

---

# THE BAREFOOT ECOLOGIST'S TOOLBOX

DR JEREMY PRINCE

LAST UPDATED : MARCH 2021

## Contents

### 5 MAKING LENGTH-BASED SPAWNING POTENTIAL RATIO ASSESSMENTS

7	Introduction
8	Completing an LBSPR Assessment
13	Interpreting an LBSPR Analysis
22	References



C H A P T E R E I G H T

# MAKING LENGTH-BASED SPAWNING POTENTIAL RATIO ASSESSMENTS

## Introduction

---

The aim of this chapter is to provide a step-by-step manual for any reader wanting to analyse size composition data with the web-based LBSPR software available for free use at: <http://barefootecologist.com.au/>.

Previous chapters described how excel spreadsheets downloaded from the biospherics.com.au website can be used to collect data, derive estimates size of maturity and life history ratios, and convert the length composition data into the .csv format required for analysis with the LBSPR software. The assumption of this chapter is that you have already worked through the previous chapters and have all the inputs required in the formats required. The intention is that you will open the software, and be trying it out on your own dataset as you work through this chapter. The chapter concludes with some discussion of how the results of the LBSPR assessments should, or can be interpreted, as well as, some discussion of potential pit-falls, and other issues users of this approach should be thinking about as they apply the technique and interpret its results.

For readers who are proficient with 'R', the algorithms to apply the LBSPR software in R, and also a 'shinyapp' that will run the application off-line, can also be freely download from the website as well. So you may well choose to incorporate that code into your own applications in R, or just to run the shinyapp with out going on line. But those who want a simpler access to the algorithms on line and to avoid working in R can simply use this web-based version.

# Completing an LBSPR Assessment

## Opening Pages

The opening page of the Barefoot Ecologist software website provides an introduction to the website. Clicking on the central 'Go to Applications' button takes you to the next page from which three software packages can be accessed. Our interest here is completing LBSPR assessments, so find and click the 'Go to App' button under the title 'LBSPR Model' on the left. This will take you to an introductory page for LBSPR assessment on which you can read about using the software. Very basic explanatory notes, compared to the version you will have read if you have worked your way methodically through this tome.

Scrolling further down this page will bring you to the bar of Tabs used to construct an assessment (figure 1). If your access to the internet has very limited bandwidth this bar of Tabs can take some time to appear, and might even fail to appear all together. My only advice is to be patient or try again at another time of day when you might expect there to be less internet traffic. When it does eventually appear click on the 'Upload Data' tab.

## Uploading Data

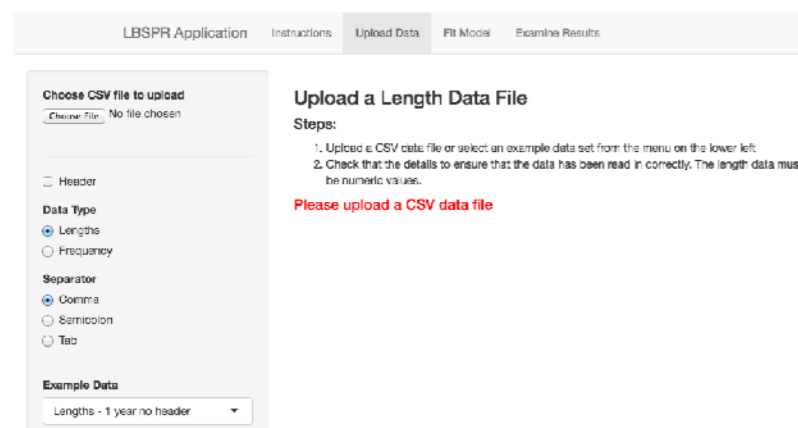


Figure 1. The Upload data page with the bar of Tabs at the top used to navigate around the LBSPR software and construct an assessment.

The Upload Data tab (figure 1) enables you to search for the .csv data file you created and saved with the Excel workbook in the last chapter. The data could have been saved as a raw list of measurements, with or without a heading. Or as a tabulated length frequency table, with a list of mid-points in Column A and counts or proportions of a sample in the subsequent columns. With both types of data multiple years of data can be uploaded

and analyzed in each data file, with the columns to the right containing the sequence of samples.

Remember the data file must have been saved in the comma separated value (.csv) format. Files in Excel format will not be recognized by this software even when you locate them on your computer.

Use the Browse button to find your data file and select the upload option. Once you have found and chosen your .csv file to upload it, the buttons below are used to describe your data so the software can correctly interpret the data.

- Do the columns have headers, or not?
- Is the data a list of lengths or a length frequency?
- Which sort of punctuation is used to separate the values, comma, semicolon or tab?

To the bottom right the software will show you a preview of how it is interpreting the first few lines of your data file. The software also provides red error messages or instructions, if it encounters aspects of data files that do not appear to match the format it is expecting. Once you have explained the data to the software's satisfaction the messages will disappear and you can proceed. Click on the 'Fit Model' tab which will take you to the page in which you will enter the input parameters.

## Uploading Input Values & Plotting data

On the 'Fit Model' page (figure 2):

1. With the 'Absolute' or 'Relative' button select the version of the software you intend using. The 'Absolute' button is used when you have an accurate independent local estimate of asymptotic size ( $L_{\infty}$ ) and the software will then give you the option to enter your estimate of  $L_{\infty}$ . Most of us barefoot ecologists can only dream about having good estimates of  $L_{\infty}$ , in which case we use the 'Relative' button. The software will then give the option of entering an estimate of  $L_m/L_{\infty}$  and using that value to infer the  $L_{\infty}$  of the assessed population from our estimate of  $L_{50}$ .

Six years spent working on a meta-analysis of growth studies have left me very aware of how difficult  $L_{\infty}$  is to estimate

accurately when fishing almost everywhere has removed most older larger animals from the studied populations. I am automatically distrustful of most locally derived estimates of  $L_{\infty}$ . In my opinion, well founded estimates of  $L_m/L_{\infty}$  and local  $L_{50}$  are easier and less technical to estimate than the  $L_{\infty}$  of populations, and will almost always end up producing a more accurate stock assessment, than going the 'Absolute' option and using a single local estimate of  $L_{\infty}$  derived for an otherwise data-poor fishery

2. Under the 'Species' heading enter the name of the file being analyzed. This does not have to be a species name, the software will use whatever you enter as a file name for any assessment output you choose to download.

3. Enter estimates of your two Life history Ratios; M/K and  $L_m/L_{\infty}$  if you have selected the Relative option, or M/K and  $L_{\infty}$  if you have selected the Absolute option

4. Enter estimates of  $L_{50}$  &  $L_{95}$

5. The software will compute and display the asymptotic size based on the  $L_{50}$  &  $L_m/L_{\infty}$  you have entered.

6. If you have entered length measurements (rather than size frequency data) you will see a slide bar at the bottom left (figure 2) which can be used to adjust the width and number of length classes the model uses to construct and analyse the length-frequency histogram with. Select a bin width.

7. Hit the Plot Data button

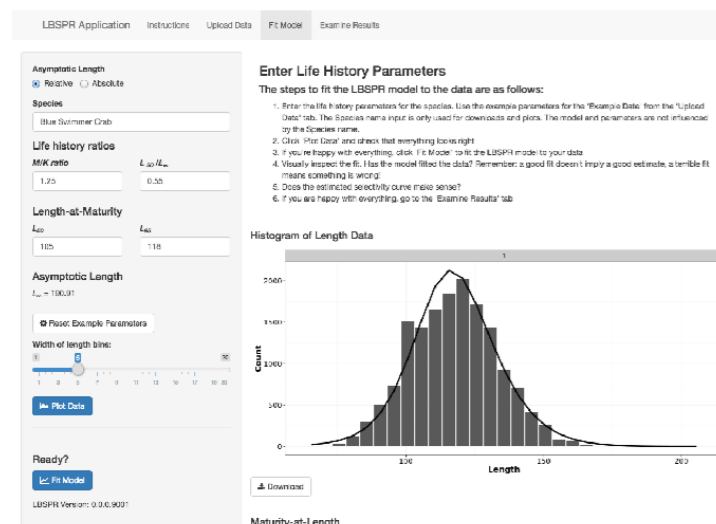


Figure 2. The Fit Model page used for entering input parameters, plotting data and fitting the data. The plotted data are shown by the black columns in the length frequency histograms in the bottom right quadrant and the curve fitted to the data by the LBSPR software is shown by the black line.

## Adjusting Bin-width of Data and Fitting Model

You will then see your size composition data plotted as black bars on the bottom right corner of the page. If you have uploaded data in the form of columns of raw length measurements you will now be able to re-adjust the width and number of length categories, if you wish, with the slide bar at the bottom left.

Remember that, ideally you should have >10 length categories or 'length bins' and your data should look like some nice smooth unimodal curve, something like the curve depicted in figure 2. The size structures being analyzed with this software are compilations of size classes from multiple age cohorts, sculpted on the left-hand side by the size-selectivity of the catching process ( $SL_{50}$  &  $SL_{95}$ ) and on the right by the fishing down of the size classes growing through the fishery. Based on the underlying model assumptions in the perfect world these processes will always result in a single more-or-less bell-shaped size composition. Use the bin-width button to change the shape of your data to be smoother or lumpier and more or less bell shaped. If you can get a reasonably smooth bell and have >10 size bins, it can be considered a relative success. Well done. If not, sorry, but your size composition is really a bit of a fail. You probably need a bigger sample size. There can be other issues as well, which will be discussed later this chapter. At this stage, your data is what it is, and you are going to push through and analyze it anyway, damn the data quality!

Once you are as satisfied as you can be with the way your uploaded data looks hit 'Fit Model' button.

## The Fitted Model with Outputs I

Within a few seconds the software will have fitted its own black curve depicting its description of the data (figure 2). Scrolling down will also bring up plots of the size of maturity curve (red) defined by your estimates of  $L_{50}$  and  $L_{95}$ , overlaid with the modelled estimates of the logistic curve that defines size selectivity (blue), the size fish are selected for catching by the fishing gear ( $SL_{50}$  &  $SL_{95}$ ).

Note the 'Download' buttons that can be used to download these images.

## Interpreting an LBSPR Analysis

---

### What Does F/M Mean?

Until this point of the book, I have deliberately minimized discussion of the second management metric estimated by LBSPR assessment; the ratio F/M. This has been for the sake of simplicity and brevity, because F/M is a little more technical fisheries concept.

As discussed in chapter 5, in fisheries assessment science, F is used to signify the instantaneous rate at which the fish are dying because they are caught, and M the instantaneous rate at which the fish are dying from natural causes (i.e. predation, disease, accident). The ratio F/M, called relative fishing mortality, is widely used as a performance indicator for fisheries management. As a general rule if fish are only vulnerable to being caught as adults (i.e.  $L_{50}=SL_{50}$  &  $L_{95}=SL_{95}$ ) maximum sustainable yields (MSY) should result when fishing mortality is similar to natural mortality. So  $F/M = \sim 0.8 - 1.0$  is considered a proxy for sustainability and  $F/M > 1.0$  is considered overfishing.

As a performance metric for management, F/M is less useful than SPR. This is because the sustainable level of F/M depends on the size selectivity of the catch ( $SL_{50}$  &  $SL_{95}$ ), relative to the size of maturity ( $L_{50}$  &  $L_{95}$ ). If fish can be caught smaller than the size they mature ( $SL_{50} \ll L_{50}$ ) then low F/M ( $< 1.0$ ) can still cause overfishing, because many fish will be caught before maturing. On the other hand, if the fish are only caught at sizes much larger than the size of maturity ( $SL_{50} \gg L_{50}$ ) high levels of SPR can be conserved even with very high fishing pressure ( $F/M > 2$ ). So as a metric of management performance F/M must always be interpreted within the context of size selectivity. The metric of SPR is much more useful because it incorporates both the effect of fishing pressure (F/M) and the size at which the fish are caught ( $SL_{50}$  &  $SL_{95}$ ).

It also turns out the LBSPR algorithms tend to produce stable, robust estimates of SPR under most circumstances, while the estimates of F/M can be much noisier, especially under regimes of heavy fishing pressure. The explanation of why this is so, gets quite technical and I do not need intend expanding on it here. The implication on both accounts is however, that the SPR estimates are easier to interpret and use than the estimates of

### Model Outputs 2

By selecting the 'Examine Results' tab you can access results in a tabulated format; estimates of SPR,  $SL_{50}$  &  $SL_{95}$  and F/M, each with estimated 95% confidence intervals in brackets. Also a color coded pie diagram of the SPR assessment of SPR, the final year in the case of a time series analysis. With an assessment of  $< 20\%$  the pie chart only shows the remaining red piece of the pie. At this level the fishery is expected to have declining levels of young fish recruiting to the stock or to be below the point of recruitment impairment (PRI). With estimates  $> 20\%$  the pie chart will develop an orange fringe to indicate that SPR is above PRI, but still below the target level of SPR 40%. At SPR 40% the pie chart will develop its green fringe indicating the stock is above the internationally accepted default level of a well-managed fishery aiming for optimal sustainable yields.

Both of these forms of output can be downloaded using their respective download buttons. The output table as a .csv table which can be opened by any spreadsheet or word processing software., and the diagram in jpeg or PNF formats.

F/M, although that does not mean they should be disregarded altogether as a guide to the level of fishing pressure being asserted on stocks.

## Things that Typically Go Wrong with LBSPR Assessments

Now our focus turns to trouble shooting some of the issues you will inevitably end up confronting if you start making regular use of the LBSPR assessment methodology.

### LOW QUALITY SIZE COMPOSITIONS

As has been discussed previously, one of the most basic assumption underlying the LBSPR technique is that the data being analysed accurately represents the adult part of the population. Once you have fitted the LBSPR model to your data you will get some sense of how well you have done in this regard, by comparing your plotted data with the line depicting the model's fit (figures 2-4). Ideally the line should pass through the top of each bar close to its centre, leaving no large jagged gaps between the data and line. The processes that create the size compositions of the adult populations under equilibrium conditions always create a relatively smooth bell like shape and consequently the LBSPR algorithms are constrained to expecting your data to conform to that sort of shape (figures 2-4). For there to be a good fit, your data will need to describe a relatively smooth bell-shaped curve, as can be seen in figure 2 & 3 but not in figure 4.

The datasets portrayed in figures 2 & 3 are described reasonably well by the LBSPR assessment model. Although not perfectly. In each case there are spikes of data that poke up above the fitted curve, and to greater or lesser extents corresponding gaps below the fitted line. With a little more work on the quality of the data being collected, and larger sample sizes, the LBSPR model could be expected to be fitting these datasets almost perfectly and producing very high-quality estimates of SPR and F/M.

In contrast things are going very wrong with the datasets depicted in figure 4. For most size classes the fitted curve passes well above or beneath the data, leaving large gaps between the

curve and data. From this we see that the LBSPR model cannot match and explain these data. Something is going badly wrong with these samples, so that the model cannot 'understand' or match the data with its fitted curve.

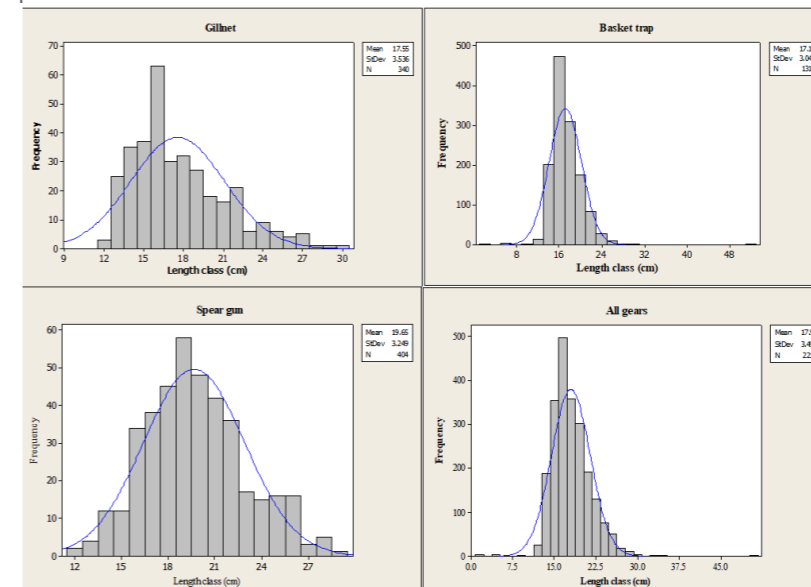


Figure 3. Four different LBSPR assessment model fits (blue line) to length composition data (grey bars). The extent to which the model can explain the data can be gauged by how closely the line passes through the centre of the top of each bar. In all these cases with relatively small sample sizes the model is fitting relatively well to the data, with the model fit in the right-hand panels being much better than in the left-hand panels. The resulting assessments are likely to be reasonably good although attention to precise measurement and collecting more data would improve them.

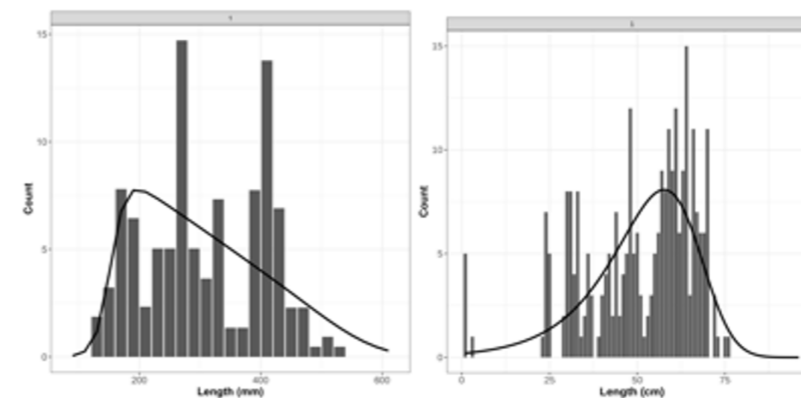


Figure 4. Two more LBSPR assessment model fits (black line) to length composition data (black bars). Note the lack of correspondence between the fitted curve and the plotted data, which reveals that the model has been unable to predict these data. The resulting assessments are likely to be inaccurate at best, and highly misleading at worst. Serious attention is needed to solve the problems with these data, simply collecting more of this type of data will not correct the issue.

### SMALL SAMPLE SIZES

Looking at the y-axis of both panels in figure 4 you will notice that most size bins having counts of <5 individuals, the overall sample sizes are quite small ~100 - 200 fish. Small sample sizes are the major cause of spiky looking data that the LBSPR algorithms have trouble fitting. Even if measuring is precise in



small samples, just by chance, some size classes will end up being under-represented while others will be over represented. If measuring is precise, as sample sizes increase many of the peaks and troughs will naturally tend to even out.

The largest size classes which are naturally rare in any population are particularly prone to being under-represented in smaller samples, which, even when small samples produce nice smooth histograms, will still cause SPR to be under-estimated and F/M to be over-estimated. This is because the LBSPR method is strongly influenced by the size of the largest fish in a sample. Statistical studies show that sample sizes of 1000 are required to ensure the largest individuals are fully represented (Erzini 1990) and ideally, samples sizes greater than 1000 individuals would always be available for analysis so that the largest individuals in each population are fully represented (Hordyk et al. 2015). In the real world of reef fish assessment, sample sizes of more than 100 fish, let alone 1000 individuals, can be difficult to accumulate, and it is necessary to use whatever data are available. My own experience with collecting length frequency data is that, if your measurers are being accurate you normally end up with a pretty smooth, rounded bell shape, by the time you have measured 150-250 fish. I have applied the methodology to reef-fish samples of ~100 fish to produce apparently robust indicative results (Prince et al. 2015, 2020). Babcock et al. (2018) and Hommik et al. (2020) report successfully applying the technique to samples as small as  $n \sim 60$ . During the development of LBSPR, our simulation testing (Hordyk et al 2015) demonstrated that indicative assessments (i.e. heavily fished, moderately fished or lightly fished) could be produced with even smaller sample sizes ( $n > 30$ ). Although it should be remembered that if sample sizes can be subsequently enlarged your estimates of SPR can be expected to increase marginally, and F/M decrease slightly, as the representation of the largest individuals in the population is improved.

In the context of small sample sizes, the 'equilibrium' assumption can be useful, because it means that samples can legitimately be accumulated for analysis over several years. We biologists tend to be focussed on monthly and yearly sampling time periods, and tend to naturally assume that aggregating data over several years is invalid. This is not the case for LBSPR.

With medium to long lived species it is valid to aggregate data over several years. The rule of thumb here is that, catch size composition can be aggregated over the time frame it takes an adult fish to grow through the fishery, because that is the time period over which a regime of fishing pressure will change the adult size compositions you are analysing. In many cases the answer to the problem of small sample size is just to double down and extend the data collection program, but before doing that, it might be worth while checking on the accuracy of the measurements being made.

#### **INACCURATE SAMPLING, FAVOURITE NUMBERS AND ROUNDING OF NUMBERS**

If samples sizes are reasonable (say  $> 250$  measurements) and your histogram is still spiky, perhaps getting spikier, then you should suspect problems with the sampling, inaccurate measuring, identification or data recording, possibly all of the above. Even if you suspect nothing you need to be aware of what commonly goes wrong with data collection, because if it they are not happening now, you will encounter them in the future. Look closely at figures 2-4 and you will see that in most cases the main data spikes correspond relatively closely to sizes that end in 0 or 5, suggesting some issue with observers rounding off their measurements. Data collectors find it difficult not to round off to even numbers or the nearest 5 or 10 units because we are all so naturally, and unconsciously good at, saving energy by cutting corners. Indirectly the gaps below the line of fit in figure 2 & 3 will have the same cause, because the rounding of measurements to numbers ending in 0 or 5, creates a shortfall in surrounding measurements, the ones that do not end with those digits.

A range of other factors can also contribute to creating horribly spiky data. I once investigated a case of corrupted data and discovered that the person writing down the measurements which were meant to be to 0.1 of a cm, was inserting zero after each 'point' so that 35.5 would be recorded as 35.05. This had the similar effect of making the data spike at cm intervals. Another issue that arises with measuring fish on a board is that the eye is drawn naturally to the mark on the tape that can be seen beyond the end of the tail, when in fact a mark hidden beneath the tail maybe closer to the end of the tail. Some measurers

may just have favourite numbers that they unconsciously round towards. This is more common when people are just fabricating the data. I once busted a cheating fish measurer by plotting up the frequency of the last digit of the data we were paying him to collect. Slightly less than 50% of his measurements ended in 3 or 7 when they each should have been about 10% of the data.

Whatever the problem is or are, if the fitted curve does such a poor job of describing the data, as is illustrated in figure 4, you should not proceed with the analysis because it will be hopelessly inaccurate. You will see this by looking at the estimated 95% confidence intervals around the SPR and F/M estimates. They will be very broad indicating the average estimates are being poorly informed by the data. You might even receive a warning message telling you the LBSPR model has failed to converge on any unique estimates. Better to put your effort into improving your data collection systems and to double down on collecting more better data.

All of the above issues can combine to screw up the data you are trying to collect and will mess up the accuracy of the LBSPR assessments. The answer to these problems is obviously first to get the measurer to actually measure the fish and not sit at home, or in the bar, filling out the forms from their own imagination. The second is to sharpen up the accuracy of the fish measures. As discussed in chapter 6, I do this by training measurers, to place the point of their knife on the measuring tape at the end of a fish's tail, to mark the length of the fish. Before removing the fish and revealing the entire measuring tape along with the knife tip exactly on its measurement, and then reading off the measurement to the nearest 1mm. With all the best will in the world, this will still only get the precision to within 2-3 mm but that should be good enough.

In the process of improving data quality, I cannot recommend highly enough working up your data as soon as you start collecting it, and repeatedly as you go along. Show the length frequency histograms to the fish measurers to make them aware of the issues they are creating with their rounding off, and inaccurate measurements. Show them some of the spiky data and make some jokes about the synchronized growth of fish their data implies, or about the fish having favourite numbers.

Find the inaccurate data transcriptions and make some jokes about the 3000 mm fish that should have been 300mm. Show them what good smooth accurate data look like so they can see the difference. The accuracy of your data will improve simply by showing your measurers that you are using their data, and if they see the results of their own work, good and bad. Engage your team of measurers in your scientific process and enable them to understand what the data they are collecting means and how it will be used. It will also build community buy-in for the results and support for reforming management in the longer run.

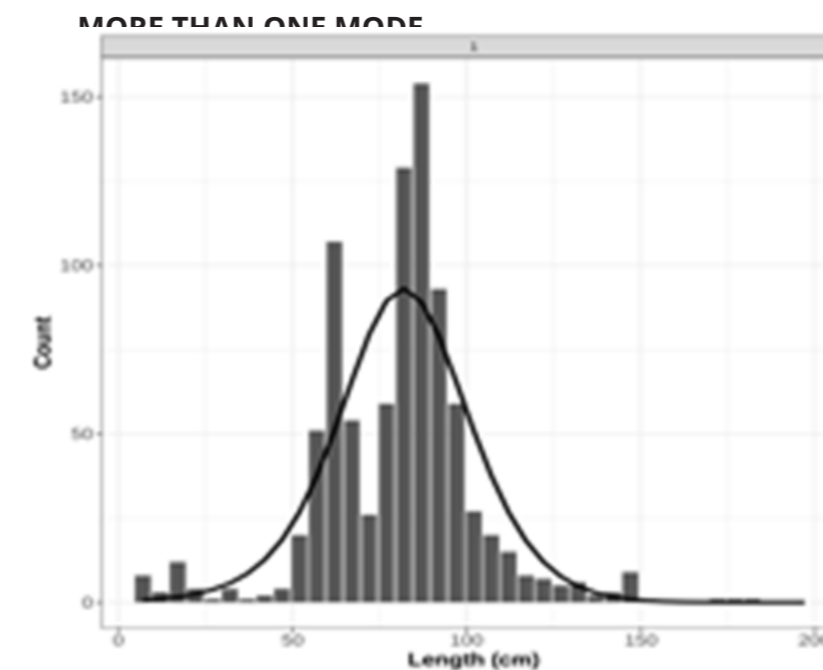


Figure 5. LBSPR assessment model fit (black line) to length composition data (black bars).

In some cases, you will end up with two smooth modes; a juvenile mode on the left, and an adult mode on the right as seen in figure 5. This is not an inaccuracy in your data but a real reflection of the abundance of juveniles in the catch. Nevertheless it will interfere with the LBSPR model's ability to fit the data, as can be seen in figure 5 by the gaps between the fitted curve and the data. As touched upon previously in chapter 6 the problem arises because, like most fisheries models the LBSPR algorithm uses the simplifying assumption that juveniles and adults have the same rate of natural mortality. When in reality juveniles often have much higher mortality rates, meaning that

they can at times be more numerous in your samples than the LBSPR model expects. Using the model's assumption a distinct juvenile mode should not be visible in the data, the data should always be described by a single bell-shaped curve. The implication of this will be that the LBSPR model attempts to fit the multiple modes beneath the one unimodal curve as is seen in figure 4 & 5. The result will inevitably be a very poor fit and inaccurate assessment with wide confidence intervals.

The solution to this issue is relatively straight forward. Fortunately, the spawning potential of a fish stock is created by the adult part of the population in the right-hand mode and not the juveniles in the left-hand mode. The solution is to truncate the data set immediately to the left of the right-hand mode and exclude the smaller juvenile size classes from the analysis. This will give you a much better fit to the data and a more accurate assessment of SPR, than trying to get the model to describe the two modes as one. Note however, that this will upwardly bias your estimates of size selectivity and F/M to some extent, because you have truncated the size composition at the lower end and effectively told the model that the impact of fishing is occurring across a narrower band of size classes. The estimate of SPR is normally, but not always, surprising unaffected by these biases you have introduced.

#### **UNDER-REPRESENTED ADULTS**

The final problem we will discuss here is less easily diagnosed and results from the under-representation of adults in your sample. This has already been discussed in chapters 5 & 6, but it is such an important issue that it warrants being covered here in this context. Most commonly the adult fish may be less vulnerable to being caught, either because they move out of the fishing area, out into deeper water, or can just escape from the fishing gear being used. In such cases the sampled catches may be mainly or almost entirely, juveniles and sub-adults, and contain few if any larger adults. This is what we call dome shaped size selectivity. In this case truncating the data does not help, if we cannot show the model the largest adults in the stock, with its in-built assumption of S-shaped size selectivity the model can only assume the largest size classes of fish do not exist, because they have been fished out, and it will compute a low SPR and high F/M as a result.

It can be very difficult to determine when dome shaped selectivity is occurring, and when it is just the wishful thinking of the community who are having trouble believing all the largest fish have been caught. You really need independent studies of size selectivity in your fishery, strong evidence that there really are large numbers of larger fish surviving in the fish grounds, or good biological evidence that your species moves away from the fishing grounds as they mature. Fishing with a different type of fishing gear, or using underwater stereo video might also be used to get this type of information. You should also be aware that even when the selectivity of the fishing gear is dome shaped, it may not be having any impact on the size composition of the catches, and your SPR estimates may still be accurate. In many cases where fishing pressure is moderate to heavy, fish do not to grow big enough to reach the potential protection provided by the right-hand side of the dome shaped selectivity. In effect only the S-shaped left-hand side of the dome-shaped selectivity has any effect on the stock.

This is obviously a complex issue to deal with, but you should be aware of it, and if you end up with very low SPR estimates and very large estimates of F/M you should ask yourself whether dome shaped selectivity could be a plausible explanation for your results, and what you might do to check that possibility. If you are fortunate enough to have the results of independent studies to specify the dome shaped selectivity, there is now a way of incorporating that into your analysis (Hommik et al. 2020) but these algorithms have not yet been incorporated into our web-based software, because in most of the situations we work, having that sort of information is almost unimaginable.

## References

---

Babcock, E.A., Tewfik, A., Burns-Perez, V., (2018). Fish community and single species indicators provide evidence of unsustainable practices in a multi-gear reef fishery. *Fish. Res.* 208, 70–85.

Erzini K. (1990). Sample size and grouping of data for length-frequency analysis. *Fisheries Research* 9:355–366.

Hommik, K., Fitzgerald, C.J., Kelly, F. & Shephard, S. (2020). Dome-shaped selectivity in LB-SPR: length-based assessment of data-limited inland fish stocks sampled with gillnets. *Fisheries Research*, 229 [online]. [Cited 16 July 2020]. doi.org/10.1016/j.fishres.2020.105574

Hordyk, A. Ono, K., Valencia, S.V., Loneragan, N. & Prince J.D. (2015). A novel length-based estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science*, 72, 217 – 231.

Prince, J.D., Victor, S. Kloulchad, V. Hordyk, A. (2015). Length based SPR assessments of eleven Indo-Pacific coral reef fish populations in Palau. *Fish. Res.* 171: 42-58.

Prince J.D. et al. (2020). Spawning potential surveys in the Western Province of the Solomon Islands. *Pacific Community Newsletter*. 162: 58 – 68.