COMPUTATIONAL OPTIONS FOR MARINE FISHERIES RESEARCH AND MANAGEMENT

J. Jayasankar
Fisheries Resource Assessment Division
ICAR- Central Marine Fisheries Research Institute

13

Introduction

Marine fisheries research and management including its core domains like stock assessment and resource simulation and forecasting has of late been driven by the explosion in the computational power and designing and development of tailor-made software. Though the activities on the computational front have peaked in the last five years or so, the seeds were sown way back in eighties with the introduction of personal computing. Here it is relevant to record that digitization of fisheries related information and datasets were much older than these software developments. Although it is nearly impossible to prepare an exhaustive list of software and routines which are presently in use by researchers and planners in Marine Fisheries Management, it is possible to categorise the computational options on a perspective note. This paper would attempt doing that.

Generic grouping of computational tools

For enhancing objectivity, the software options under focus can be grouped into the following four categories:

(i) Custom made software for specific fishery related issues

(ii) General purpose software routines which are of high relevance to fisheries research

(iii) Software and digital options for information and data processing in fisheries research and

(iv) Miscellaneous options

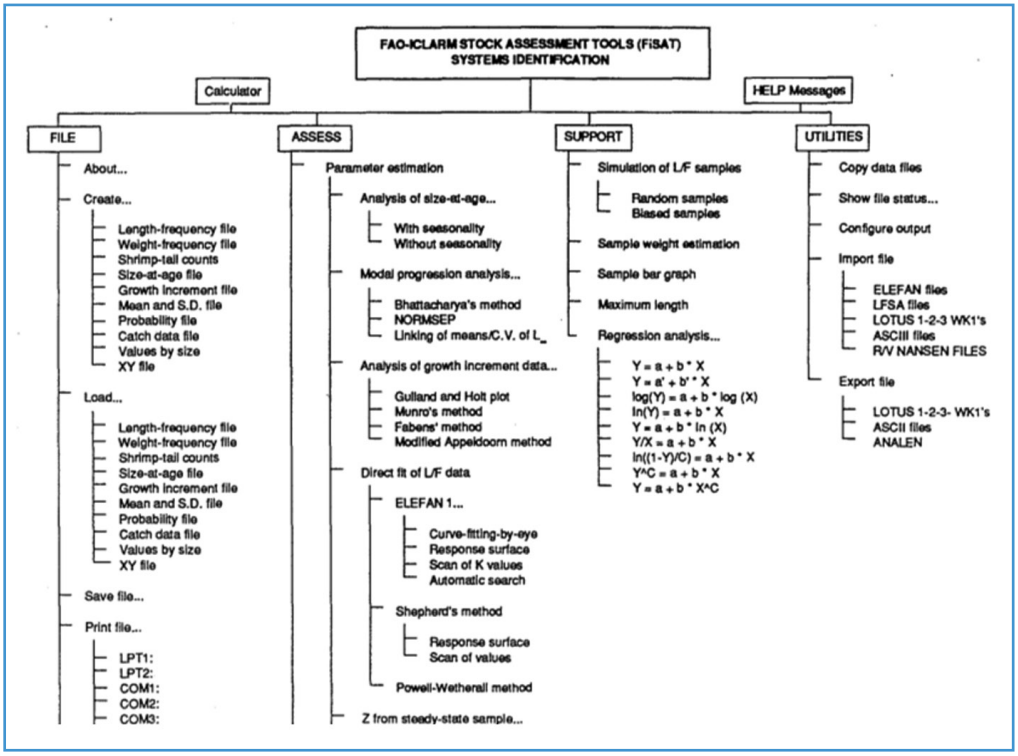
Custom Made Software

This group is the most important and diversified one amongst all computations tools available. These include very specific tools like Electronic Length Frequency Analysis (ELEFAN) and the routines of similar nature used in quantitative fish stock assessment which have later been enshrined in the FAO- ICLARM Stock Assessment Tools (FiSAT).

FiSAT

A beautiful introduction to the thought process that preceded the creation of this software has been given in Pauly and Sparre (1991). The software was the major first such effort in the field of fish stock assessment that brought the various possibilities arising out of established conceptualisations that could throw light on the quantification of the growth and reproductive performance of an average fish of a species and stock under one roof with a common data initialization. The spreadsheet based data feeding was fully focused upon and the tools were further grouped under a bouquet titled ASSESS wherein VBGF type growth, LCC based mortality, spawning stock- recruitment to yield per recruit to Thompson and Bell models could be applied on suitable datasets which can be fed in half a dozen formats. The major standout feature of this software is the ease of performance of modal progression analysis to separate out the cohorts from a mixed bag of length-weight data using the linearized differentiation of normal densities. This process again has been simplifies almost to the spreadsheet level with added advantage of letting the researcher to select and omit the possible candidate data points.

The arrangement of options and tools under the software is best depicted by the tree chart (partially reproduced here) by Pauly and Sparre (1991).



This probably is the most deployed software for any fish population dynamic analytics and till date serves as the benchmark for other new age software. The ease of use of this software partially stems from the non-stochastic approach to the analysis and thereby ensuring replication of results on even datasets. This deterministic approach is the one which has lead to myriad of more comprehensive and broad based analytical framework for the data sampled.

Marine Resources Assessment Group (MRAG) Initiative Under Fishery Management Science Programme (FMSP)

LFDA Version 5.0

The Length Frequency Distribution Analysis (LFDA) package is a PC-based computer package for estimating growth parameters and mortality rates from fish length frequency distributions. Version 5.0 of LFDA includes methods for estimating the parameters of both non-seasonal and seasonal versions of the von Bertalanffy growth curve. It includes three methods for estimating growth parameters. These are Shepherd's Length Composition Analysis (SLCA) method, the projection matrix (PROJMAT) method, and a version of the Elefan method. A facility is provided that allows conversion of length frequencies to age frequencies using the estimated growth curves. In addition to methods for estimating growth parameters, the package also includes three methods for estimating the total mortality rate Z, given estimates of the von Bertalanffy parameters. A function allowing simulation of length frequency data under a variety of assumptions is also included. As with previous versions of LFDA, the package includes a comprehensive context-sensitive Help system and a detailed example analysis. The download file also includes the graphics server programme required to plot the data.

CEDA Version 3.0

The Catch Effort Data Analysis package (CEDA) is a PC-based software package for analysing catch, effort and abundance index data. Version 3.0 allows calculation of estimates of current and unexploited stock sizes, catchability and associated population dynamics parameters. Both depletion and several types of stock production (biomass dynamic) models can be fitted, using one of three different assumptions about the distribution of residuals. Both point estimates and bootstrap confidence intervals for the estimated parameters can be calculated. CEDA also includes the facility to do projections of stock size into the future under various scenarios of catch or effort levels, so that different management strategies can be investigated. Output is presented both graphically and textually, and can be printed or saved to disk for further use. As with previous versions of CEDA, the package includes a comprehensive context-sensitive Help system and a detailed example analysis. The download file also includes the graphics server programme required to plot the data.

Yield Version 1.0

Yield Version 1.0 is a program for calculating fishery yields and stock biomasses, on an absolute or per-recruit basis, and for calculating biological reference points associated with these. On starting the program, users are asked to enter values of biological parameters (e.g. growth, mortality, age at maturity and stock-recruitment relationship) and fishery parameters (e.g. length at first capture, fishing season). For each parameter, either a single value can be entered, or a probability distribution can be specified to allow for uncertainty. When calculating yields and yields per recruit, the program takes explicit account of specified parameter uncertainties, presenting results in terms of histograms. Transient projection and reference point calculations can also be made, once the extent of stochastic recruitment variability has been specified. As with CEDA and LFDA, the package includes a comprehensive context-sensitive Help system and a detailed example analysis. The download file also includes the graphics server programme required to plot the data.

ParFish Version 2.0

Participatory Fisheries Stock Assessment (ParFish) Software is a PC-based software package that uses Bayesian Statistics and Decision Theory to assess the state of a fishery stock and estimate limit and target control levels. The software supports the overall approach which is described in the accompanying ParFish Guidelines. The guidelines provide an overview of all six steps in the approach including: i) understanding the context; ii) engaging stakeholders; iii) undertaking the stock assessment; iv) interpreting the results and giving feedback; v) initiating management planning and vi) evaluating the process. The ParFish software is currently based on the logistical biomass growth model and requires information on four parameters: Current Biomass, Unexploited Biomass, Catchabilty and Growth rate. Interview data from fishers are used to construct 'priors' for the model parameters which can be combined with other available information to provide best estimates. This information is then used, together with preference data from fishermen, to calculate the current stock level and the control levels that will provide the most preferred catch rates for fishers. The programme takes explicit account of uncertainty in the data, presenting results as probability density functions (with accompanying mean, median, mode and confidence intervals). The Software is accompanied by a manual which gives step-by-step guidance on inputting data and running the analysis. There area also additional reference sheets which assist with the interviews and other data collection methods.

RAPFISH- A Rapid Analysis Tool for Fishery Sustainability

This is an unique but important software tool, which could be tagged under research as well as management of fisheries, wherein Multi Dimensional Scaling, a multivariate statistical dimension reduction tool, has been put to use to rank fisheries simultaneously on biological,

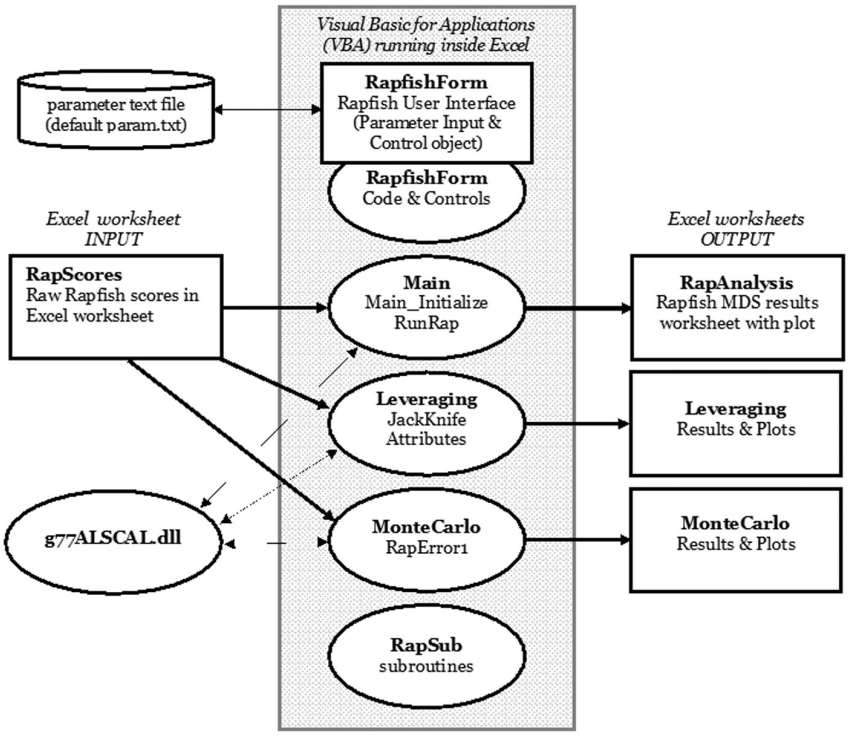
technological, economic, ethical and sociological fronts by ranking different fisheries under various contributory aspects falling under these five dimensions. Visual Basic for Applications (VBA) codes have been developed which would guide through the data entered through Excel spreadsheet to the development of report on the status of sustainability of the fisheries under focus and their combined unidirectional ranking. The framework of the software has been described by Patricia, K and Tony J. Pitcher (2004) as follows:

The original rationale for developing for *Rapfish* was toevaluate sustainability, and examples from that modality are largely used in this document. Fisheries scientists grade fisheries according to a large set of 'attributes'. Attributes are grouped in ecological, economic, social, technological, and ethical categories, or 'evaluation fields', so that 'sustainability' can be considered from various points of view. The *Rapfish* technique is flexible such that other modalities of status may be used, such conformity with a set of specified objectives or compliance with a code of conduct . *Rapfish* applies a statistical ordination technique called Multi-Dimensional Scaling (MDS) to reduce the NxM matrix of fisheries statistics for N fisheries and M attributes into a N x 2 dimensional space which has similar distance properties as the N x M statistics. In this 2D attribute space, one dimension (x-axis) is the score representing the status (degree of sustainability) from 'bad' to 'good', and the other dimension (y-axis) represents other factors, unrelated to sustainability (or whatever status is being scored), which distinguish fisheries. The MDS routine ALSCAL in the statistical package SPSS was used in the development and testing of the *Rapfish* technique. SPSS batch programming facilities software were written (Kavanagh 1999) to automate the *Rapfish* procedure, including routines for attribute leveraging and Monte Carlo error analysis. Problems with this software were inflexibility and awkwardness in re-configuring parameters due to limitations in the SPSS command language.

This report describes a more portable and easy-to-use *Rapfish* software implementation, implemented in Microsoft Excel and its programming language, Visual Basic for Applications (VBA). Excel is a popular and low-cost application and the majority of fisheries scientists are familiar and comfortable using it for statistical data analysis. The original ALSCAL FORTRAN code for multi-dimensional scaling has been re-written and built as a dynamic link library routine (DLL) called from an Excel/VBA program. This Excel/VBA/FORTRAN implementation of *Rapfish* is portable (users need only Excel and not SPSS), is easy to programme for various repeat analyses such as leveraging and Monte Carlo, and has a handy graphical user interface to control processing and visualize results.

The Rapfish software architecture can be diagrammatically explained as follows:

General purpose software routines which are of high relevance to fisheries research. Under this category fall the analysis environments like R, WinBUGS, GAMS etc. which have



many customization options of very direct and high utility value in the area of fisheries management. Of these the open source R software stands out in dishing out half a dozen routines referred to as libraries which have been.

TropFishR

The TropFishR package uniquely adds further data-limited method capacity (Table 1) by including tradi- tional and updated versions of the Electronical LEngth Fre- quency ANalysis (ELEFAN) method, used in growth parameter estimation, with new optimisation techniques (Tay- lor & Mildenberger 2017), Millar's nonlinear selectivity mod- els (Millar & Holst 1997), and a complete set of methods for fisheries analysis with LFQ data. This compilation allows a stock assessment routine to derive reference levels (e.g. FMSY, F0 1) by means of yield per recruit modelling, which may be based on a single year of LFQ data. Until now the preferred software for single species stock assessment with length-fre- quency data has been the windows-based programme FiSAT II (Gayanilo, Sparre & Pauly 1996) due to its user-friendly, click-based interface. The software is, however, limited in its ability to import data and perform automated analyses. The TropFishR package aims to remedy these shortcomings by allowing further expansion and flexibility. Although wider in scope, the main methods follow

those outlined in the FAO manual 'Introduction to tropical fish stock assessment' (Sparre & Venema 1998). Many of the same examples and datasets featured therein are included in the package (Table 1) and documented in accompanying help files, which facilitates use in training and teaching. Finally, output from various functions can be passed to plotting functions, allowing for export as publication-quality figures.

For historical reasons, and the link to the above-mentioned book by Sparre & Venema (1998), the package's name reflects the fact that the methods have often been applied to tropical fisheries, although they are equally applicable to other regions with data-poor stocks for which LFQ data is available. Typically, the workflow of a data-poor stock assessment with LFQ data would include: (i) estimation of biological stock characteristics (growth and natural mortal- ity), (ii) fisheries performance aspects (exploitation rate and selectivity), and (iii) stock size and status. The order of the methods is important as they build upon each other in a sequential way. If some or all of the vital parameters for stock assessment are already known, the user may skip the data-poor approaches for their assessment and can directly proceed with yield modelling applications.

CatDyn: Fishery Stock Assessment by Generalised Depletion Models

As a recourse to viewing the stock dynamics through catch rather than the population, which is of course used as an index for the latter, routines have been developed to assess, model and predict stock health using Generalised Depletion models. The entire gamut of parametrisation, modelling and forecasting has been made handy by the R library CatDyn. As per the introduction given by the author(s) of CatDyn, the library is capable of the following:

Based on fishery Catch Dynamics instead of fish Population Dynamics (hence CatDyn) and using high-frequency or medium-frequency catch in biomass or numbers, fishing nominal effort, and mean fish body weight by time step, from one or two fishing fleets, estimate stock abundance, natural mortality rate, and fishing operational parameters. It includes methods for data organization, plotting standard exploratory and analytical plots, predictions, for 77 types of models of increasing complexity, and 56 likelihood models for the data.

The concept of depletion modelling is set into motion using the following parametrization. The process equations in the Catch Dynamics Models in this package are of the form

$$C_t = ke^{-\frac{M}{2}}E_t^a N_t^b$$

$$N_t = N_0 e^{-Mt} - e^{\frac{M}{2}} \sum_{i<t} C_{t-1} e^{-M(t-i-1)} + \sum_j P_j e^{-M(t-j)}$$

where C is catch in numbers, t, i are time step indicators, j is perturbation index (j=1,2,...,100), k is a scaling constant, E is nominal fishing effort, an observed predictor of

catch, *a* is a parameter of effort synergy or saturability, *N* is abundance, a latent predictor of catch, *b* is a parameter of hyperstability or hyperdepletion, and *M* is natural mortality rate per time step. The second summand of the expanded latent predictor is a discount applied to the earlier catches in order to avoid an *M*-biased estimate of initial abundance. Perturbations to depletion represent fish migrations into the fishing grounds or expansions of the fishing grounds by the fleet(s) resulting in point pulses of abundance. In transit models (limited to one fleet) there are also emigration events happening at specific time steps for each perturbation. In 2 fleet cases the fleets contribute complementary information about stock abundance, and thus operate additively; any interaction between the fleets is latent and affects the estimated values of fleet dependent parameters, such as *k*, *a*, and *b*.

The observation model can take any of the following forms: a Poisson counts process or a negative binomial counts process for catch recorded in numbers, an additive random normal term added to the continuous catch (in weight) predicted by the process (normal and adjusted profile normal), a multiplicative exponential term acting on the process-predicted catch such as the logarithm of this multiplier distributes normally (lognormal and adjusted profile lognormal), and Gamma (shape and scale parameterization).

The library CatDyn takes care of almost all the parameterisation issues and dishes out the type of output which would magnify the status of fisheries as seen from the macro dynamic level in such a way to aid the policy makers.

Other R Libraries

There are a few more libraries in R viz. FSA, Fishery Libraries in R (FLR), fishMod etc. which have specific routines or functions that could be applied under one type of assessment protocol. Amongst these FLR seems to be a multifaceted effort wherein almost all aspects of fishery including bycatches, discards as well as economics of fishery are being simulated/ analysed. Most of these libraries have explicit or derived leads to the arriving at of crucial Biological Reference Points (BRP), like MSY, F0.1 etc. which would help the fishery manager to take a call on the precautionary or knife edge type calls on effort moderation so that the stock health is saved.

Software and Digital Options for Information and Data Processing in Fisheries Research

This third type of software are basically data driven and hence could be basically front-ends of huge data repositories. Let us have a look at a few of them:

FAO's FishStatJ

The FishStatJ application provides users with access to a variety of fishery statistical datasets. Any data having yearly time series and coded dimensions can potentially be stored and processed by **FishStatJ**. The system consists of a main application module and workspaces

which include the datasets and can be loaded by the user.FishStatJ is a Java-based desktop application. This is quite helpful in getting data on global capture fisheries and aquaculture fish production on aggregated and disaggregated at various levels of granularity.

R language- Rfishbase

This is a library provided interface to various types of customised data tapping from the repositories of fish biology related information on the lines of Fishbase. Despite functioning as the programmatic interface to Fishbase <http://www.fishbase.org>, re-written based on an accompanying 'RESTful' API. Access tables describing over 30,000 species of fish, their biology, ecology, morphology, and more. This package also supports experimental access to <http://www.sealifebase.org> data, which contains nearly 200,000 species records for all types of aquatic life not covered by 'FishBase.'

R language- rfisheries

This yet another database interface in R which gives updated landings of various countries. It is a programmatic interface to 'openfisheries.org'. This package is part of the 'rOpenSci' suite (http://ropensci.org).

Apart from these there are many more open access digital repositories like ICOADS, APDRC, SeaWIFS (NOAA), which dish out various bio- geo- chemical and oceanographic datasets on spatio-temporal tagging base which are of high relevance in climate based modelling of Fishery dynamics.

Miscellaneous

Apart from these types of end to end solution providing software or coding platforms, there are quite a few general purpose routines which have immense use in the field of analysing fishery information. A few of them are listed below:

(i) WinBUGS: A Windows based Bayesian analysis tool using Gibbs Sampler which supports Markov Chain Monte Carlo (MCMC) algorithms is a great tool, wherein fishery growth models can be analysed with additional information on the trends observed in the estimated trajectories of important parameters like virgin biomass, carrying capacity, intrinsic rate of growth, which were otherwise considered as constant unique valued functionals in frequentist concept.

(ii) Generalised Algebraic Modelling System (GAMS) / Data Envelopment Analysis (DEA) for performing optimisation of bioeconomic models on the lines of constraint shored minimisation/ maximisation goal followed in Linear Programming.

(iii) Routines and libraries which could carry out Automatic Differentiation Model Building (ADMB), which by far has been recorded as the most suited one for

complicated fishery optimisation issues using which crucial pre- BRP parameters could be estimated with more precision.

(iv) There are quite a few comprehensive packages which could analyse marine communities on multivariate biotic and abiotic variables over a series of sampling points, thereby comparing the diversity gradient and the spacio-temporal innuendos thereof. The most prominent one is Plymouth Routines in Multivariate Ecological Research (PRIMER), Clark and Warwick (2001).

(v) Other advancements like the analysis of slope of size spectra, which has a very high level of application in assessing marine ecosystem especially under the realms of the average trophic levels is another addition to the analysis tools basket. Size spectrum models have emerged from 40 years of basic research on how body size determines individual physiology and structures marine communities. They are based on commonly accepted assumptions and have a low parameter set, making them easy to deploy for strategic ecosystem-oriented impact assessment of fisheries. They are rooted on the fundamental principle of food encounter and the bioenergetics budget of individuals. Within the general framework, three model types have emerged that differ in their degree of complexity: the food-web, the trait-based, and the community models. The implementations of size spectrum models on these lines flag important variations concerning the functional response, whether growth is food-dependent or fixed, and the density dependence imposed on the system.

(vi) Another booming area of research armoured by computational power is Stock Synthesis. Stock Synthesis (SS) provides a statistical framework for calibration of a population dynamics model using a diversity of fishery and survey data. It is designed to accommodate both age and size structure in the population and with multiple stock sub-areas. Selectivity can be cast as age specific only, size-specific in the observations only, or size-specific with the ability to capture the major effect of size-specific survivorship. The overall model contains subcomponents which simulate the population dynamics of the stock and fisheries, derive the expected values for the various observed data, and quantify the magnitude of difference between observed and expected data. Some SS features include ageing error, growth estimation, spawner-recruitment relationship, movement between areas. SS is most flexible in its ability to utilize a wide diversity of age, size, and aggregate data from fisheries and surveys. The ADMB C++ software in which SS is written searches for the set of parameter values that maximize the goodness-of-fit, then calculates the variance of these parameters using

inverse Hessian and MCMC methods. A management layer is also included in the model allowing uncertainty in estimated parameters to be propagated to the management quantities, thus facilitating a description of the risk of various possible management scenarios, including forecasts of possible annual catch limits. The structure of Stock Synthesis allows for building of simple to complex models depending upon the data available.

Conclusion

The options thrown up by the giant strides made by computational advance are immense in recent times and so are the opportunities and new challenges posed by unravelling the more and more complicated facets of intricate dynamics of oceanic flora and fauna. Though any one software or method can be singled out as THE SOLUTION, the adoption of more than one for same set of data would give an idea about the sensitivity/ robustness of the inferences, thereby making forecasting a more reassuring assignment.

Suggested Reading

Beddington, J. 2005. Capacity building in the use of FMSP stock assessment tools and management guidelines. DFID Fisheries Management Science Programme, Final Technical Report, project R8468. MRAG Ltd. 73 pp.

Beddington, J., 2005. Capacity building in the use of FMSP stock assessment tools and management guidelines. DFID Fisheries Management Science Programme, Final Technical Report, project R8468. MRAG Ltd. 73 pp.

Clarke, K.R., Warwick, R.M. 2001. Change in marine communities: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E, Plymouth, 172pp.

Hindson, J., Hoggarth, D.D., Krishna, M., Mees, C.C., and O'neill, C., 2005. How to Manage a Fishery - A Simple Guide to Writing a Fishery Management Plan. Centre for Environmental Education, Ahmedabad, Gujarat, India. 86 pp.

Hoggarth, D.D., Abeyasekera, S., Arthur, R.I., Beddington, J.R., Burn, R.W., Halls, A.S., Kirkwood, G.P., Mcallister, M., Medley, P., Mees, C.C., Parkes, G.B., Pilling, G.M., Wakeford, R.C., Welcomme, R.L. 2006. Stock assessment for fishery management: A framework guide to the stock assessment tools of the Fisheries Management Science Programme (FMSP) FAO Fisheries Technical Paper. No. 487. Rome, FAO. 261 pp. [Includes a CD-ROM].

Hoggarth, D.D., Mees, C.C., O'neill, C., Hindson, J., and Krishna, M. 2005. A Guide to Fisheries Stock Assessment using the FMSP Tools. Centre for Environmental Education, Ahmedabad, Gujarat, India. 96 pp.

Kavanagh, P. and Pitcher, T.J., 2004. Implementing microsoft excel software for rapfish:a technique for the rapid appraisal of fisheries status . Fisheries Centre Research Reports Vol 12 (2): 75 p.

Mildenberger, Tobias, K., Taylor, M.H. and Wolff, M., 2016. "TropFishR: An R package for fisheries analysis with length-frequency data." *Methods in Ecology and Evolution*. https://doi.org/10.6084/m9.figshare.4212975.v1.

Pauly, D and Sparre, P., 1991. A Note on the Development of a new software package, the FAO-ICLARM Stock Assessment Tools, Fishbyte 9(1): 47-49

Sparre, P., and Venema, S.C., 1998. "Introduction to Tropical Fish Stock Assessment." *FAO Technical Paper*.