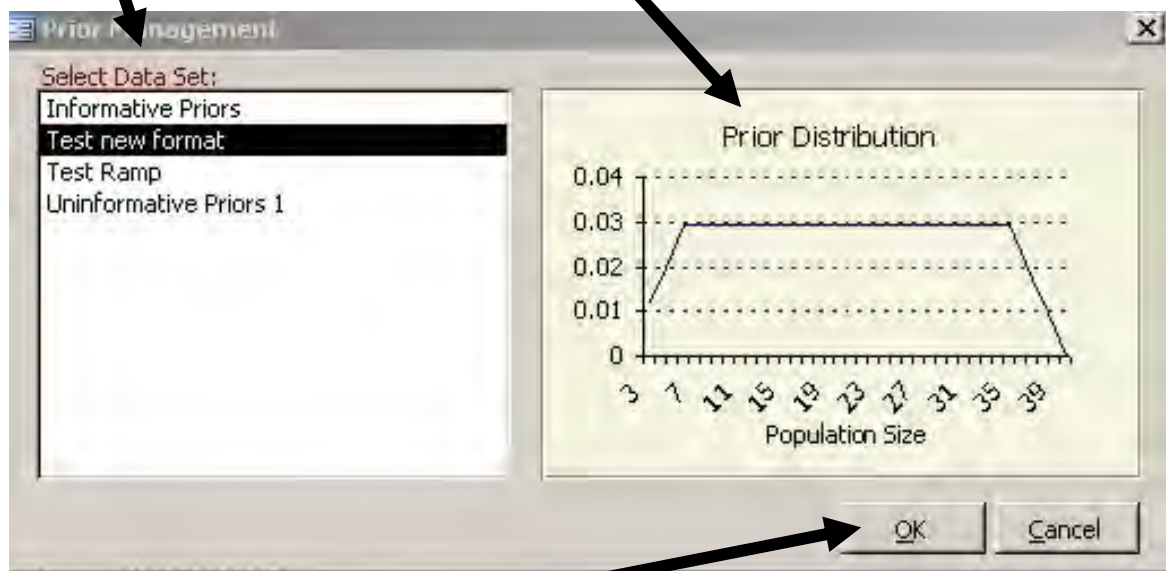
Click the "OK" button when you have selected the saved priors you want to use. Then click the "Next >" button to proceed.

**INTERPRETING A PRIOR PROBABILITY GRAPH**

A "prior" probability distribution is the starting probability distribution used in a Bayesian analysis. These probabilities are updated and continually improved in response to the evidence (sightings data, special code information etc.) to produce a final "posterior" probability distribution from which the population estimates and Bayesian equivalent of confidence levels are derived.

Interpreting prior graphs

The interpretation of prior and posterior graphs is the same. The vertical Y axis gives the probability that the population segment is each of the range of possible sizes (shown on the horizontal X axis). The key principle to grasp is that the bigger the Y axis value, the greater the chance the corresponding value of X is the population size.

Uninformative priors are shown by a straight line – which in effect often means "we haven't a clue what the ID independent population size is, so we will start by specifying that every possible value of X (population size) from 1 up to the maximum prior has an equal chance of being the actual population size.
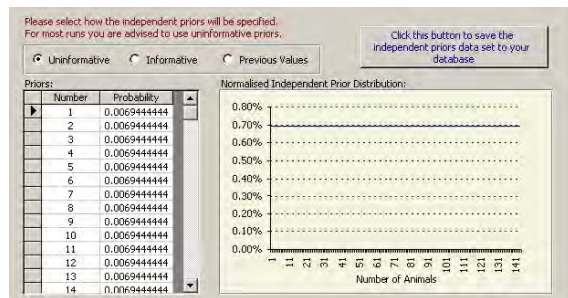
Informative priors can take a range of shapes, but some values of X will start with higher probabilities.
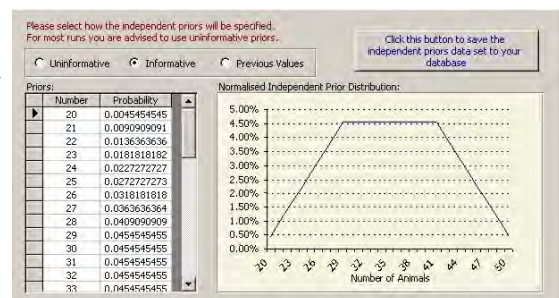
In both cases the minimum and maximum prior values (of X) represent the bounds within which we are specifying the true population size must lie.
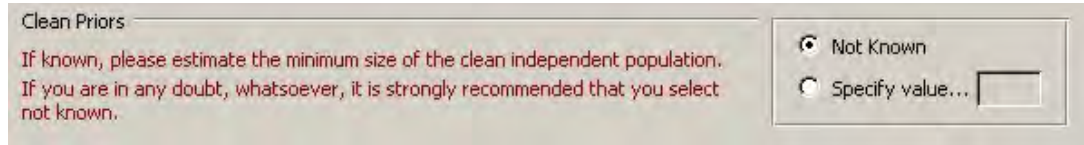
PROCEEDING WITH ANALYSIS – SPECIFYING THE MINIMUM CLEAN PRIOR..

When you have selected your ID independent priors, and the normalized independent prior distribution is shown in the yellow window on the main priors form, you can now optionally enter the minimum clean prior using the box at the bottom of the page.

Specifying minimum clean prior

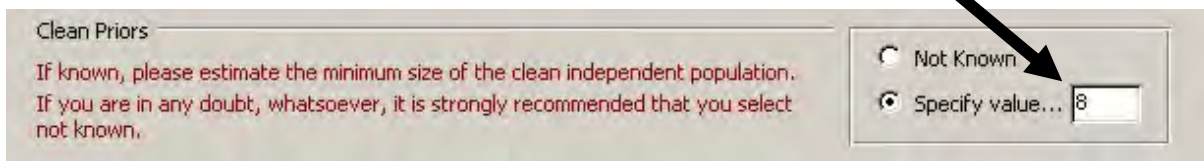The default clean minimum value used by RHINO is 1 (the "Not Known" radio-button will be selected).

If you are happy to use the default clean prior of 1, click the "Next>" button to proceed.

However if you have some knowledge of the minimum clean prior, you can enter another number. This is a form of more informative clean prior. For example supposing that you saw 2 clean F age rhino together before driving seven kilometers and seeing another 3 clean rhino cows with different aged calves, and a clean E aged bull and 2 clean F aged bull (with different horn configurations). Thus you know there must be at least 8 clean independent animals in the population. By selecting the "Specify value" radio-button and specifying 8 as the minimum clean independent prior, RHINO cannot estimate there are fewer clean animals than this minimum number you know are present in the population.

One advantage of using Bayesian methods like RHINO is that you can use information you have to reduce the chance of getting daft answers (such as a lower confidence levels which are lower than the known minimum). Information on clean animals identifiable only to a key observer using more minor features (such as horn configurations, tiny ear nicks, small scars, photographs of eye and nose wrinkles etc.) can also be used in determining the minimum clean prior. Sightings of such animals are still treated as clean sightings by RHINO.

Click the "Next>" button to proceed
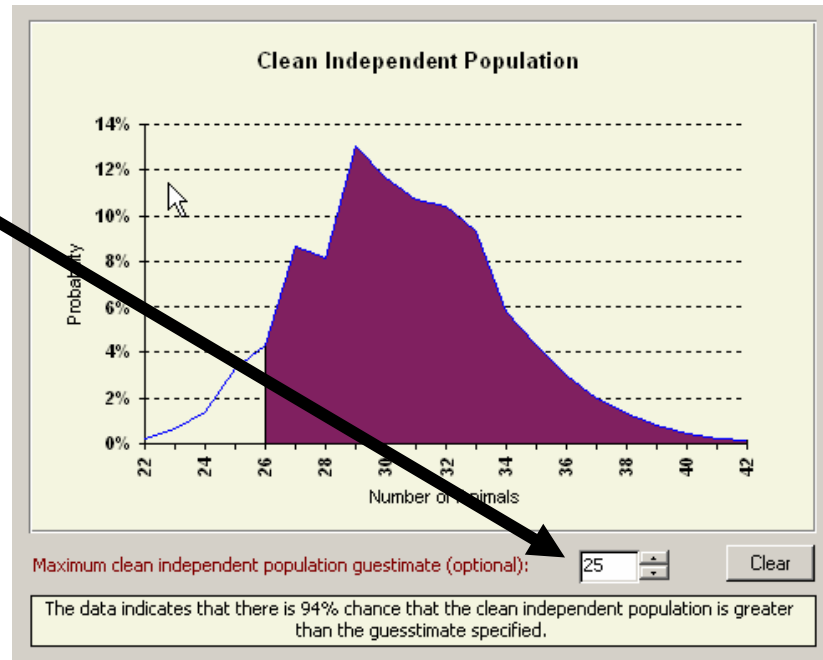
Specifying Priors — Changes in RHINO 2.0

CHANGES IN SPECIFYING PRIORS BETWEEN RHINO 1.21 and RHINO 2.0 ..

In RHINO 2.0, users no longer need to specify the minimum uninformative ID prior as this is automatically set at 1.

In RHINO 2.0 users can no longer select the upper clean prior. RHINO by default sets this at 20,000. The reason for this change, is that some users of RHINO 1.21 may have biasing clean estimation by setting upper clean priors which were likely (based on the data) to be well below the actual number of clean animals. Taking a simple example – Suppose the ratio of ID: Clean animals in a population is exactly 50:50, and the ID population estimate is 40 animals. It follows that one could expect there to be a similar number of clean animals (in this case around 40). However a user of RHINO 1.21 could chose to specify that the maximum clean independent number was only 10. If so then the maximum clean population size that RHINO can estimate will be 10 (in this case only about a quarter of the probable true clean population size). To eliminate the chance of this happening, users can no longer specify the upper clean prior.

However, on the Clean independent graph form later, users can specify a suggested maximum clean independent number, and RHINO 2 will calculate the probability of the population being greater than this value given the evidence (data).

To avoid confusion deselect the shade CPI tick box to see the part of the distributon bigger than the specified guesstimate of the maximum number clean.

Clean Independent Population

Maximum clean independent population guesstimate (optional): `25`   Clear

The data indicates that there is 94% chance that the clean independent population is greater than the guesstimate specified.

Management of Saved Priors

MANAGEMENT OF SAVED PRIORS...

If you want to delete any previously saved ID independent priors then go back to the initial
RHINO start-up screen and main menu.   Click on the "Tools" icon in the main menu topic box,
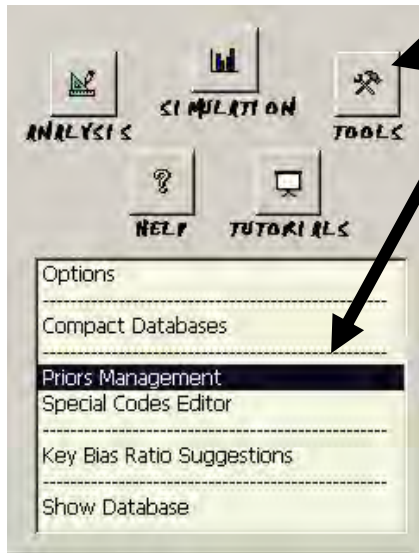
Then select "Priors Management" in the right hand menu .

A "Priors management" form will appear listing the saved
prior files in the top box with a graph of the currently
selected set of saved priors in the yellow box beneath. Click
on the saved prior file names to select different saved
priors.
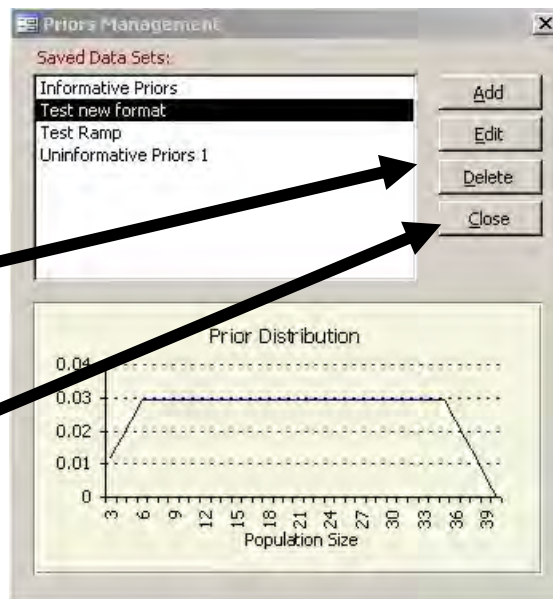
Click on the buttons to the right to either   "Add",
"Edit", "Delete" or "Add" Priors".

When you have finished click on the
"Close" button.

BAYESIAN MARK RECAPTURE

RHINO then will proceed to undertake the Bayesian mark recapture analysis.

Results of Bayesian mark-recapture

Results of Bayesian Mark-Recapture— Identifiable Indpendent Posterior Probability distribution and associated statistics.

Once *RHINO* has completed the Bayesian Mark-Recapture analysis of the non-trap-happy ID independent data, the resultant posterior probability distribution is shifted up by the number of trap-happy ID animals dropped from the dataset prior to Mark-Recapture to obtain the final ID independent posterior probability distribution. For details of the Bayesian mark-recapture calculations please refer to the technical manual. The form that is displayed presents the results of the analysis.

The blue graph shows the final posterior probability distribution for the ID independent segment of the population (dropped trap-happy animals having been added back in).
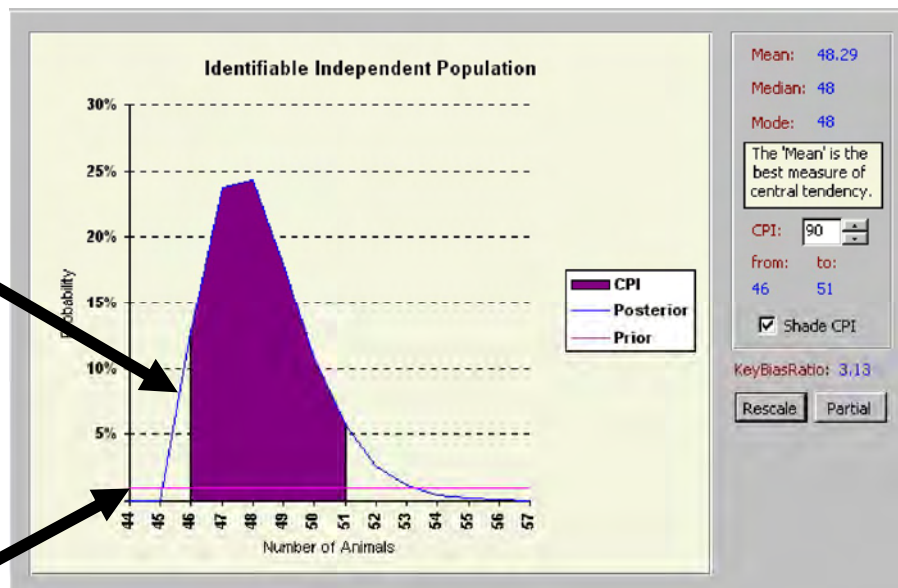

The light purple line

shows the prior distribution probabilities for the X axis range being graphed. You should rescale the graph (see below) if you want to see the whole prior probability distribution.

Final ID independent posterior distribution

ID INDEPENDENT POPULATION ESTIMATES ..

Population estimates (mean, mode and median) are calculated from the posterior probability distribution, and are shown in blue a box at the top right of the form. The "mean" value is the average value of X for the distribution; the "median" represents the value of X which falls exactly in the middle of the distribution (50th percentile), and finally the "mode" represents the value of X with the greatest probability of occurrence.
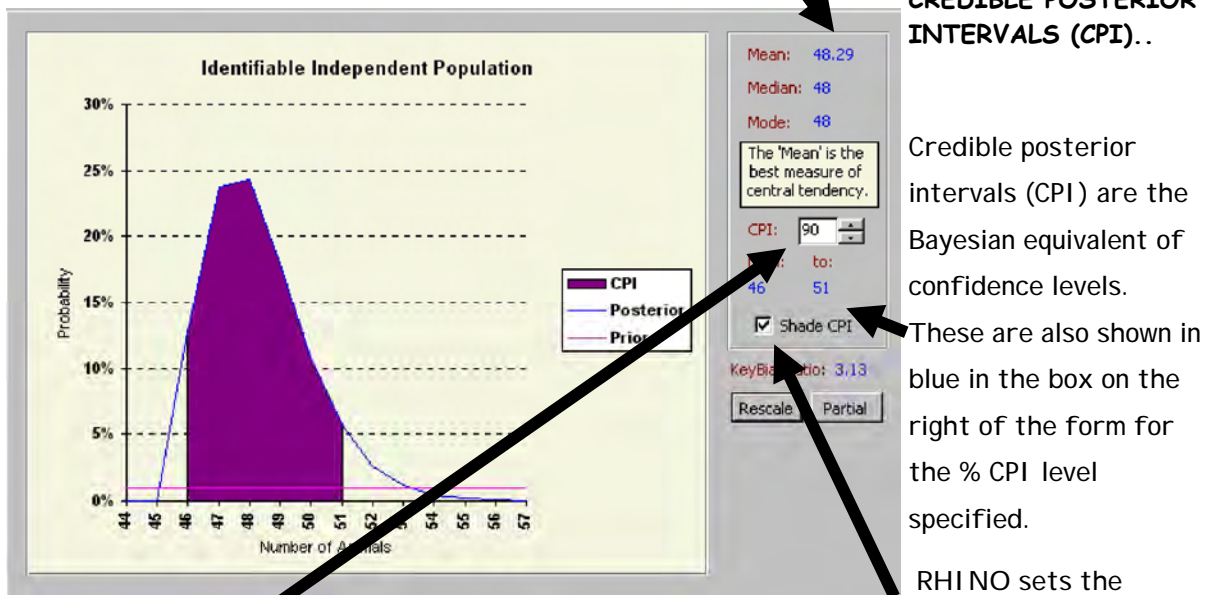
CREDIBLE POSTERIOR INTERVALS (CPI)..

Credible posterior intervals (CPI) are the Bayesian equivalent of confidence levels. These are also shown in blue in the box on the right of the form for the % CPI level specified.

RHINO sets the default CPI level at 90% (with CPI values representing the numbers of rhino at the 5th and 95th percentile in the distribution). To shade the CPI on the graph click the Shade CPI tick box. The graph will look like the one above in Access 97 and Access 2000. For some as yet unknown reason shading the ID independent CPI intervals in Access 2003 on an XP machines also shades the prior probabilities.

SHADING CPI "CONFIDENCE" LEVELS..

If you want to shade the area of the graph covered by the selected CPI levels, click to check the shade CPI tick-box on the right of the form. By repeatedly clicking the tick box you can toggle between shading the CPI or not.

Changing the CPI "Confidence" level

SETTING DIFFERENT CPI "CONFIDENCE" LEVELS..

If different "confidence levels" are required (e.g. 95%), simply type in the level required in the CPI cell. The blue CPI levels shown below the CPI cell will change in response to the new CPI level selected.

INTERPRETING A POSTERIOR PROBABILITY GRAPH..

A "posterior" probability distribution such as the one shown by the blue line in the graph on the previous page is the final probability distribution after processing all the evidence (data). The vertical Y axis gives the probability that the population segment (in this case ID independents) is each of the range of possible sizes (shown on the X axis). The greater the probability on the Y axis, the greater the chance the corresponding value of X is the population size.

Population estimates (mean, mode and median) are calculated from each posterior distribution.

SELECTION OF WHICH MEASURE (MEAN, MEDIAN or MODE) TO USE TO ESTIMATE POPULATION SIZE

Each of the above measures (mean, median or mode) will give the best estimate under different circumstances.

RHINO 2.0 has improved upon *RHINO 1.21's* relatively crude "statistical advisor system" (which was based on limited simulation modeling). A simple expert system has been developed (based on the results of 65,000 simulations) which advises the user on the probable best measure (mean, median, mode or none of these) to use. A yellow box in the top right of the form advises users on which measure they should use in each case.

Mean: 74.47
Median: 74
Mode: 73

The 'Median' is the best measure of central tendency.

Additional statistical advice is inserted into the reports produced at the end of the run.
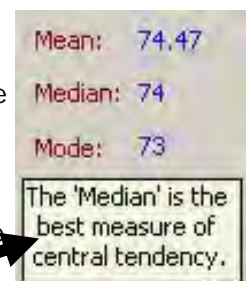
RHINO's Stats Advisor & Rescaling graphs

The built in statistical advisor in *RHINO* also gives users an approximate idea of possible biases in their estimate (documented in the *RHINO* run report). *RHINO* will also warn users when they are likely to not have enough data and that *RHINO* estimates will be unreliable. In such cases *RHINO* estimates should not be used until additional data have been collected.

In the next major version of *RHINO* (ver 3.0), it is hoped to include a much more sophisticated statistical advisor system based on Neural Network analysis of very large simulated data sets. The results of these simulations are also currently being used by the developer to develop some improved statistical advice rules, which in the interim can hopefully be incorporated in an updated RHINO 2.1 version.

RESCALING THE GRAPH ..

RHINO 2.0 now automatically scales the posterior graph to exclude unlikely values of X. However the user can rescale any graph in RHINO by clicking the Rescale button. Users simply select maximum and minimum values for the X axis, and a maximum value for the Y axis. Click on the "OK" button and the graph will be re-drawn. You may want to rescale the graph if you want the X axis to cover the whole range of priors selected. Rescaling is also advisable if you have selected the Partial statistics option.
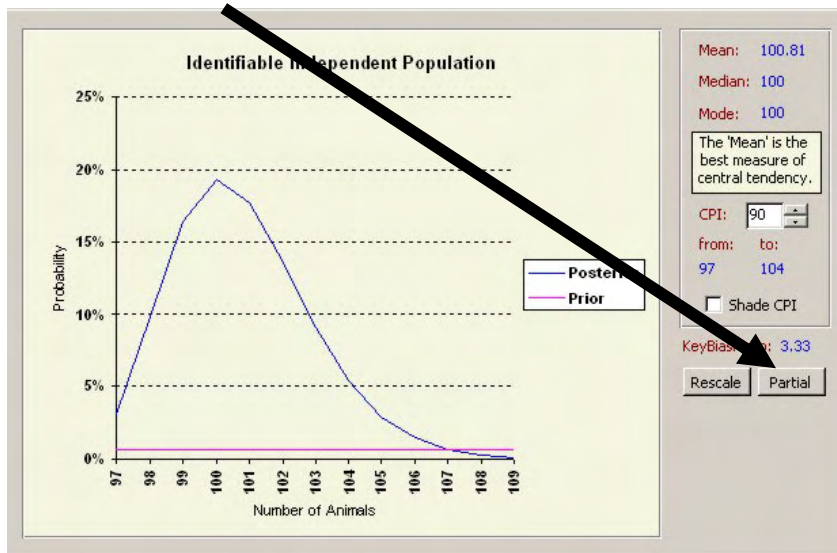

The final graph displayed on this form is the one used in the final RHINO run report

Partial Statistics

PARTIAL STATISTICS..

Click on the "Partial" button if you also want to display posterior distributions after each fifth of the ID independent dataset has been analysed. This is primarily useful in training as users

can see the effect of adding more data on the shape of the posterior distribution. In most cases the partial statistics graphs will becoming more precise as you analyse more data. The "Partial" button is disabled if the minimum prior is not set to 1. If you have selected the "Partial" option and want to return to the original graph simply click the "Full" button. The example below shows how the answer converges towards an actual population size of 100 and precision increases as the amount of data analysed increases.
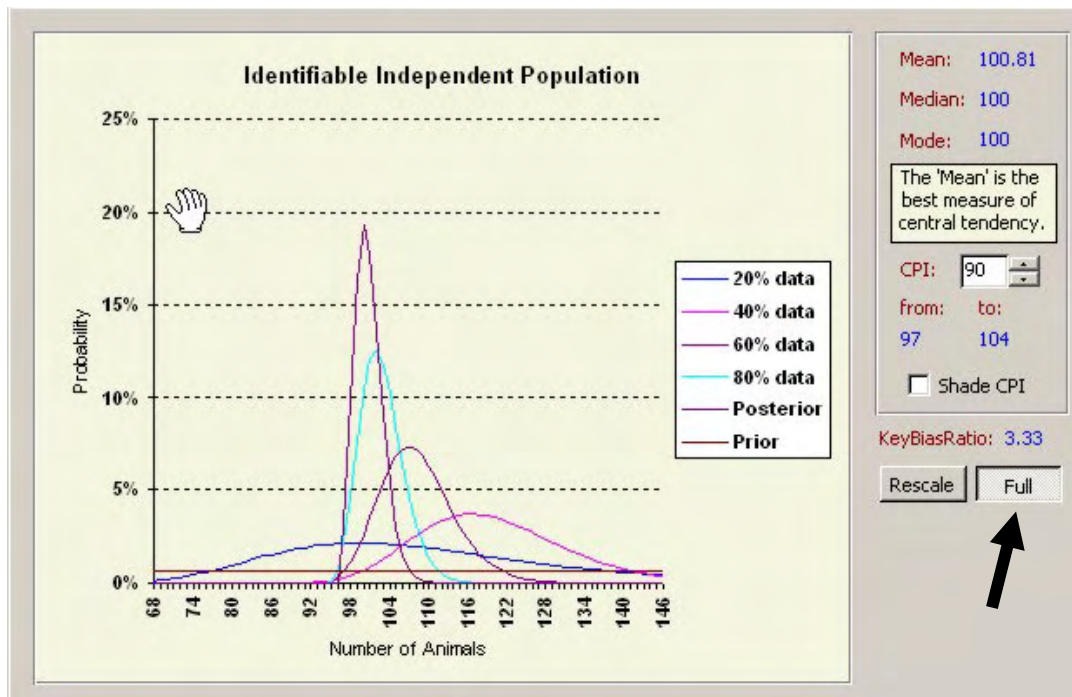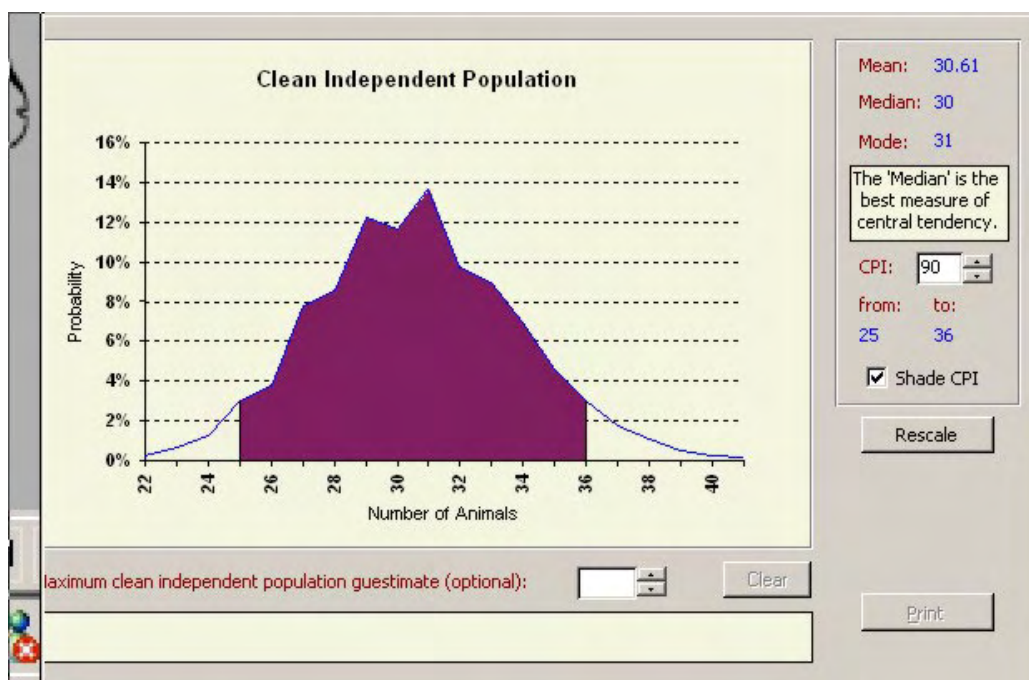
Clean Independent Graph

Clean Independent Graph

CLEAN INDEPENDENT POSTERIOR PROBABILITY GRAPH..

Once you click the next button on the ID Independent segment results page, RHINO proceeds to estimate the Clean Independent population size, and presents the results of the analysis in the form of a Clean Independent Posterior distribution.

The blue line shows the probabilities of the clean population segment being each of a range of X axis values. Population estimates (mean, mode and median) are again calculated from this distribution, as are the Bayesian equivalent of confidence levels (CPI) which can be shaded by checking the Shade CPI tick box. The CPI level can be changed by changing the number in the CPI box on the form. The graph can be rescaled by clicking the "rescale" button.

RHINO's 2.0's improved statistical advisor expert system advises the user on the probable best measure (mean, median, mode or none of these) to use. The yellow box in the top right of the form advises users on which measure they should use (the median in this above example).

Estimating clean independent number

In the next major version of RHINO (ver 3.0), it is hoped to include a much more sophisticated statistical advisor system based on Neural Network analysis of very large simulated data sets. The results of these simulations are also currently being used by the developer to develop some improved statistical advice rules, which in the interim can hopefully be incorporated in an updated RHINO 2.1 version.

CLEAN ESTIMATION METHODOLOGY..

A key assumption RHINO uses to estimate the clean population size is that the sighting frequency levels for clean animals will be similar to those for ID animals. RHINO uses the final ID independent posterior distribution to derive weighting factors which quantify the relationship between ID sample number and the ID population estimate. These weighting factors are used together with information about possible binomial sampling error to derive the final Clean independent distribution. Correction factors are used in the calculations, if the clean number has changed during the reporting period due to deaths, marking exercises, removals etc. For full details of the methods used to estimate clean number please refer to the technical manual.

### FREQUENTLY ASKED QUESTION – WHAT ABOUT TRAP HAPPY CLEAN ANIMALS ?

Users commonly ask how RHINO deals with the possibility of trap-happy clean animals. By virtue of these animals being clean, we do not know the sighting frequencies for individual clean animals and therefore which clean animals are trap happy. RHINO simply assumes that the proportion of clean trap-happy sightings in a given dataset is the same for clean animals as it is for ID animals. It does this by using weighting factors in the clean population estimation calculations which are derived from the final ID independent distribution (which includes trap-happy animals). These weighting factors reflect the relationship between sample number and population estimates based on the sightings of all ID animals (i.e. including trap happy animals). hese weighting factors in effect adjust the clean population estimate so that it is based on the assumption that a similar proportion of clean animals are trap-happy as ID animals.

Probability of clean population being less than X

PROBABILITY OF THE CLEAN POPULATION BEING BIGGER THAT A SUPPLIED GUESSTIMATE..

In *RHINO 2.0*, users no longer have the option of specifying an upper clean independent prior. This removes the chance of users biasing the results (underestimating clean numbers) by supplying a maximum clean prior that the data strongly indicates is likely to be well below the likely clean population size.

On this form, users still have the opportunity to enter their best guesstimate of the maximum clean number if they wish. Simply enter a number in the box and press the "Enter" key on your keyboard.
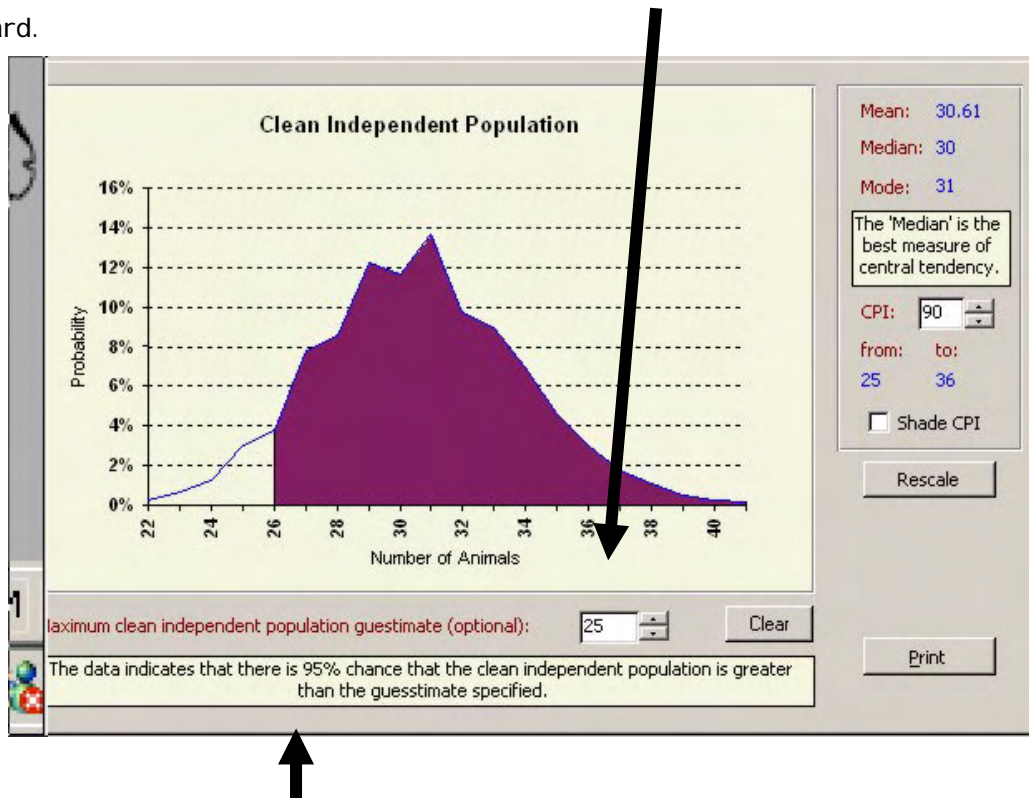
RHINO then calculates the probability based on the evidence (i.e. data), that the clean independent population is greater than the specified maximum guesstimate. Probabilities above the supplied maximum guesstimate are also shaded on the graph. To view this graph best you may want to "toggle-off" shading the CPI interval if you have this option selected. To clear this box and related graph and text click the "Clear" button.

Total Independent Graph

Total Independent Graph form

TOTAL INDEPENDENT (ID & CLEAN) POSTERIOR DISTRIBUTION AND ASSOCIATED POPULATION ESTIMATE AND CPI VALUES

The next graph shows the total Independent segment posterior probability distribution. This is derived by combining the ID independent and clean independent posterior distributions. The blue line shows the probabilities of the total independent population size being each of a range of X axis values.
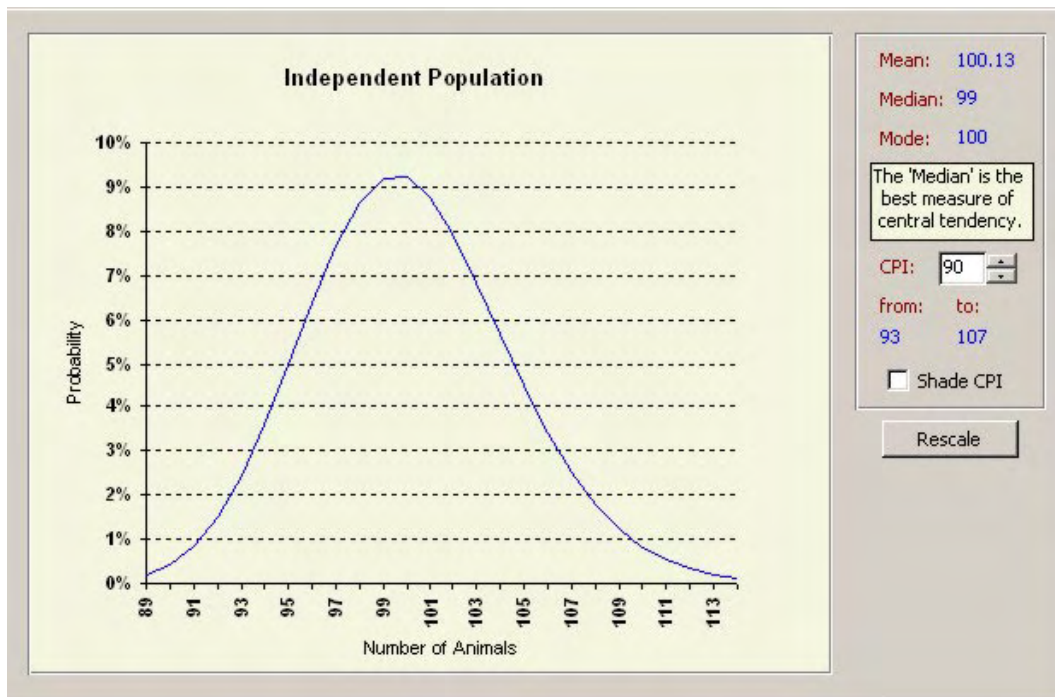
Population estimates (mean, mode and median) and the Bayesian equivalent of confidence levels (CPI values) are calculated from this distribution in the same way as in previous graphs. As with previous graphs, users can chose to re-scale the graphs, change CPI levels, and shade CPI levels. In this case the Shade CPI tick box is not checked and so the 90% CPI level from 93-107 is not shaded. To shade the CPI you would simply check this tick box. RHINO's statisitical advisor also advises on the best estimate to use (Median in this example) To proceed click the "Next >" button. To return to the previous form click the "< Back" button.

Total Independent Estimates in a Multi-Area Analysis

TOTAL INDEPENDENT POPULATION GRAPH IN A MULTI-AREA SIMULATION..

In a multi-area simulation the total independent population posterior distribution (shown below) is determined by combining the total ID independent and total Clean independent distributions. In a multi-area simulation these two combined total distributions are also shown on this form. They are accessed by simply clicking on the tab names at the top.
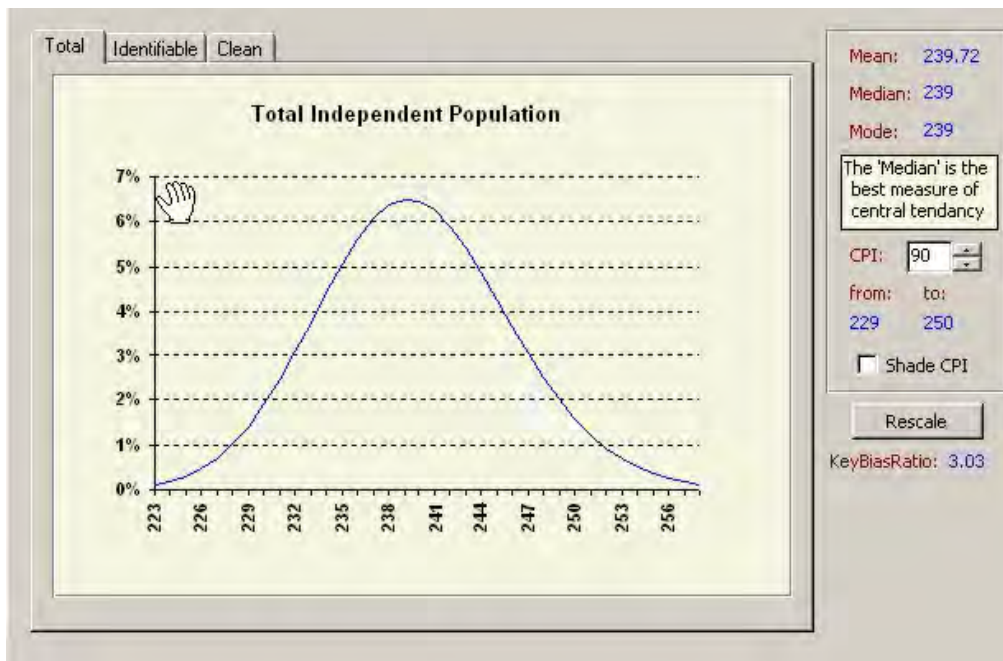
At the end of a multi-area analysis the "Identifiable" tab displays the graph of the total combined ID independent population for all areas (. It is derived by combining the various individual area ID independent cross boundary movement bias corrected posterior distributions.
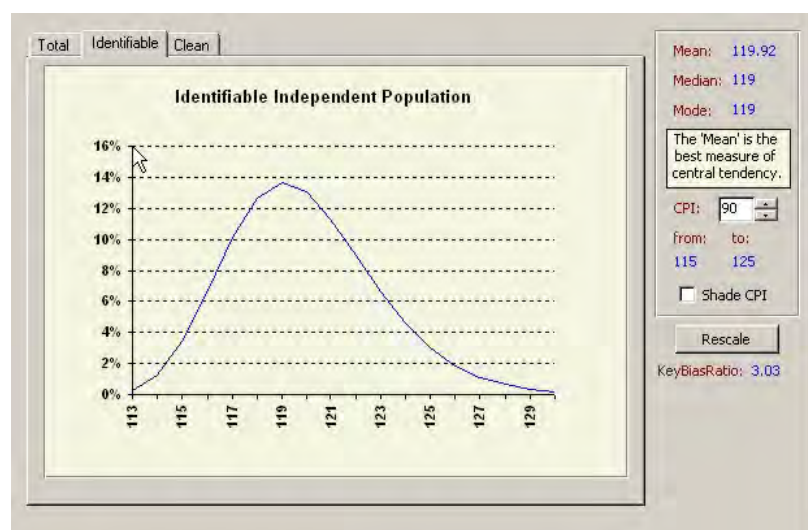
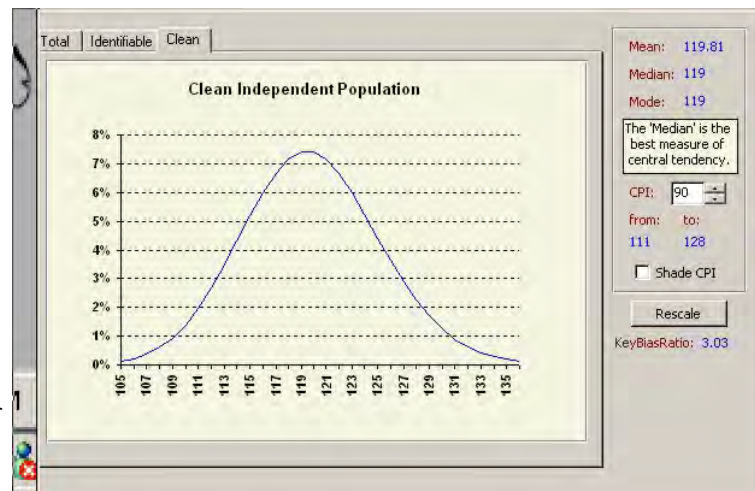At the end of a multi-area analysis the " Clean" tab accesses the graph of the total combined Clean independent population for all areas. It is derived by combining the various individual area Clean independent cross boundary movement bias corrected posterior distributions.

In a multi-area simulation the "< Back" button is disabled.

Dependent Graphs

After viewing the total independent distribution click on the "Next>" button to bootstrap estimates of the number of dependents.

USE THE TABS TO ACCESS THE THREE DIFFERENT TABS ..

By clicking on the tabs at the top left of the graphs, one can view any of the three dependent graphs associated with this form. When you click on a tab that graph will come to the front and be displayed. In all three graphs (as with previous posterior probability graphs), the blue line shows the probabilities of the chosen population segment being each of a range of X axis values. Population estimates (mean, mode and median) and the Bayesian equivalent of confidence levels (CPI values) are calculated from these graphs in the same way as before. Once again RHINO's statatistical provides advice on the best measure to use on each graph. As in previous graphs, users can chose to re-scale the graphs, change CPI levels, and shade CPI levels.

Document TitleTotal Dependents Graph

"TOTAL" DEPENDENTS GRAPH

The default graph initially shown by RHINO shows the posterior probability distribution for all dependents of independents (this represents the total number of A-D age calves in a rhino population). This graph is obtained by combining the posterior distributions for the other two graphs on this form (dependents of ID independents and dependents of clean independents).
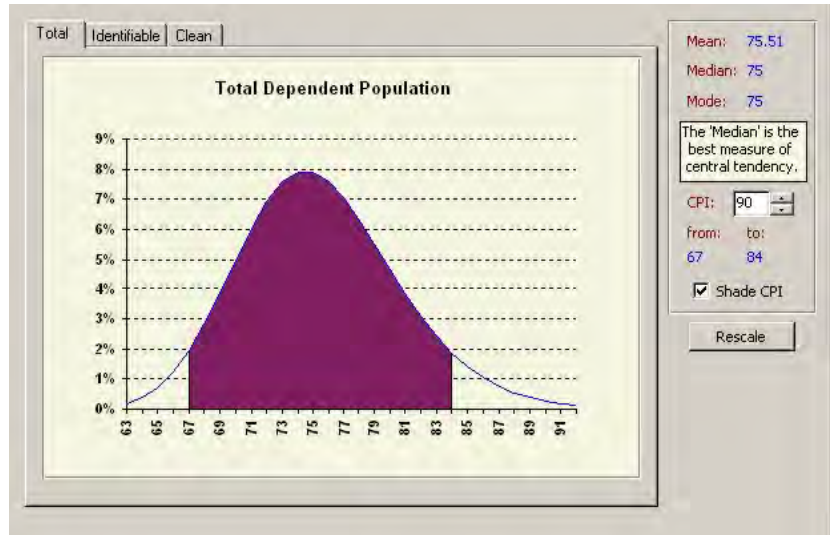
The tab for this graph is labeled "Total". This graph will refer to the whole area in a single simulation or to a single area in a multi-area simulation (if you are busy with the individual area analyses).

COMBINED DEPENDENTS OF INDEPENDENTS GRAPH AT THE END OF A MULTI-AREA ANALYSIS.

The total dependent estimate (for all areas) produced at the end of multi-area analysis is obtained by combining the combined ID dependent distribution for all areas, and the combined clean dependent distribution for all areas.
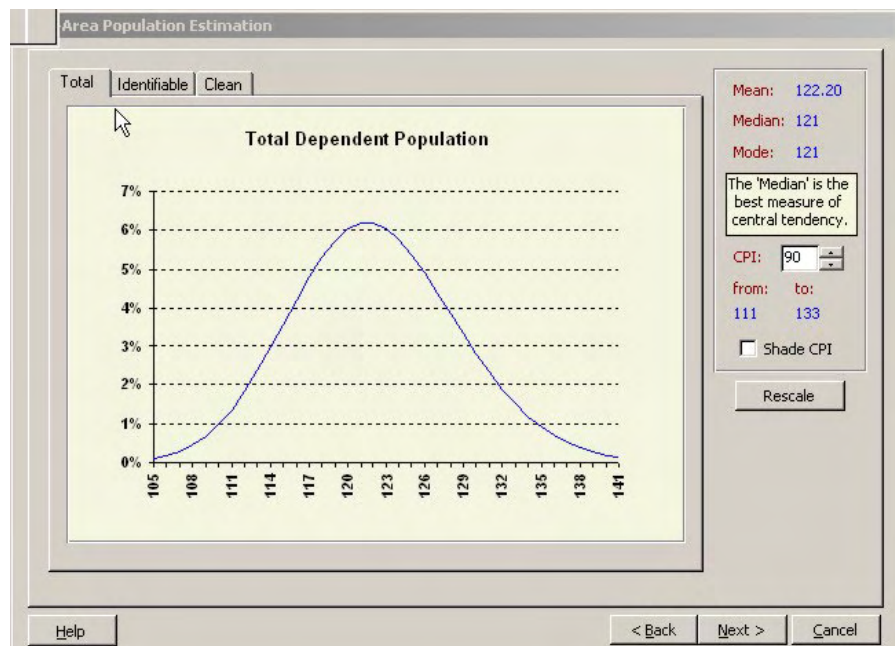
Dependents of ID Graph

GRAPHS DURING AND NEAR THE END OF A MULTI-AREA ANALYSIS …

If you are currently analyzing an individual area in a multi-area analysis the graphs will refer to this area only. However once you have finished analyzing all areas separately in a multi-area analysis, combined total dependent graphs will be generated for all areas combined.

DEPENDENTS OF ID INDEPENDENTS GRAPH

The second (middle) graph on the form with the tab labeled "identifiable" shows the posterior distribution of the number of dependents of ID independents (i.e. in rhino the terms the number of A-D age calves of identifiable adults). N.B. in this case "Identifiable" refers to the calves' mothers, rather than the calves themselves (which may or may not be identifiable).
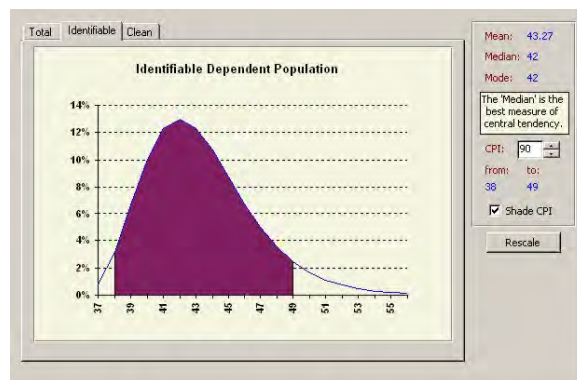
DEPENDENTS OF CLEAN INDEPENDENTS GRAPH

The third (right) graph with the tab labeled "Clean" shows the posterior distribution of the number of dependents of ID independents (i.e. in rhino the terms the number of A-D age calves of clean adults). N.B. in this case "Clean" refers to the calves' mothers rather than the calves themselves (which may or may not be clean). Note that recommended best measures are not the same for the above two graphs (median for ID dep and mode for clean dep.)
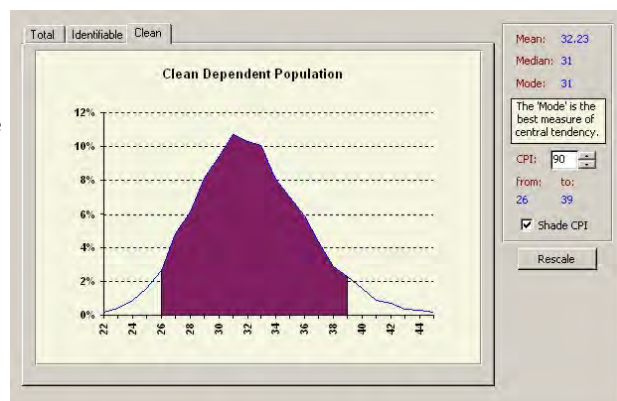
TO PROCEED...

To proceed click the "Next >" button. To return to the previous form click the "< Back" button.

In a multi-area simulation the "< Back" button is disabled.

Total Population Estimate, Reporting and Closing run

TOTAL POPULATION ESTIMATES ..

The next graph shows the posterior probability distribution for the whole population (or individual area or all areas combined in a Multi-area analysis) and is obtained by combining the relevant total Independent and total Dependent posterior distributions.

As with previous posterior probability graphs, the blue line shows the probabilities of the chosen population segment being each of a range of X axis values.
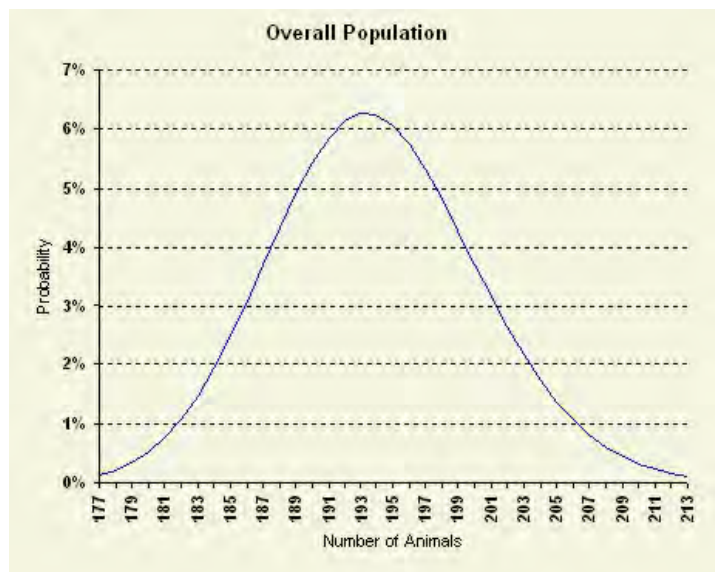
Population estimates (mean, mode and median) and the Bayesian equivalent of confidence levels (CPI values) are calculated from these graphs in the same way as before. As in previous graphs, users can chose to re-scale the graphs, change CPI levels, and shade CPI levels.

Producing a RHINO analysis report

PRODUCING RHINO ANALYSIS RUN REPORTS..

Reporting has been significantly improved in *RHINO 2.0.* Sightings frequency /truncated Poisson fit and posterior probability graphs are now automatically inserted into the report which is in the form of a MS Word document. To produce a report click the "Report >" button at the bottom of the overall population estimate form.

Due to limitations in the size of the report allowable in Access, the *RHINO* run report in *RHINO 2.0* is produced using MS Word and MS Excel. Users therefore have to have Word and Excel loaded on their computers to generate the reports. The graphs for the report are produced using Excel and inserted into a Word document. The creation of the report may take a little time (especially if you have been doing a multi-area analysis with six areas). For this reason *RHINO* asks you to confirm you want to produce a report. Please be patient.
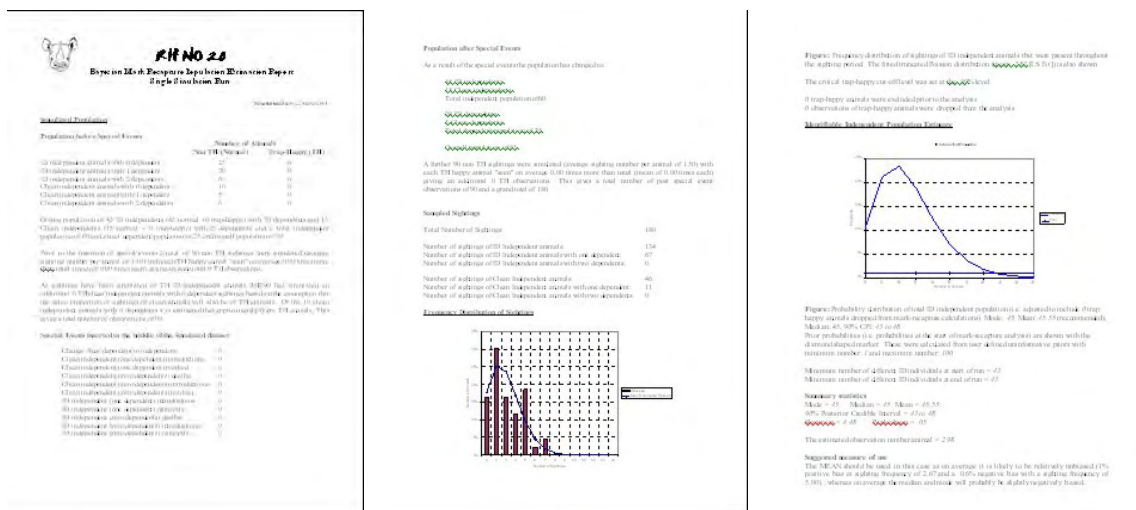
At the end of simulations users can also select to create a simulation run report. The above example shows the first 3 pages from a *RHINO* simulation report .

During a multi-area analysis you will have the option of generating separate individual area reports after you have analysed the data for each area. You should only select to do this if you want to distribute separate individual area reports, because the content of the individual area reports will be included in the final multi-area report produced at the end of the analysis.

Saving simulated data

PROPOSED IMPROVEMENTS IN FUTURE

In the next major revision of *RHINO* (version 3.0) it is hoped to produce the reports using Delphi and so users will not have to have Word and Excel to produce reports.

ENDING THE RUN...

To end the run after completing a single or multi-area analysis (and clean all the datatables used by *RHINO*) click the "Close" button at the bottom right of the form. This also lets *RHINO* clean and empty temporary tables used in the analyses.

OPTION OF SAVING A SINGLE-RUN SIMULATION DATASET
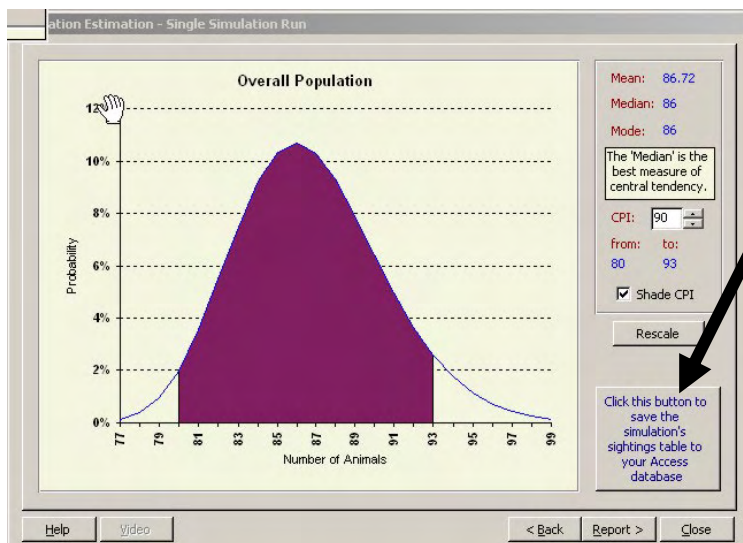
If you have been running a single-run simulation, you now have the option of saving your simulated data (which you can then re-analyse at a later date). Simply click the box button at the bottom right of the table and follow instructions to save your simulated multi-area data.

TO PROCEED IN A MULTI AREA SIMULATION ...

If you are analyzing data at an individual area level the graph will show the total estimate for the current area being analysed. The area's name will appear in the form heading. When you are ready to proceed click the "Next >" button.

Multi-area analysis

TOTAL COMBINED POPULATION GRAPH FOR ALL AREAS IN A MULTI-AREA SIMULATION..

In a multi-area simulation the total population posterior distribution is also obtained by combining the total independent and total dependent distributions.

Thus in a multi-area analysis total population estimates will be produced for each area separately as well as for all-areas combined at the Park level.

In a multi-area simulation the "< Back" button is disabled.

OPTION OF SAVING MULTI-AREA SIMULATION DATA

If you have been running a multi-area simulation, and have generated the total combined population estimate for all areas combined, you will have the option of saving your simulated data which you can then re-analyse at a later date. Simply click the box button at the bottom right of the table and follow instructions to save your simulated multi-area data.

Area abbreviations to be used in Multi-area cross tabulation table.

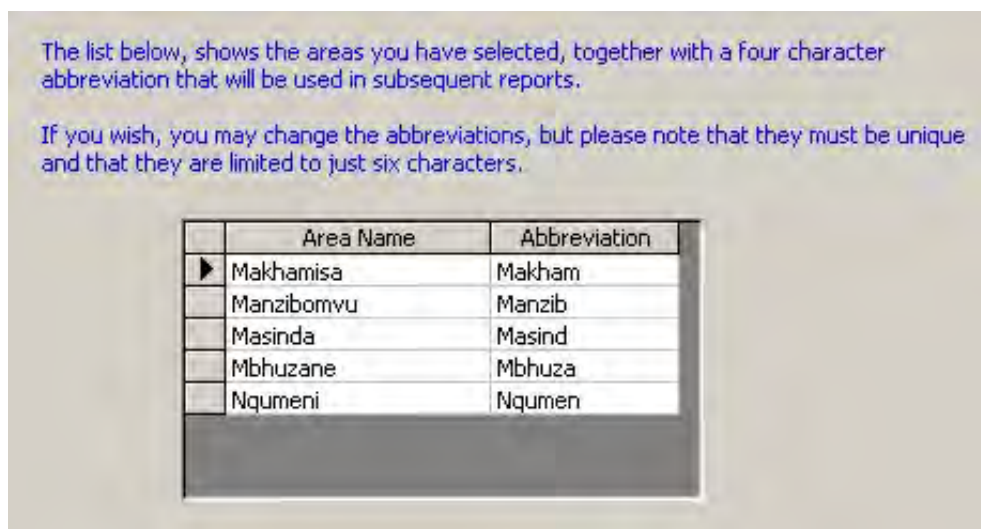DEFAULT ABBREVIATIONS ..

If you select to undertake a multi-area analysis this form is used to show the default 6 character abbreviations RHINO will use in tables to refer to different areas (unless you specify otherwise).

Multi-area Abbreviations

NEED TO ENSURE THAT ALL ABBREVIATIONS ARE UNIQUE ..

A list will now appear showing the six character acronyms *RHINO* is proposing to use as area name column headings in a sightings summary table. By default *RHINO* selects the first 6 characters of each area name. If you want to use different abbreviations please edit the abbreviations in the table. *RHINO* automatically checks to see that each abbreviation is unique, and if not will warn the user. To proceed with a multi-area analysis simply click on the "Next >" button on the bottom of the form.

The list below, shows the areas you have selected, together with a four character abbreviation that will be used in subsequent reports.

If you wish, you may change the abbreviations, but please note that they must be unique and that they are limited to just six characters.

Area Name	Abbreviation
► Makhamisa	Makham
Manzibomvu	Manzib
Masinda	Masind
Mbhuzane	Mbhuza
Nqumeni	Nqumen

EDITING ABBREVIATIONS ..

You can use this form to change any of the default abbreviations suggested by RHINO either to make them more recognizable or to ensure each abbreviation is unique. Simply click on the abbreviation you want to change and type in your new abbreviation.

TO CONTINUE,,

When you are happy with the abbreviations (as shown on the form) click the "Next > button to continue".

Multi-area sightings summary table

Multi-Area sightings summary table, Allocation of ID animals to areas for analysis and Cross boundary correction factors.

CROSS TABULATION OF SIGHTINGS OF INDIVIDUAL ID ANIMALS BY AREA…

This table shows the results of a cross-tabulation of the number of sightings of each ID animal seen in each area. In the example below there are five areas.

Area Population Estimation

Rhino Name	Analysis Area 1 Makham	Analysis Area 2 Manzib	Analysis Area 3 Masind	Analysis Area 4 Mbhuza	Analysis Area 5 Nqumen	Analysis Area 6 -	Area all Sightings allocated to...	Total sighting number/ animal	Sightings in allocated area
Brs592	3		2	2			Makham	7	42.86%
W870	1			1			Makham	2	50.00%
Brs569	9		4				Makham	13	69.23%
W851	5			2			Makham	7	71.43%
Brs586	6			2			Makham	8	75.00%
C414	3		1				Makham	4	75.00%
Brs572	17		2	1			Makham	20	85.00%
Brs568	6		1				Makham	7	85.71%
W862	13			2			Makham	15	86.67%
Brs600	20			3			Makham	23	86.96%
Brs585	8			1			Makham	9	88.89%
Brs597	16			2			Makham	18	88.89%
Brs543	3						Makham	3	100.00%

Original data (before reallocation of cross boundary sightings)

Total:	277	181	291	642	176			1567	
Animals:	52	30	58	87	30				

After reallocation of cross boundary sightings for analysis purposes

Total:	264	175	279	665	184			1567	
Animals:	33	26	47	70	28				
XbndryCF:	1.05	1.03	1.04	0.97	0.96				

Print

Help   Video   < Back   Next >   Cancel

Allocation of sightings to areas

A "Print" button has been attached to the form so you can print out copies of this table to provide feedback to the people responsible for rhino monitoring in the different areas of a Park. This button may not work on some machines.

NUMBER OF AREAS THAT CAN BE ANALYSED IN A MULTI-AREA ANALYSIS..

In the Multi-area simulation only two areas are defined. However when analysing data users can analyse up to 6 different areas in RHINO 2.0.

ALLOCATION OF SIGHTINGS OF ID ANIMALS TO A SINGLE AREA PRIOR TO ANALYSIS..

For analysis purposes (primarily to reduce bias caused by cross-boundary movement of animals) RHINO allocates all the sightings of each animal to the area it was most often seen in. In the event of a tie, RHINO uses random numbers to decide which area to allocate sightings to. As a result, if you re-run a Multi-area analysis, sightings of animals with tied maximum numbers of sightings in different areas may be allocated to different areas for analysis purposes. Thus the results of repeat analyses of the same dataset may differ slightly as a result.

The table also shows the percentage of the total sightings of each ID animal in the area all its sightings were allocated to, as well as specifying the areas to which all sightings of the different ID animals have been transferred to for analysis purposes.

SUMMARY OF PRE- AND POST-REALLOCATION SIGHTING NUMBERS...

At the bottom of the table, the total number of ID animals and sightings both before and after sightings re-allocation are given.

Cross boundary movement correction factors

Note that the number of different animals seen in each area is usually higher before re-allocation of sightings. This is because cross boundary animals are "seen" in more than area (and hence double counted) before (but not after) re-allocation of sightings. If not corrected for, this can bias estimates upwards.

CALCULATION OF AREA CROSS BOUNDARY CORRECTION FACTORS (XbndryCF)...

Finally, the form gives cross-boundary correction factors (XbndryCF) calculated as the number of sightings before re-allocation divided by the number of sightings after re-allocation. If an area ends up with more ID sightings after re-allocation than before, its XbndryCF will be < 1. Similarly if ends up with fewer sightings the XbndryCF will be >1. If an area has exactly the same number of ID sightings before and after re-allocation the XbncryCF will be 1.

USE OF XbndryCF TO CORRECT ID AND CLEAN POSTERIOR DISTRIBUTIONS AND HENCE CORRECT ESTIMATES...

If proportionately more sightings have been allocated to an area, this may lead to an over-estimation of the number of ID animals in that area. Similarly, the allocation of fewer ID sightings to an area, will lead to an underestimation of the number of ID animals in that area.

ID population estimates can be corrected by multiplying estimates by the XbndryCF to get a corrected mean estimate. If there are more ID sightings in an area after re-allocation than before, the XbdryCF will act to reduce the ID population estimate. Conversly it will increase it if fewer sightings have been added to an area. The differences in un-corrected and XbndryCF corrected mean ID estimates for the different areas are then rounded to the nearest integer, and the ID posterior distributions for the different areas moved up or down by these amounts to produce cross boundary bias corrected final posterior probability distributions from which corrected ID population estimates for the different areas can be obtained. By correcting for possible bias, these corrected area estimates are likely to more realistically estimate the average density of rhinos present in each area.

Cross boundary movement bias correction

If more ID sightings have been allocated to an area after re-allocation, this will increase the ratio of ID sightings to clean sightings. This in turn will lead to an underestimation of clean numbers in the area. To correct for this the Clean independent mean is divided by the XbndryCF to get a (slightly increased) corrected mean estimate.

The differences in un-corrected and corrected mean Clean independent estimates are then rounded to the nearest integer and the Clean posterior distributions are moved up or down by half this amount to produce cross boundary bias corrected final posterior probability distributions from which corrected Clean population estimates for the different areas can be obtained.

Conversely the XbndryCF will act to reduce clean estimates if an area has fewer ID sightings than before after reallocation.

Thus, if more ID sightings have been allocated to an area in a multi-area analysis, the derived XbndryCF for the area will act to adjust the ID independent estimate downwards and the Clean independent estimate upwards (to correct for any biases introduced by the re-allocation of cross-boundary sightings for analysis purposes). Similarly if fewer ID sightings have been allocated to an area in a multi-area analysis, the derived XbndryCF for the area will act to adjust the ID independent estimate upwards and the Clean independent estimate downwards.

Area to be analysed next in Multi-area analysis ...

AREAS IN THE MULTI-AREA ANALYSIS AND AREA TO BE ANALYSED NEXT

The box at the top of the form shows the areas selected in a multi-area analysis and their abbreviations. A multi-area analysis proceeds by analysing the data for the first area in the list, then the second are, and so on. When data for the last area have been analysed, RHINO proceeds to combine the probability distributions from the different areas to produce total Park level population estimates.

The next area to be analysed is highlighed on the form for your information. No user input is required here except to click the "Next>" key to proceed or the "< Back" key to return to the previous form.

Text detailing how to use simulation options to be added….

Form 17 – Specifying Multi Area simulation

Form 18 – Specifying Multiple Simulation

Form 19 – Summary of pop size at the start and button to start multi simulations

Form 20 –   Multi-simulation results summary – ID independents

Form 21 –   Multi-simulation results summary – Clean independents

Form 22 –   Multi-simulation results summary – Total independents

Form 23 –   Multi-simulation results summary –

Dependents of ID independents

Form 24 - Multi-simulation results summary - Dependents of Clean independents

Form 25 - Multi-simulation results summary - Total dependents

Form 26 - Multi-simulation results summary - Total population including report generation, saving of results, and closing

Form 27 - Single run simulation - entering of basic specs

Form 28 - Single run simulation - entering of special codes and popn summary.

Form 29 - Single run simulation - summary form before and after special codes.