**A Simple User Guide for LBB** (LBB\_33a.R)

R. Froese, H. Winker, G. Coro, N. Demirel, Athanassios C. Tsikliras, Donna Dimarchopoulou, Giuseppe Scarcella, Wolfgang Nikolaus Probst, Manuel Dureuil, Daniel Pauly

published online at <http://oceanrep.geomar.de/xxxxx/> in July 2019

This is an accompanying document for “Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A., Dimarchopoulou, D., Scarcella, G., Probst, W.N., Dureuil, M., Pauly, D. (2018). “A new approach for estimating stock status from length frequency data” published in the *ICES Journal of Marine Science*. The length-based Bayesian biomass estimator (LBB) is a new method for the analysis of length frequency data, preferably from the commercial fishery. LBB works for species that grow throughout their lives, such as most fish and invertebrates, and requires no input apart from length frequency data. It estimates asymptotic length (Linf), length at first capture (Lc), relative natural mortality (M/K) and relative fishing mortality (F/M) over the age range represented in the length-frequency sample. With these parameters as input, standard fisheries equations are used to estimate depletion or current exploited biomass relative to unexploited biomass (B/B0). In addition, these parameters allow the estimation of the length at first capture that would maximize catch and biomass for the given fishing effort (Lc\_opt), and estimation of a proxy for the relative biomass capable of producing maximum sustainable yields (Bmsy/B0). Relative biomass estimates of LBB were not significantly different from the “true” values in simulated data and similar to independent estimates from full stock assessments. This is an updated version of the LBB code which i) uses median maximum length across available years instead of absolute maximum length as default start value for estimating a prior for Linf, ii) allows users to specify in the ID file whether they want to apply a correction for the pile-up effect which may occur in species with continuous reproduction, and iii) shows an LBB analysis for each year with data and stops execution if LFs in the examined year are unrealistic.

This document provides a simple step-by-step guide for researchers who want to apply LBB to their own data.

The required R-code (LBB\_33a.R or higher) and some example input files (Example\_Dat.csv and Example\_ID.csv) can be downloaded from [http://oceanrep.geomar.de/43182/](http://oceanrep.geomar.de/xxxx/) as Archive (Zip file with LBB User Guide LBB\_33a, R-code and test data).

Installation instructions

1) Install a recent version of R on your computer. LBB was tested under R versions 3.4.4, 3.5.1 and 3.5.2, available from <http://www.r-project.org/>, but newer versions should also work.

2) We suggest using RStudio as an R development environment. RStudio is a free software that is available for several Operating Systems (Windows, OS, Linux, ...) and can be downloaded at <http://www.rstudio.com/products/rstudio/download/> .

3) Install the Gibbs sampler JAGS for your Operating System from the following web site: <http://sourceforge.net/projects/mcmc-jags/files/JAGS/4.x/> .

*If you are using an Operating System different from MS Windows and encounter errors in execution, please comment out (#) all the lines in the code containing the windows(..) function. If you are using an Online RStudio IDE, comment out all X11() commands.*

4) In order to run the code, several R-packages are required. These should be installed and activated automatically when you run the code the first time. If not, check the first lines of the code and install any required libraries manually.

5) Two different data files are required by LBB, i.e. the Stock ID file (ID.File) and a file containing data of “length-frequencies in length bins”, which should be placed in the same directory as the script. The names of these files should be specified as input to the script in the following way: the ID.File should be indicated in the code (line 25), whereas the length bins file should be set as first entry in each line of the Stock ID file. One file can be used to collect all length bins data. Examples are provided with the code and their structure is specified in the next section.

6) Make sure that the source file and the downloaded R script are in the same directory.

7) Open LBB in RStudio. Use the tab “Session” and select “Set Working Directory” -> “To Source File Location”, so that the code will find the data files.

8) If you want to use your own input files, just change the file name of the ID.File in line 25 and the name in the File column of the ID file for the respective stock. If you create your own input files, make sure you use the same headers (case sensitive) as in the provided example files. Make sure you are using comma-delimited (.csv) files (look at the data in a simple text editor such as Notepad to check for consistent use of commas; semi-colons are not accepted).

9) To specify the stock to analyze just enter the unique name or identifier of the stock in row 22 (e.g. Stock <- “DPS\_GSA22”).

10) In RStudio, click on “Source” (or press Ctrl+A followed by Ctrl+R or Ctrl + Shift + S) to execute the code.

11) When the analysis is complete, results can be found in the console window as well as in the LBB graphs window (which can be saved manually).

# Structure of the input files

**Structure of the data file (.csv)**

For each stock, the following information must be specified (in the corresponding columns):

**Length-frequencies in length bins**

**Stock**: a unique fish stock name or identifier (e.g. “DPS\_GSA22”), repeated for each year

**Year**: the reporting year of the length data (e.g. 2004), repeated for each length class.

**Length:** lower border of length class to which CatchNo refers, in millimeters (mm)

**CatchNo**: number of individuals (e.g. 12345 or 0.123) with lengths falling into the length class.

**Structure of ID file (.csv)**

For each stock, the following information must be specified (in the corresponding columns):

**File:** The name of data file in csv format (e.g. “My\_Dat\_1.csv”)

**Stock**: a unique stock name or identifier (corresponding to the one in the ‘Stock’ column in the data file)

The following information is optional (default NA or FALSE)

**Species:** the scientific name of the species, e.g. “*Phycis blennoides*”

**StartYear**: the first year with LF data to be used, e.g. 2000

**EndOfYear**: the last year with LF data to be used, e.g. 2004

**Years.user**: a string of years to be analyzed, e.g. 2000,2002,2004 [Note: no space only comma between years]

**Year.select:** a year for which you want B/B0 with confidence limits to be printed in the console, and which you want to be shown in the graphical output (upper middle panel in Figure 3), e.g. 2008

**Gears.user**: Gears to be analyzed, e.g. “trawl1,LL2,trap3” without space between commas

**Lcut.user:** lower threshold for length data. Data will be restricted to those L >= Lcut.user (in cm)

**Lc.user:** user-specified prior for length at 50% first capture, e.g. 27, in cm

**Lstart.user**: length where gear retention is larger than 95%, e.g. 85, in cm

**Linf.user**: Linfinity or asymptotic length of the von Bertalanffy growth function, e.g. 110, in cm

**MK.user:** user-specified M/K prior, e.g. 2.0

**mm.user:** a Boolean value to specify if analysis is to be done in millimeters; default is FALSE (but Length must be in mm in any case)

**GausSel:** a Boolean value to specify if gill net selection is used; default is FALSE

**MergeLF**: a Boolean value to aggregate LFs with previous year; default is FALSE (if TRUE, first and second year will have identical LFs)

**Pile:** Indicates whether the correction for the pile-up effect should be used; 0 is No (recommended), 1 is Yes, 999 is partly (determined by the fit); the on-screen results will show the Pile factor value in the last row.

**Lm50**: length at which 50% reach maturity, e.g. 28, in cm

**Comment**: a comment on the stock or the quality of the analysis or special settings. This comment is shown in the output.

**Source:** Indicate the source of your length frequency data.

Remember that the files must be saved in “csv” (comma delimited) format. Double-check that indeed a comma (and not a semi-colon) is used as delimiter.

# Results of LBB analysis

The screen output for the deep-water rose shrimp (*Parapenaeus longirostris*) in the Aegean Sea (DPS\_GSA22) is

below:

Lmax = 64 , median Lmax = 40 cm, for potential setting of Linf.user in ID file

Years in data set (for potential cut & paste into Years.user in ID file):

1996,1997,1998,1999,2000,2001,2003,2004,2005,2006,2008,2013,2014

If error without hint occurs, copy years into Years.user and delete next year to be

processed from string

Running Jags model to fit SL and N distributions for Parapenaeus longirostris

----------------------------------------------------------------------

LBB results for ***Parapenaeus longirostris***, stock **DPS\_GSA22**, 1996-2014

Files:Example\_ID.csv, Example\_Dat.csv

-----------------------------------------------------------------------

Linf prior= 40, SD=0.4 cm Lmax=64, median Lmax=40

Z/K prior = 2.4, SD=1.4, M/K prior=1.5, SD=0.15

F/K prior = 0.926 (wide range with tau=4 in log-normal distribution)

Lc prior = 15.3, SD=1.5 cm, alpha prior=13.8, SD=1.4, Lm50=NA cm

General reference points (median across years):

Linf = 39.3 (39-39.9) cm

Lopt = 28 cm, Lopt/Linf=0.71

Lc\_opt = 24 cm, Lc\_opt/Linf=0.6, Lmean if F=M 24.3 cm

M/K = 1.25 (0.971-1.56)

F/M = 1.3 (0.91-2.25), F/K=1.54 (1.24-1.87), Z/K=2.81 (2.65-3.04)

B/B0 = 0.26 (0.13-0.41), B/B0 F=M Lc=Lc\_opt 0.38

Y/R' = 0.051 (0.025-0.076), Y/R' F=M Lc=Lc\_opt 0.059

Estimates for 2014 (mean of last 3 years with data):

Lc50 = 18.9 (18.6-19.3) cm, Lc/Linf=0.48 (0.47-0.49)

Lc95 = 25.7, alpha=0.434 (0.424-0.445)

Lmean/Lopt= 0.86, Lc/Lc\_opt=0.8, L95th=37.7 cm, L95th/Linf=0.96, Mature=NA%

F/M = 1.7 (1.3-2.6), F/K=2.3 (2-2.8), Z/K=3.7 (3.3-3.9)

Y/R' = 0.046 (0.029-0.073)

B/B0 = 0.21 (0.13-0.34), best LF fit year 2006=0.207 (0.14-0.29)

B/Bmsy = 0.57 (0.35-0.91)**, selected B/B0 2008 = 0.23 (0.15-0.32)**

RF: Mostly very good LF patterns! Accepted LBB prior for Linf.prior=40; selected 2008 as year with good fit and reasonable B/B0 compared to adjacent estimates.

**Graphical output**

LBB produces plots of the raw data per year. This is meant to help identify and exclude years that appear unfit for analysis. The years fit for analysis can then be specified in the ID file using StartYear and EndOfYear, or by specifying a string of years in Years.user (see above).

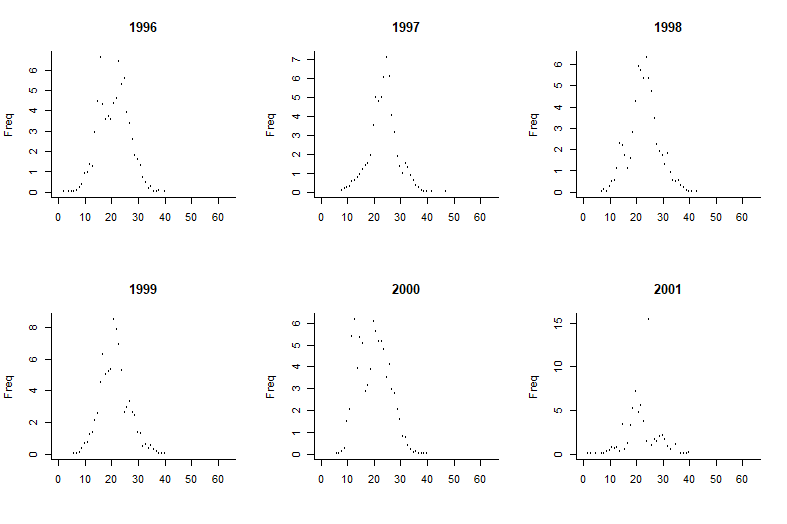


Figure 1. Length frequency data as provided to LBB. Exclude years that do not show the expected pattern. For example, 2001 looks strange with the one high outlier and could be excluded.

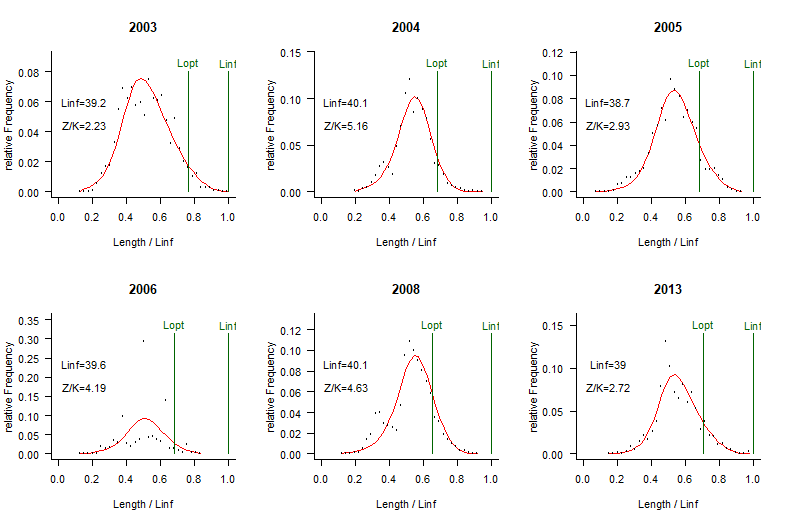


Figure . LBB fits to length frequency data in the provided years. Note that some years have already been excluded. In this example, most fits are good but 2006 looks bad and should be excluded. Very bad fits are indicated when the code is executed and processing is halted. Remove those years as follows: first, copy the string of years shown in the R-console above the results (highlight the text, right-click, select “Copy”). Paste these years into the “Years.user” column in the ID file. Delete years from the string which you do not want to use. The code prints instructions in the console if processing is halted. Sometimes, if the data are very bad, execution stops with an unintelligible error message. In such case, delete the year that was about to be processed and rerun. Note that you can use forward and backward in the menu in the graph window by selecting: History, Previous or Next.

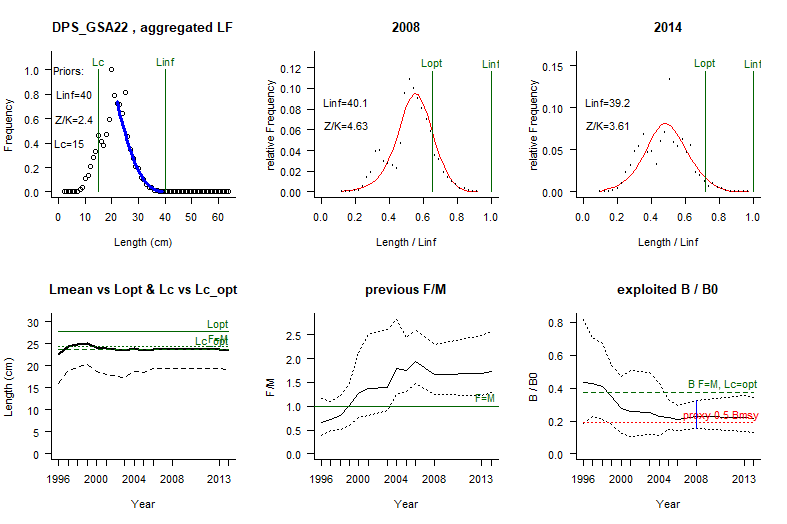


Figure . An example of graphical summary output produced by LBB, here for deep-water rose shrimp (Parapenaeus longirostris) in the Aegean Sea, for the years 1996 to 2014. The upper left panel shows the accumulated LF data used to estimate priors for Lc, Linf and Z/K. The upper middle and right panels show the LF data for the first (or selected) and last year in the time series. The red curve shows the fit of the LBB master equation, which provides estimates of Z/K, M/K, F/K, Lc, and Linf. From Linf and M/K, Lopt is calculated and shown as reference. The lower left panel shows Lmean (bold black curve) relative to Lopt, and Lc (dashed black curve) relative to Lc\_opt. The lower middle panel shows relative fishing pressure F/M (black curve), with approximate 95% confidence limits (dotted curves), with indication of the reference level where F = M (green horizontal line). The lower right panel shows relative biomass B/B0 (black curve) with approximate 95% confidence limits (dotted black curves), with indication of a proxy for Bmsy (green dashed line) and a proxy for Bpa or 0.5 Bmsy (red dotted line). The blue vertical line indicates the confidence limits of the selected year, here 2008.

**Good LBB practice**

This is a guide on how to best use LBB:

1. Install everything as described above and run the code with the example data. Store code and example data in a safe place, so you can go back to them in case you run into errors that you cannot resolve.
2. If the example data give the results shown in this document, you are ready to go. Save the code under a new name (e.g. LBB\_33a\_MyCode.R), so you can make changes to it without messing up the original.
3. Open the Example\_Dat.csv file, add your own data at the end following the examples given, and save the file e.g. as “My\_Dat.csv”. Be careful if you use Excel, because depending on your regional settings, Excel may replace comma with semi-colon. You can either change the regional settings to force Excel to use a point as decimal indicator and a comma as column separator, or use instead the free Star-Office Calc software.
4. Open the Example\_ID.csv file. Start a new entry below the last occupied row. In column A: “File”, replace “Example\_Dat.csv” with your new file name. Fill in the other columns following the examples given in the existing rows and the explanations given above. When done, save the ID file under a new name, such as “My\_ID.csv”.
5. In the LBB R-code in row 25, replace “Example\_ID.csv” with your new file name. Run the code.
6. If everything was done correctly, the console should indicate a first printout with the maximum length Lmax in your data and the median Lmax value across all years with data. When you are done with the analysis and the estimated Linf between years varies more than +/- 10%, then you may use the median Lmax – Lmax range to set the Linf.user prior in the ID file, somewhere in that range, depending how good you think the largest fish are represented in your sample: if good choose median Lmax, if bad chose Lmax, if you are uncertain, use the mean of the range. If there is not much variance in Linf between years (Figure 2), keep the default prior (do nothing).
7. Look at the length-frequency plots of the raw data (Figure 1). If there are only few points per year, set MergeLF in the ID file to TRUE. This will combine length-frequencies of subsequent years before the LBB analysis is done.
8. Look again at the length-frequency plots of the raw data (Figure 1). If in some years the plots do not follow the expected pattern, exclude that year. This is best done by copying the string of years from the output in the console and paste it in the Years.user column in the ID file. Then delete the ‘bad’ years from that string.
9. After you have excluded all ‘bad’ years, rerun the code. Now LBB will try again to fit the master equation to the remaining years (Figure 2). Execution will stop if the fit is deemed unrealistic. In that case, exclude the year as described above. Note that you can move backward in the window with the annual LBB fits by selecting ‘History’ in the menu of the window.
10. When all length frequency plots are satisfactory and successfully fitted by LBB, the results of the analysis are printed in the console and a window with summary graphs is displayed (Figure 3). If there is only one year with length frequency data, only the upper panel of Figure 3 is displayed, showing that year twice.
11. The LBB results can be used either directly for management (see main LBB paper for guidance) or as providing an objective biomass prior B/B0 for use with other models, such as CMSY or AMSY. For the latter purpose, select a year for which you want to estimate the B/B0 prior and indicate it in the Year.select column in the ID file. Typically, you want to select a year that is neither the lowest nor the highest but rather a year that has a good length frequency fit and gives B/B0 results similar to adjacent years.
12. You are done! For your report, you can copy and paste the text from the console (highlight, right-click, Copy) into a Word document. You can also copy and paste there the graphical output such as shown in Figure 3.

If you have any questions, contact us:

Rainer Froese [rfroese@geomar.de](mailto:rfroese@geomar.de)

Gianpaolo Coro [gianpaolo.coro@isti.cnr.it](mailto:gianpaolo.coro@isti.cnr.it)

Henning Winker [henning.winker@gmail.com](mailto:henning.winker@gmail.com)

Nazlı DEMİREL [nazdemirel@gmail.com](mailto:nazdemirel@gmail.com)

Enjoy using LBB.