Project no. **502572**

FISBOAT

FISHERIES INDEPENDENT SURVEY-BASED OPERATIONAL ASSESSMENT TOOLS

Instrument : STREP

Thematic Priority : 8.1

FINAL ACTIVITY REPORT

Period covered: from 01 March 2004 to 30 June 2007

Start date of project: 01 March 2004 Duration: 40 months

Project coordinator name : Pierre Petitgas
Project coordinator organisation name : IFREMER

Revision : draft 1

FISBOAT final activity report

Publishable Executive Summary ……………………………………………………. 2
Overview of Project Results …………………………………………………………. 3
Plan for the Use and Dissemination of Knowledge ……………………………. 6
Document 1: Indicator-Based Assessment – Methods ………………………. 11
Document 2: Indicator-Based Assessment – Application of methods ……. 148
Document 3: Survey data only Assessment models …………………………. 181
Document 4: Simulation Evaluation with FLR - Methods ……………………. 211
Document 5: Simualtion Evaluation with FLR – applications ………………. 228
Document 6: Simulation Evaluation with ALADYM – Methods ……………... 244
Document 7: Simulation Evaluation with ALADYM – Applications ………… 265
Document 8: Comprehensive Assessments ……………………………………. 272

Individual case study reports are annexed to this report in a separate volume

Publishable executive summary

The Fisboat project was aimed at developping fish stock assessment tools that are based on fishery independent research survey data only and evaluate how these tools perform in providing diagnostics and advice in different management contexts. The survey-based assessment included indices of demography, total mortality, spatial occupation, biological traits leading to comprehensive stock diagnostics. The project involved several disciplines: population biology, survey methods, stock assessment, management. The project case studies spanned a diversity of European stocks and regional seas : Barents sea cod, North Sea cod and herring, Baltic sea cod, Bay of Biscay hake and anchovy, Thyrrenean sea red mullet, Ionian sea hake, Aegean sea hake.

The project has developed fishery-independent survey-based methods and tools to assess on fish stocks. The project has developed the capacity to calculate fish populations' indices of abundance, vital traits and spatial distribution, monitor changes in their time series and formulate comprehensive indicator-based diagnostics. The successful application of methods and tools to all project case studies proved the feasability of the procedures and the operationality of the tools in providing fishery-independent survey-based assessment and advice. Methods and applications were compiled as pubished manuals (ICES CM 2007/O:27 and O:16). The Fisboat indicator-based procedures suggest a way to achieve an operational and comprehensive monitoring system of fish stocks with an ecosystem perspective.

The project has also developed survey-data-only assessment models which span a diverse range of data requirements, from aggregated biomass to length-structured and age-structured models. These models allow for the estimation of abundance, catchability and mortality indices. Models preformances were bench-mark tested using simulated test data sets with known characteristics. A manual of methods compiling models documentations and performances was produced (ICES CM 2007/O:04).

Because survey-based assessment procedures used indices and indicators, simulation evaluating their performance was a natural complement. The project developed under the FLR framework a simulation evaluation loop comprising an age-structured population model, a survey-like observation model, a harvest model as well as graphical and statistical outputs summarising simulation results. The tools and their case study applications were documented in specific manuals. The simulation experiments that were run on the case studies allowed to investiage key issues including which are the harvest rules that are robust to uncertainties in the population dynamics as well as in the precision of survey indices. The FLR simulation plateform was appropriate for the current TAC-based management context within ICES waters. Another simulation plateform (ALADYM) was also developed. It used a more biologically complex population model, which was useful in other management situations, e.g. Mediterranean waters, where fishery landings are not controled and where so called 'technical measures' are envisaged as management options. The ALADYM simulation plateform allowed to investigate combinations of fish stock biological traits with management measures on the long-term sustainability of the population. Methods, tools and results of applications to case studies were reported in documents produced as manuals.

In all, the project developed operational tools and applied these on case study applications with success, thus demonstrating the possibility to monitor fish stocks using fishery-independent survey-based procedures and provide advice in different management contexts. The comprehensive indicator-based diagnostics combined with simulation evaluation tools had the potential to increase the reliability of the diagnostics and advices. Ways on how to create comprehensive assessments have been reported in a document cross-cutting all project aspects. All project products are available on the Fisboat website at http://www.ifremer.fr/drvecohal/fisboat/.

Overview of project results

Conventional fish stock assessment methods are dependent on commercial catch data. But there are more and more situations in which alternative assessment methods are needed. For instance, for stocks that are collapsed, the fishery might be closed thus fishery catches are unavailable or unreliable. Also, misreporting and unaccounted discards generate difficulties to convert reliably landings to effective catches at sea. Last, advice is asked for an increasing number of stocks, and for many of them historical series of fishery catches disaggregated by age do not exist, making conventional assessments impossible. In that context, research survey-based measurements made at sea in known conditions represent an invaluable set of fishery-independent data on which to base an assessment. What type of assessment do these data lead to and how could such assessment be useful alongside existing methods?

To investigate these questions, the FISBOAT project involved several disciplines: population biology, survey methods, stock assessment, management. The project case studies spanned a diversity of European stocks and management contexts: cod in the Barents sea, the North Sea, the Baltic sea, hake in the Bay of Biscay, the Ionian sea, the Aegean sea, herring in the North Sea, anchovy in the Bay of Biscay and red mullet in the Tyrrhenian sea.

The FISBOAT project developed methods and tools (software and documentation) for assessing the status of fish stocks as well as provide advice on management strategies, using only fishery-independent information from research surveys. Three categories of methods were developed: (i) assessment models, (ii) monitoring procedures based on indicators of stock attributes, (iii) simulation evaluation tools. The two first category of methods are essentially diagnostic methods that provided relative diagnostics. Simulation tools allowed to investigate the effect of management options and were complementary to the diagnostic tools. The developed procedures were applied on case study applications with success, which demonstrated the possibility to monitor fish stocks using fishery-independent survey-based procedures and provide advice in different management contexts. The methods and their application tools were documented and developed as scripts in language R.

The present final activity report compiles synthesis documents which describe the methodological developments achieved, their applications to case study stocks and how a comprehensive approach to stock assessment and advice could be set up. Individual case study reports on the application of the methods are annexed to this report in a separate volume. Most of the Fisboat documents have been disseminated at ICES ASC or expert groups and at GFCM. The present report has been concieved to serve as a manual of methods and applications to support the development of an alternative approach to fish stock assessment using survey data. The Plan on the Use and Dissemination of Knowledge documents that purpose. Further, all project products are available on the Fisboat website at http://www.ifremer.fr/drvecohal/fisboat/.

Indicator-based assessments are treated in Documents 1 and 2. Survey-based assessment models are the subject of document 3. Documents 4 and 5 are concerned with the FLR simulation evaluation plateform and its applications. Documents 6 and 7 treat of the ALADYM simulation tools and their applications. Finally Document 8 makes suggestions for a comprehensive assessment approach.

Document 1 compiles the methods for an indicator-based assessment. First, methods were developed to calculate survey derived indicators of biological and spatial attributes of fish stocks. An extensive list of indicators was considered for monitoring abundance, vital traits and spatial distributions. The indicators were documented using a standardized format. Second, a variety of methods were considered to statistically evidence trend or change in the indicator time series. These methods consituted a monitoring approach to fish stock assessment with set risks of detection and false alarm. The assessment used the wide range of biological information and not just the abundance at age. In addition to using a collection of univariate indicator time series, methods to construct multivariate indicators were considered and documented. Multivariate methods also served to select those indicators which carried the signal of change in the time series. Last, the indicator-based assessment was achieved by combining the results of the analyses of the indicator time series in specific diagnostic tables. The construction and interpretation of these tables was also documented.

Document 2 compiles the application of the indicator methods to all the project case sudies. Document 2 is based on individual case study reports that reported applications with a standard template. The template is concieved to document how changes in indicators time series are identified and how diagnostics are achieved. It could serve as an example for delivering an indicator-based assessment. Results of applications show that the performance of the different methods varies depending on the nature of the variability along the time series. Short time series with high variability are statistically less easily interpretable. Indicators of spatial distributions often alert on stock status as well as the index of age at maturity. The wide range of biological and spatial indicators considered may provide advance warning in comparison to using abundance indices alone. The indicator-based assessment compared with conventional assessments along the past series, meaning commercial catch data are not necessary to assess fish stock status. The applications were performed using the R script tools developed. The application to all project case study stocks demonstrated the potential of the tools for delivering indicator-based assessments in operational mode in a wide range of situations. Procedures could now be applied to provide indicator-based assessments to expert groups for any stock that is monitored with research surveys.

Document 3 documents and tests survey-based assessment models that used survey indices of abundance only. A variety of models were considered that had different input data requirements (e.g., aggregated biomasse, abundance at length, abundance at age) and had different assumptions for the mortality across ages and years. The models were tested for their ability to capture changes in abundance time series, in different scenarios of stock (depletion/recovery) and survey performance (selectivity, catchability, noise in indices). The tests used published NRC simulated data with known properties. The survey-based assessment models behaved as smoothers for noisy indices and were able to reliably capture the major signal in biomass and recruitment, although they smoothed out transient changes. Based on survey data only, models could not provide absolute estimates of stock size but tests indicated that they would provide useful indications on trends, to which managers might wish to react.

Document 4 describes the models and tools of the FLR simulation evaluation loop, with reference to the FLR website when appropriate. The framework implements the following models that interact in a dynamic simulation loop: an operating model that provides the underlying stock dynamics, an observation model that provides (based on the operating model outputs) survey indices with specified error and bias, a harvest control rule model (HCR) that provides the management options and substract from the population fishery catches. Also documented are graphical display tools that enable visual analysis of the simulation outputs. The range of HCRs that can be considered depend to some extent on the biological complexity in the operating model. Here the operating model was a classical VPA-like population model (age-based with stock recruitment relationship) with a yearly time step, which was appropriate for the type of data used in ICES assessment working groups. The HCRs considered were rules for defining TAC based on survey abundance indices or on Z and were appropriate for the current ICES management context.

Document 5 compiles the application of the FLR simulation evaluation tools on case study stocks (North Sea herring,North Sea cod, North East Arctic cod, Bay of Biscay anchovy). Management plans for these stocks are documented. The capability of HCRs to management the stock using survey data only is tested different uncertainties in the survey indices, in the stock dynamics or misreporting. The performance of different HCRs are compared (Z-based, TAC-based, TAC-based with addition triggering indicator alarm). For these purposes a list of performance statistics is defined. Results show that it is possible to obtain well performing HCRs based on survey-derived information only. Commercial catch data are therefore not necessary to provide management advice. HCRs need to be more conservative for controlling the system when its variability is higher, regardless of whether the variability comes from external factors (e.g., misreporting, noise in survey data) or the stock itself (recruitment). Rather than maximise the catch based on modelling assumptions, results show that it is better and possible to design robust HCRs that will perform well given the many uncertainties.

Document 6 is a suite of papers describing the Aladym simulation models and tools. Aladym is an age-based population operating model that considers age to length conversion and also formulates length-dependent natural mortality, fecundity and gear selectivity. The model time step is the month allowing, which allows to consider progressive length-dependent recruitment to the population during the year. The model is appropriate for the management context currently in place in the Mediterranean. With the Aladym model, the HCRs that can be considered range from TAC to mesh size regulation and seasonal closures. A stochastic version of Aladym was developed to account for uncertainty in the biological parameters, which are then considered as random variables. Metrics were defined to characterise stock status and fishing pressure. Last, survey-derived information that is input to the Aladym model is documented.

Document 7 compiles the application of the simulation evaluation Aladym tools on case study stocks (red mullet in the Tyrrhenian sea, hake in the Aegean sea, hake in the Bay of Biscay, cod in the Baltic sea). The long-term evolution of these stocks was simulated using model parameter values estimated from survey data. Simulations were performed for different complementary purposes. Correlation between indicators of fishing pressure and biological parameters was analysed in order to test whether they could be used to alert on a degrading stock status. These analyses are complementary to the indicator-based assessment (indicator testing), and in particular when survey time series is short. Simulation experiments also allowed to estimate reference values for Z that were compatible with the long-term stock viability given uncertainties in recruitment and biological parameters. Last, complex multi-annual management plans were considered in which seasonal closures added to reduced TACs. Such plans performed better than total fishing ban during too few years and were also considered more realistic.

Document 8 summarises and puts into perspective all project developments and results obtained on the case study stocks. It suggests a protocole for implementing a comprehensive assessment procedure based on survey data. The different steps of the protocole are to agree on management objectives, select indicators and reference values, select methods to detect changes in time series and how to elaborate the assessment based on their combination. The indicators under consideration should not only concern the biology of the stock but also fishing pressure and survey quality. The survey-based assessment being based on statistically identified changes in indicator time series, it can only be relative to a reference stock status. But there is no reason why management decisions might not be based on relative assessments. The last step of the protocole is therefore to agree on management response to good and bad assessment results. The document further describes how indicator-based asssessment and simulation evaluation are complementary, for testing indicators, estimating reference values as well as management responses to indicator-based assessments. It also suggests how to use comprehensive survey-based assessments along side conventional assessments and the increased reliability gained by using a wider base of knowledge for assessing fish stocks. Last, the consideration of a wide range of indicators opens the way to assessing fish stocks with an ecosystem perspective.

PLAN FOR USING AND DISSEMINATING KNOWLEDGE

Overview table

Date	Type	Type of audience	Countries addressed	Size of audience	Partner responsible / involved
April 2004	Conference	GFCM-FAO	European + Mediterranean	60	Sibm
June 2004	Fisboat website http://www.ifremer.fr/drvecohal/fisboat/	Public	worldwide	worldwide	Ifremer
October 2004	Fisboat Leaflet	Public and Research	Europe	≈ 700	Ifremer
February 2005	Conference	Fisherman Associations	Italy	20	Sibm 1 paper
May 2005	Conference	Fisherman Associations	Italy	15	Sibm 1 paper
May 2005	Conference	Marine Biology Scientists	Italy	40	Sibm 1 paper
Sept. 2005	Conference ICES ASC	Research	International	≈ 700	Armines, Ifremer 2 papers
Jan. 2006	Working group Association Française d'Halieumétrie	Research	France	20	Armines , Ifremer 1 paper
June 2006	Conference ICES Symposium	Research	International	≈ 500	Imperial College, Ifremer 2 papers
Sept 2006	Conference ICES ASC	Research	International	≈ 700	Armines, Ifremer, I.Coll. 3 papers
Nov. 2006	Working group ICES WGACEGG	Research	International	≈ 20	Ifremer, Azti 3 papers
March 2007	Working group GFCM stock assessment	Research	Mediterranean countries		Sibm 1 paper
March 2007	Working group ICES WGMGM	Research	International		Cefas 1 paper
March 2007	Working group ICES HAWG	Research	International		Frs 1 paper
June 2007	Conference Association Française d'Halieumétrie	Research	France	≈ 100	Ifremer 1 presentation
July 2007	Press	European Parlement	European countries		Ifremer interviewed 1 article
Sept 2007	Conference ICES ASC	Research	International	≈ 700	Armines, Ifremer, Cefas, Sibm 5 papers
2007	Publications	Research	International		Cefas, Ifremer, Armines 6 articles
2008	Publication of a special journal volume expected	Research	International		Ifremer, Cefas, all partners

Research articles

Cotter, J., Mesnil, B. and Piet, G. 2007. Estimating stock parameters using year class curves. ICES Journal of Marine Science, 64: 234-247.

Kell, L., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M., Poos, J. and Scott, F. 2007. FLR : an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640-646.

Poulard, J.-C. and Trenkel, V. (accepted). Do survey design and wind conditions influence survey indices ? Canadian Journal of Fisheries and Aquatic Sciences

Trenkel, V., Rochet, M.-J. and Mesnil, B. 2007. From model-based prescriptive advice to indicator-based interactive advice. ICES Journal of Marine Science, 64: 768-774.

Trenkel, V. (in revision). A biomass random effects model (BREM) for fish stock assessment and management with application to Bay of Biscay anchovy. Canadian Journal of Fisheries and Aquatic Sciences.

Woillez, M., Poulard, J.-C., Rivoirard, J., Petitgas, P. and Bez, N. 2007. Indices for capturing spatial patterns and their evolution in time with an application on European hake (Merluccius merluccius) in the Bay of Biscay. ICES Journal of Marine Science, 64: 537-550.

ICES CM Papers

Cotter, J., Fryer, R., Mesnil, B., Needle, C., Skagen, D., Spedicato, M-T. and Trenkel V. 2007. A review of fishery-independent assessment models, and initial evaluation based on simulated data. ICES CM 2007/O:04.

Cotter, J., Petitgas, P. et al. 2007. FISBOAT manual of indicators and methods for assessing fish stocks using only fishery independent survey data. ICES CM 2007/O:27.

Petitgas, P., Poulard, J.-C., Radtke, K., Spedicato, M.-T., Ibaibarriaga, L., Politou, C.-Y., Korsbrekke, K. , Deernberg, C. and Fernandes, P. 2007. Comprehensive indicator-based diagnostics of fish stocks using fishery-independent survey data: the FISBOAT report. ICES CM 2007/O:16.

Pomarede M., Simmonds E. J., Hillary R., McAllister M., Kell L., Needle C., 2006. Evaluating the management implications of different types of errors and biases in fisheries resources surveys using a simulation-testing framework. ICES CM 2006/I:28.

Poulard, J.-C. and Trenkel, V. 2005. Relationship between survey indices and survey design and wind conditions: Bay of Biscay groundfish survey. ICES CM 2005/Z:02

Spedicato, M.-T., Woillez, M., Rivoirard, J., Petitgas, P., Carbonara, P. and Lembo, G. 2007. Usefulness of the spatial indices to define the distribution pattern of key life stages: an application to the red mullet (Mullus barbatus) population in the south Tyrrhenian sea. ICES CM 2007/O:10.

Trenkel, V. 2007. A biomass random effects model (BREM) for stock assessment using only survey data: application to Bay of Biscay anchovy. ICES CM 2007/O:03.

Woillez, M., Petitgas, P., Rivoirard, J., Poulard, J.-C., and Bez, N. 2005. Indices for capturing spatial pattern and change across years of a fish population: an application on European Hake (Merluccius merluccius) in the Bay of Biscay. ICES CM 2005/L:16.

Woillez, M., P. Petitgas, J. Rivoirard, J.-C. Poulard, P. Fernandes, R. ter Hofstedte, K. Korsbrekke, A. Orlowski, M.-T. Spedicato and C.-Y. Politou. 2006. Relationships between population spatial occupation and population dynamics. ICES CM 2006/O:05

Woillez, M., Rivoirard, J. and Fernandes, P. 2006. Evaluating the uncertainty of abundance estimates from acoustic surveys using geostatistical conditional simulations. ICES CM 2006/I:15

Woillez, M., Rivoirard, J., Petitgas, P. and Deerenberg, C. 2007. Selecting and combining survey based indices of fish stocks using their correlation in time to make diagnostics of their status. ICES CM 2007/O:07.

Conference papers

Pomarede M., Hillary R., Kell L., Needle C, Simmonds E. J., McAllister M., 2006. Evaluating the relative merits of fishery dependent and independent data in fisheries management. ICES Symposium on Fisheries Management Strategies, June 27th-30th Galway, Ireland.

Trenkel, V., Rochet, M.-J. and Mesnil, B. 2007. From model-based prescriptive advice to indicator-based interactive advice. ICES Symposium on Fisheries Management Strategies, June 27th-30th Galway, Ireland.

Working Documents to ICES and GSCM Expert Groups

Cotter J., Fryer R., Mesnil B., Needle C., Skagen D., Spedicato M.-T. and Trenkel V. 2007. A review of Fishery-Independent assessment models, and initial evaluation based on simulated data. Working Document to the ICES Working Group on Methods of Fish Stock Assessment, Woods Hole, March 2007

Ibaibarriaga L. and Petitgas P. 2006. Catchability analysis between abundance estimates with DEPM and Acoustic methods. Working Document to the ICES Working Group on Acoustic and Egg surveys for sardine and anchovy in ICES areas VIII and IX, Lisbon November 2006.

Petitgas, P., Massé, J., Beillois, P. and Coppin, F. 2006. Proposition for a common data base structure for acoustic surveys. Working Document to the ICES Working Group on Acoustic and Egg surveys for sardine and anchovy in ICES areas VIII and IX, Lisbon November 2006.

Spedicato M.T., M. Woillez, J. Rivoirard, P. Petitgas, P. Carbonara, G. Lembo. 2007. Usefulness of the spatial indices to define the distribution pattern of key life stages: an application on the red mullet population in the south Tyrrhenian sea. GFCM-SAC-Sub-Committee Stock Assessment. Workshop on trawl survey based monitoring fishery system in the Mediterranean, Rome, Italy, 26-28 march 2007. 15 pp.

Trenkel, V. 2006. Combining acoustic and DEPM survey indices in the biomass random effects model for stock assessment. Working Document to the ICES Working Group on Acoustic and Egg surveys for sardine and anchovy in ICES areas VIII and IX, Lisbon November 2006.

Website

The FISBOAT web site http://www.ifremer.fr/drvecohal/fisboat/ was an important tool for dissemination and is expected to carry on that role. The website is expected to be maintained and will give access to project outcomes and related matters. All project products are available on the website: software codes and their documentation, data, case study reports, manuals of methods, project deliverables. All meeting documents were also posted on the website (agendas, meeting reports) as well as project reports (interim and final). Six documents were produced that compiled the project outcomes: a Manual of indicators and methods and a Report on their application to case studies, a Report on survey-data-only assessment models and their performance, a Manual on the FLR simulation evaluation loop and a Report of its application to case studies, a Report on how to create comprehensive assessments.

ICES 2007 Theme Session O

IFREMER organised and co-chaired the ICES 2007 Annual Science Conference Theme Session O on ‘Flying outside the ICES Assessment WG paradigm – Alternative approaches to providing fisheries management advice’. The Theme session was intended to be a forum for presenting alternative methods using fishery-independent information and stakeholder involvement for provide effective means to diagnosing the status of marine resources and communities and identifying management alternatives. The report available at : http://www.ices.dk/iceswork/asc/2007/themesessions.asp). Papers from the Fisboat project presented reviews of fishery-independent methods (O:4, O:16), described the manual produced by the project (O:27), provided application to case studies (O:10), and examples of the application of techniques (O:3, O:7). The final discussion led to suggest that ICES considers alternative stock assessment methods alongside traditional methods for selected stocks as a means of complementing current methods.

Publication project - Fisboat special volume

IFREMER and CEFAS organised to publish jointly the novel fishery-independent assessment methods developed and their case study applications in a special volume of the journal Aquatic Living Resources. The special volume is expected to be written in a Manual style. Publication is expected for the end of 2008. The contents of the volume has been organised as follows.

Manual of Fish Stock Assessment using Surveys and Indicators
Eds. P. Petitgas (Ifremer), J. Cotter (Cefas), V. Trenkel (Ifremer), B. Mesnil (Ifremer)
Special volume of Aquative Living Resources

Introduction

Article no.01 : Fish Stock Assessment using Surveys and Indicators : an overview (Petitgas)

Section 1 : Surveys and indicators

Article no.02 : Choices of surveys and indices (Cotter, Trenkel)
Article no.03 : Manual of biological indicators (Trenkel, Cotter et al.)
Article no.04 : Manual of spatial indicators (Woillez, Rivoirard)
Article no.05 : Combining raw indicators into multivariate indicators (Petitgas, Poulard, Rivoirad)

Section 2 : Methods to analyse time trends and changes

Article no.06 : Non parametric methods for trends (Cotter)
Article no.07 : Detection of recent trends and power analysis (Trenkel, Bogaards)
Article no.08 : A statistical process control approach to detect change (Mesnil, Petitgas)
Article no.09 : Analysis with Min/Max autocorrelation factors (Woillez, Rivoirard)

Section 3 : Fishery-indepent assessment methods and management strategies

Article no.10 : Indicator-based assessment and forecasting (Trenkel, Petitgas, Woillez)
Article no.11 : Fishery-independent assessment models (Mesnil, Cotter, Trenkel, Needle et al.)
Article no.12 : Fishery-independent management strategies and control rules (Cotter, Bogaards et al.)

Section 4 : Simulation methods
Article no.13 : Manual on FLR tools (Hillary et al.)
Article no.14 : Manual on ALADYM (Lembo, Abella, Fiorentino, Spedicato)

Section 5 : Illustrative applications

Article no.15 : Indicator-based methods applied to case studies (Petitgas, Poulard, Radtke et al.)
Article no.16 : FLR tools applied to case studies (Apostolaki, Ibaibarriaga, Bøthun, Bogaards, Pomarede)
Article no.17 : ALADYM applications to case studies (Spedicato, Poulard, Radtke, Politou et al.)

Concluding articles

Article no.18 : Comprehensive assessments and management strategies (Cotter, Petitgas, Mesnil, Abella et al.)
Article no.19 : Perspective from an invited author external to the Fisboat project

Incorporation of FISBOAT methods into the Fisheries advisory toolbox

Fisboat simulation evaluation tools have been used during the course of the project in working groups of assessment bodies: FLR has been used for testing harvest rules on North Sea herring (ICES) and ALADYM for testing technical measures on red mullet (GFCM). Indicator-based methods are expected to be applied within ICES in 2008 for contributing to the assessment of data poor species at the ICES Working Group on Assessment of New MoU Species (WGNEW). Little information is known on these species but time series of survey data exist and the European Commission is asking for scientific advice on them. The ICES Study Group on Management strategies (SGMAS) may be a forum for continued collaborative methodological development of fishery-independent methods, in particular on what management rules to develop based on indicator-based assessments and methods and tools for evaluating harvest control rules. In addition, the EU project IMAGE (Indicators for fisheries management in Europe) may allow to continue on some of the topics developed in the Fisboat project. Also, the new science program of ICES is expected to develop coordinated research activity on the Identification of indicators, models and methods to ensure high quality advice for integrated management under the ecosystem approach.

Thus fisheries management is now clearly envisaged in an ecosystem approach and there is an emerging need for indicator based assessments for data poor species but also for integrated ecosystem management. The fisheries advisory tool box of methods is therefore expected to expand and the fishery-independent assessment methods developed in the Fisboat project are thus expected to be in the list of new methods. Further developments of fishery-independent methods for ecosystem and fish stock monitoring,assessment and advice are expected to take place in the coming years, which will build on the Fisboat project outcomes. The Fisboat Manual of methods and applications (special volume) is expected to contribute to that purpose.

Document 1: Indicator-Based Assessment – Methods

ICES CM 2007/O:27

MANUAL OF INDICATORS AND METHODS FOR ASSESSING FISH STOCKS USING ONLY FISHERY-INDEPENDENT, SURVEY BASED INFORMATION

Contribution to EC research project:

Fishery Independent Survey Based Operational Assessment Tools (FISBOAT),

DG-Fish, STREP n° 502572 (2004-2007)

Co-ordinator: P. Petitgas, IFREMER, Nantes, France

compiled by
J. Cotter, CEFAS, Lowestoft, Suffolk, NR33 0HT, UK
email: john.cotter@cefas.co.uk

from contributions by :
B Mesnil, IFREMER, Nantes, France
V Trenkel, IFREMER, Nantes, France
M-J Rochet, IFREMER, Nantes, France
P Petitgas, IFREMER, Nantes, France
M Woillez, Ecoles des mines, Fontainebleau, France
J Rivoirard, Ecoles des mines, Fontainebleau, France
J Cotter, CEFAS, Lowstoft, England
A Uriarte, AZTI, Pasaia, Spain
P Witthames, CEFAS, Lowstoft, England
M T Spedicato, COISPA Tecnologia & Ricerca, Bari, Italy
G Lembo, COISPA Tecnologia & Ricerca, Bari, Italy
H Bogaards, IMARES, IJmuiden, The Netherlands

Contents

Contents 12
Abstract 14
Keywords: 14
1. Introduction 15
2. Surveys and estimation 19
2.1 Introduction 19
2.2 Potential limitations of survey data as the unique source of information for stock assessment and management 20
2.3 Estimation of indicator values from surveys 24
3. Biological indicators 27
3.1 Introduction 27
3.2 Intrinsic population growth rate, r 28
3.3 Total mortality, Z 29
3.4 Numbers-at-length, numbers-at-age: NaL, NaA 31
3.5 Spawning Stock in number: SSN 33
3.6 Length statistics: Lbar, L25, L50, L75 35
3.7 Survey catch weight: W 38
3.8 Condition: C 41
3.9 Gonadosomatic index: GSI 43
3.10 Length and age at maturity: LaM50, AaM50 45
4. Spatial indicators 47
4.1 Introduction 47
4.2 Centre of gravity: CG 49
4.3 Inertia: I 51
4.4 Anisotropy and isotropy: An, Is 53
4.5 Global index of collocation: GIC 55
4.6 Number of spatial patches: NOP 57
4.7 Positive area: PA 59
4.8 Spreading area: SA 61
4.9 Equivalent area: EA 63
4.10 Microstructure index: MI 65
4.11 References for spatial indicators 66
5. Methods 68
5.1 Introduction 68
5.2 Simulating population dynamics 70
5.2.1 ALADYM(v 08) 70
5.2.2 Estimating indicators and reference points 81
5.3 Indicator time-series methods. 85
5.3.1 Nonparametric method for determining recent trends 85
5.3.2 Assessing the power to detect future trends 93
5.3.3 Statistical Process Control (SPC) schemes 96
5.3.4 Nonparametric statistical methods for assessing trends 102
5.4 Construction of multivariate indicators. 125
5.4.1 Principal Components Analysis (PCA) and biological indicators 125
5.4.2 Multi-factorial analysis (MFA) and spatial indicators 129
5.4.3 Min/Max autocorrelation factors (MAFs) and time continuity 133
5.5 Diagnosing stock status from indicator series. 136
5.5.1 Combining trend signals using a cause-effects table 136

5.5.2 A 'traffic light' procedure based on Cusum out-of-control tables. 140
5.5.3 A multivariate statistical procedure. 143
Format for documenting indicators... 146

Abstract

This manual discusses the use of indicators for assessing the state of a fish stock. The indicators envisaged are estimated from trawl surveys, or trawl-supplemented surveys such as egg and acoustic surveys, in which catchability is maintained constant so far as practically possible. A preliminary section considers factors that determine the appropriateness of a survey for each species and indicator, and makes suggestions for statistical estimators for quantitative and descriptive indicators. A widely applicable selection of biological and spatial indicators is then documented in a standardised format that includes references to examples of their use and modifications. The biological indicators include those relating to quantity of fish, size, and reproduction. The spatial indicators characterise the geographic distribution of a stock and make allowance for low or zero densities of fish at some stations. The final section of the manual presents several presentational and statistical methods for assessing and interpreting trends in indicators. A stock simulation model is described that may assist with determination of reference points for biological indicators. Since indicators tend to be highly specific and normally many would be used to assess the state of a stock, multivariate methods form an important part of this section. Indicators offer valuable biological and geographic information for supplementing existing model-based stock assessments. They are also likely to form an important component of an ecosystem approach to fishery management, or they could be used pragmatically to tune harvest control rules in a form of adaptive management.

Keywords:

fish stock assessment; trawl survey; fish survey; biological indicator; spatial indicator; stock status;

1. Introduction

Indices are measurable quantities that can be used to characterise a fish stock, for example, the mean length of the fish or the centre of gravity of its geographic distribution. When variation in a particular index relates to a process in the stock that one wants to follow, the index becomes an indicator of that process. This manual discusses indicators that are estimated from consistent time-series of results from scientific surveys conducted by research vessels (Anonymous 2004), or by other fishing vessels being operated purely for the ends of a survey and not for the pursuit of catches for commercial gain (since that goal would almost certainly cause bias). The surveys envisaged use a standardised trawl (Anonymous 2006) with small-mesh codends such that selectivity can be assumed to be reasonably constant for all fish larger than the selection range of the codend mesh. The trawling may only have a supporting role in the survey, for example for acoustic or egg surveys. For surveys using other catching or sampling methods, the applicability or otherwise of each indicator should be carefully considered in relation to selectivity. The fish species envisaged breed and recruit annually; some of the indicators may not be suitable for species that do otherwise, e.g. tropical species. A single species is assumed unless stated.

The relevance of indicators to the management of fisheries has become more widely acknowledged in recent years. In Northern Europe at least, the scientific community advising on fisheries management has mostly relied on quite complicated models to assess the state of fish resources and make recommendations. This approach entails significant costs in terms of data volume and quality, and of expertise as well, and is only affordable for the top-valued species and fisheries. Nowadays, other species that were of secondary interest in the past have become the primary targets of fishers, and scientific advice on the conditions for their sustainable exploitation is also sought by managers. In most instances, however, the rich data bases required for the conventional assessment models are lacking; other routes have to be considered, such as indicator-based methods. Moreover, States have embraced the "ecosystems approach to fisheries management" which, however defined, basically implies that the interactions between ecosystems' states or processes and fisheries have to be accounted for when decisions are made for fisheries. There is a broad consensus within the scientific community that building detailed quantitative models, continuing the traditional fishery science approach, to advise ecosystems-based management is simply not an option given the gaps in knowledge and the huge cost of getting the appropriate data. The alternative route is to identify and monitor a suite of indicators that reflect the state of, and possibly the human pressure on, the marine systems; management action is then advised based on observed changes in the indicators. Considerable scientific efforts are underway concerning indicators, as evidenced by a burgeoning literature.

There are at least two potential difficulties in this context. Firstly, like other environmental issues, fisheries issues are highly controversial. It can be expected that when systems of indicators – or indicator-based assessments – eventually get included in the management decision framework, they will be exposed to strong reactions, perhaps more so than traditional fish stock assessments, if only because they will be novel. A prime concern, therefore, is that the process to infer the state of fish stock, ecosystems, etc. from indicators should be formalised, in the sense of being rational, objective, defensible, and replicable by others; it should be amenable to non-ambiguous descriptions that are intelligible to stakeholders and managers.

Secondly, marine ecosystems and fish stocks are notoriously subject to high variability. Yet, it is quite likely that several indicators will be based on relatively limited samples (in terms of size, frequency, seasonal or area coverage). This will be particularly true for indicators collected during scientific sea surveys, which are already an important recourse for monitoring fish stocks when commercial fishery data are lacking or deteriorating (the topic dealt with by the FISBOAT project), and will be the prime – or even the sole – source for several ecosystems indicators. Surveys mobilise costly vessels and typically involve a limited number of stations and samples of moderate size, thus survey indicators often have large CVs. In any case, the signal-to-noise ratio is likely to be low. Procedures therefore have to be found to avoid casting measurement noise straight into the advice, and triggering undue action with all the political fuss that may follow.

For the present, indicators could be used to supplement existing methods of model-based, single-species stock assessment and management (Demaré 2006). This would incorporate some additional biology into what is otherwise mainly a computational exercise. In due course, indicators are likely to form a fundamental part of an 'ecosystem approach to fisheries' (EAF) (Garcia and Cochrane 2005), Jennings (2005), Cury and Christensen (2005). Indicators might also be used pragmatically to inform a management system based on harvest or effort control rules negotiated relatively from year to year between management and industry. A suite of well-chosen indicators are envisaged to tune such a system so that, after an initial period of trial and error, indicators relevant to the health of the stock respond somewhat predictably to management actions. This would be a form of adaptive management (Walters 1986).

This manual is intended as a contribution to the necessary formal structure for using indicators to assist the management of fisheries. A preliminary, general section, section 2, discusses the appropriateness of the surveys used, and some possible estimators that can give different results from the same survey. Section 3 summarises a small selection of potentially useful indicators of the biological state of a stock of fish. Section 4 considers indicators of the spatial state of a fish stock, most of which allow for the occurrence of low or zero densities of a species of fish at some stations, making them very generally applicable. The last section, section 5, presents methods to assist with the interpretion of time-series of different indicators, singly and in combination. This aspect is important because indicators tend to be highly specific, so that use of many is often necessary to gain a full picture of a stock (Rice and Rochet 2005). Additionally, the interpretation of indicator series often depends more on their trends up or down over time than on their absolute values (Jennings and Dulvy 2005; Trenkel *et al.* 2007). The general problem is how to assemble all the different results in a way that is informative and suitable for justifying decisions about management of the fishery.

Table 1.1 is a list of the state indicators described in this manual, the marine environmental processes and population characteristics that they relate to, and the primary authors from the FISBOAT project in each case. The indicators are described briefly using a standardised format; each relates to a specific biological or spatial characteristic of the stock, or of selected age or length classes within it. In several cases, the indicator described is one example from a suite of related indicators; where possible, references are provided to allow further information to be followed up. The manual includes but is not limited to indicators and methods trialled during the FISBOAT project, 2005 to 2007. The methods trialled were limited to those applicable to most of the project case studies with the available data and within the available time. The FISBOAT project also considered non-indicator based methods of stock assessment using surveys; these methods are described in a related project manual entitled 'Review of fishery-independent assessment methods' prepared by Mesnil (ICES CM2007/O:04).

References

Anonymous (2004) Report of the workshop on survey design and data analysis (WKSAD). *International Council for the Exploration of the Sea*, Copehagen. ICES CM 2004/B:07, Ref. D.G., 261pp.

Anonymous (2006) Report of the study group on survey trawl standardisation (SGSTS). *International Council for the Exploration of the Sea*, Copenhagen. ICES CM 2006/FTC:05, 67 pp.

Cury, P.M. and Christensen, V. (2005) Quantitative ecosystem indicators for fisheries management. *ICES Journal of Marine Science* **62**, 307-310.

Demaré, W. (2006) Fisheries-based management and advice in Europe. In: *The knowledge base for fisheries management.* ed Motos, L. and Wilson, D.C., Elsevier**:** 295-305.

Garcia, M. and Cochrane, K.L. (2005) Ecosystem approach to fisheries: a review of implementation guidelines. *ICES Journal of Marine Science* **62**, 311-318.

Jennings, S. (2005) Indicators to support an ecosystem approach to fisheries. *Fish and Fisheries* **6**, 212-232.

Jennings, S. and Dulvy, N.K. (2005) Reference points and reference directions for size-based indicators of community structure. *ICES Journal of Marine Science* **62**, 397-404.

Rice, J.C. and Rochet, M.-J. (2005) A framework for selecting a suite of indicators for fisheries management. *ICES Journal of Marine Science* **62**, 516-527.

Trenkel, V.M., Rochet, M.-J. and Mesnil, B. (2007) From model-based prescriptive advice to indiator-based interactive advice. *ICES Journal of Marine Science* **64**, in press.

Walters, C. (1986) *Adaptive management of renewable resources.*, Macmillan Publishing Co., New York

Table 1.1 List of the indicators documented in this manual, the marine environmental processes and population characteristics that they are thought to relate to, and the primary authors of text.

Indicator and abbreviation	Processes affecting	Population characteristics	Contributors
Intrinsic population growth rate, *r*	Fishing, natural mortality, reproduction	Numerical abundance summed over all ages	Trenkel
Total mortality, Z	Fishing, natural mortality, migrations with age into or out of survey area	Rate of dying, migrations related to age	Rochet, Trenkel, Cotter
Spawning stock in number, SSN	Maturation, fishing, natural mortality, nutrition	Abundance of potentially breeding fish, sustainability of the stock	Mesnil, Uriarte, Witthames
Length statistics, Lbar, L25, L50, L75	Recruitment, growth, fishing, natural mortality	Growth, length frequency distribution, recruitment	Trenkel, Mesnil, Cotter
Total weight caught, *W*	Fishing, natural mortality, growth, feeding	Numerical abundance, age composition, growth	Cotter
Condition, *C*	feeding, growth	Nutritional status of individuals, reproductive fitness	Cotter, Witthames
Gonadosomatic index, GSI	feeding, maturation	Nutritional status, reproductive fitness	Cotter, Witthames
Length and age at maturity, LaM50, AaM50	maturation, fishing mortality, evolutionary selection	Size and age of potentially breeding fish.	Rochet, Trenkel, Witthames, Cotter
N-at-length, N-at-age, NaL, NaA	Recruitment, growth, fishing, natural mortality	Length and age frequency distribution	Cotter
Centre of gravity, CG	Migrations, climate change, fishing, population size	geographic location of the whole population or of concentrations of fish	Woillez, Rivoirard, Petitgas
Inertia, I	Dispersal, environmental change, migrations	Changing population size, migrations, climate and environmental changes	Woillez, Rivoirard, Petitgas
Anisotropy, An, Isotropy, Is	Depth, currents, proximity to shore	Alignment of the population in relation to environmental gradients	Woillez, Rivoirard, Petitgas
Global index of collocation, GIC	Competition, genetic differences, dispersal	Geographical overlap of two populations	Woillez, Rivoirard, Petitgas
Number of spatial patches, NOP	Dispersal, common attractants, lack of mixing	Patterns of movement, foraging strategies, population size	Woillez, Rivoirard, Petitgas
Positive area, PA	Dispersal without regard to variations of abundance	Population size, habitat preferences, food availability	Woillez, Rivoirard, Petitgas
Spreading area, SA	Dispersal with regard to variations of abundance	Population size, habitat preferences, food availability	Woillez, Rivoirard, Petitgas
Equivalent area, EA	Dispersal assuming uniform abundance	Population size	Woillez, Rivoirard, Petitgas
Microstructure index, MI	Small-scale variability of habitat, abundance	Relationship of population to environment	Woillez, Rivoirard, Petitgas

2. Surveys and estimation

2.1 Introduction

This preliminary section is intended to be general to the use and interpretation of indicators from fish surveys. In section 2.2, Trenkel warns of dangers arising from placing excessive reliance on survey results for the estimation of indicators without first carefully considering whether the survey will in fact provide appropriate data for the species and indicators of interest. In section 2.3, Cotter discusses alternative estimators that can give widely differing results for an indicator using the same survey data.

2.2 Potential limitations of survey data as the unique source of information for stock assessment and management

V. Trenkel
IFREMER, Nantes, France

Introduction

There are several issues that need to be considered when using survey data as the unique source for carrying out stock assessments and providing management recommendations. The main issues leading to bias or uncertainty regarding evolution of the stock are

1. the surveyed area does not encompass the stock area;
2. the size/age classes sampled are not representative of the stock;
3. variation in survey catchability.

Survey area ≠ stock area

There are various reasons why survey areas might not encompass stock areas, in addition to the problem of stock boundaries not being well known, or survey areas varying between years. The simplest reason is that part of the stock, or certain age classes, are outside the survey area. No single survey will cover the whole stock area for geographically wide spread species such as northern hake. For other species the problem might be that certain age groups are too deep to be caught by the survey gear, or too shallow for the survey vessel to access them, or they are not accessible to the survey gear because their habitat is for example not trawlable. Anchovy in the Bay of Biscay is an example of a species with a variable proportion of recruits too close to the coast and thus in too shallow water for the survey vessel being used. The visible effect of this is that numbers at age 2 are higher than numbers at age 1 in the previous year (Fig 2.2.1a). Similarly for cod in the North Sea IBTS survey as shown by the negative mortality rates for age 1 in many years (Fig 2.2.1b). The negative Z for cod age 5 are probably an effect of small sample sizes. The same effect can of course be caused by gear selectivity so that in each case it is necessary to find the most plausible explanation.

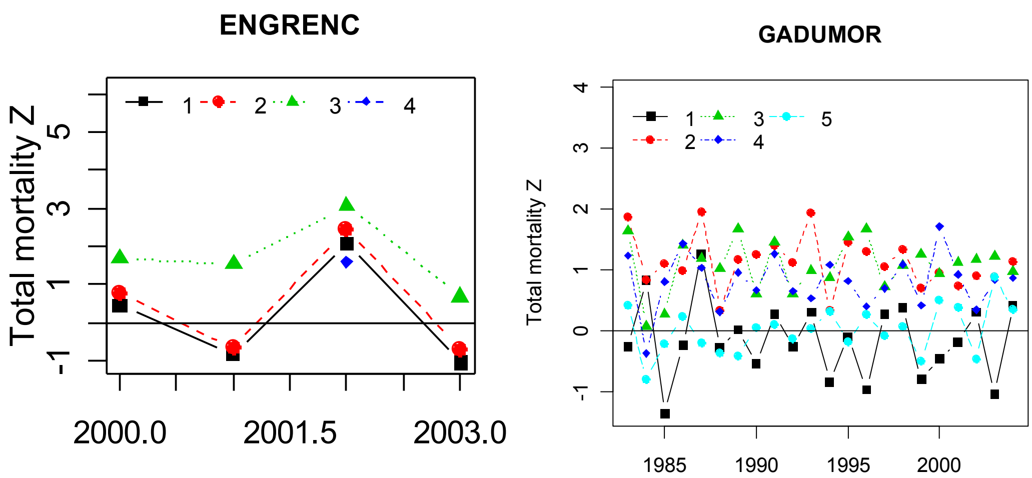


Figure 2.2.1. Total mortality estimates Z by age and year derived from survey numbers-at-age. Z estimates for age a in year t refers to the total mortality between age a in t and age a+1 in t+1. a) anchovy in Bay of Biscay; b) cod in North Sea.

In addition to the above issues, stocks might move out of, or into the survey area in response to changing environmental conditions. Diel migrations or other activity patterns can also lead to variability in availability to the survey gear.

Size classes sampled not representative of stock

Many scientific surveys have been designed as young fish surveys. For example, what is called today the *International Bottom Trawl* survey (IBTS) started off as the *International Young Herring Survey* (IYHS) in the North Sea, then became the *International Young Fish Survey* (IYFS) before finally obtaining its current name. The change in objectives reflected in the varying names did not imply any change in design, rather a modification of the list of species for which information was collected. Hence in response to the initial objectives, a sampling trawl designed for catching young fish is still used today (GOV 36/47). The time of year of the survey was decided similarly. The initial survey took place in the first quarter as herring juveniles are then available. Currently a third quarter survey is also carried out.

When the IBTS survey was extended to the Bay of Biscay and Celtic Sea (French EVHOE survey), the same GOV trawl was adopted despite the fact that substrates are often more difficult for trawling and the GOV is more suited for soft bottoms. The GOV was slightly adapted by removing the exocet Kite and replacing it by 6 additional floats. As the main target species are hake, megrim and monkfish, the survey is carried out in the fourth quarter when the recruits of those species become accessible.

The consequence of designing surveys to target recruits is that there can be the problem of exploited size classes not being well represented in the survey catches. This can be due to the survey gear being used (selectivity), the vessel speed, or of course an area mismatch dealt with above.

As an example, consider the length distribution of hake in landings and in the EVHOE autumn groundfish survey (Fig 2.2.2). In the survey the bulk of individuals is smaller than 20 cm while in the landings the distribution is rather flat between 30 and 60 cm. Note that the legal landing size is 27 cm. The survey catches very few individuals in the size range targeted by the fishery. Thus, for northern hake, it seems that the EVHOE survey might be suitable to provide recruitment (age 0) and perhaps age 1 estimates, but it is unlikely that variations in total stock abundance or other indices relating to the adults will be captured reliably.

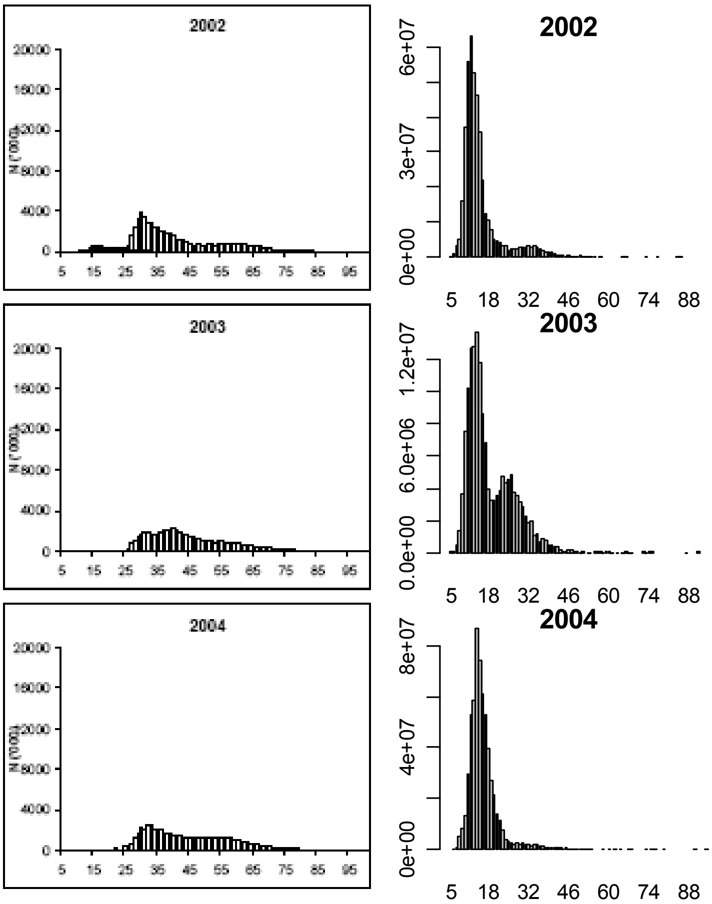


Figure 2.2.2. Length distribution of hake landings for the northern stock (left) and hake caught in the autumn groundfish survey in the Bay of Biscay and Celtic Sea (right).

Variations in survey catchability

A range of factors can make survey catchability vary between hauls and interannually. Between-haul variability will most likely reduce the precision of survey indices while interannual variation might bias estimates and affect time trends. The latter might be called a year-effect in survey catches. A study of the potential year-effect in survey catches for the EVHOE autumn groundfish survey taking place in the Bay of Biscay showed that, on average, 20% of interannual variation in abundance indices could be explained by survey conditions for benthic species, 11% for demersal, and none for pelagic species (Poulard and Trenkel, submitted) (Fig. 2.2.3). In contrast, survey conditions explained a smaller and decreasing part of the interannual variability in the coefficients of variations of these abundance indices and in species mean weight. Thus survey conditions might bias survey indices. In the same study it was found that taking account of survey conditions could alter time trends in species' abundance indices and, as a consequence, influence the stock assessment based on survey information.

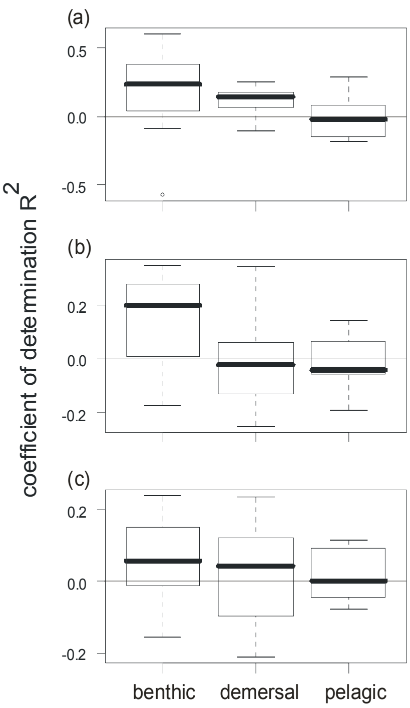


Figure 2.2.3. Box plots with coefficients of determination R^2 by species for best fitting models explaining survey indices by environmental (wind) conditions and survey design: (a) average survey density; (b) CV of average survey density and (c) individual mean weight. Results are grouped by habitat type: benthic, demersal (near sea floor) and pelagic.

References

Poulard, J.-C. & Trenkel, V. M. (submitted). Do survey design and wind conditions influence survey indices ? *Canadian Journal of Fisheries and Aquatic Sciences*.

2.3 Estimation of indicator values from surveys

J. Cotter
Cefas, Lowestoft, UK.

Introduction

When using a survey to estimate average values of an indicator over a stock or a sampling stratum of a survey, the statistical method of estimation used can significantly affect the computed results. In this regard, indicators of quantity and attributes present slight differences in theory. This is most easily seen with simple numerical examples.

Indicators of quantities of fish

Firstly, consider indicators of abundance (or weight) and suppose for simplicity that there are just three fishing stations in the survey area, yielding 1, 3, and 0 fish. The mean density per station over the survey area is $4/3 = 1.33$. This estimate is responsive both to the number of zero-yielding stations and to the densities of fish observed when fish are found to be present. On the other hand, the domain occupied by the stock implicitly includes only the two non-zero values, so that the mean density per station over the stock domain is $4/2 = 2$. This different estimate is only responsive to the densities of fish when found to be present, while reduction of stock abundance is estimated by geographic contraction of the stock domain. Of course, the catching of zero fish does not confirm that fish are absent, so the estimated stock domain is subject to sampling error as for the estimated mean density. We could distinguish the two estimates by calling them the *survey mean density per station* and the *stock mean density per station*. Indicators of the area occupied by a stock are presented in the section on spatial indicators.

Indicators of attributes of the fish

Next, consider indicators based on measured attributes of the individual fish caught, e.g. their lengths, and suppose for simplicity, using the same example, that the single fish at the first station was 50 cm, and the three fish at the second station were 10, 12, and 15 cm for which the station average is 12.33 cm. Including the third station where zero fish were caught makes no sense when averaging attributes, so the 'survey mean length' is unimportant. However, we can estimate a *stock mean length* as (1) a mean for the fish or (2) a mean for the stations, i.e.

1. as $(50+10+12+15)/4 = 21.75$, or
2. as $(50+12.33)/2 = 31.17$.

Note that the first estimate is weighted towards values observed at the station yielding most fish; it is a ratio estimator because it uses the number of fish at a station as a covariate of length to improve the precision of the estimate. The second estimate gives equal weight to the average value at each station without regard to how many fish were caught (given that at least one was caught). Choice between the two estimators may depend on the degree of within-haul correlation (Pennington and Vølstad 1994). A low degree, i.e. a good mix of lengths (in this example) at each station, suggests that the first estimator will be best because more fish implies more information about the stock. A high degree, i.e. long fish predominate at some stations, short at others, suggests the second because, by contrast, more stations implies more information about the stock. We could distinguish (1) and (2) as the *ratio estimator*, and the *station estimator*, respectively, of stock mean length.

A complication arises when attributes such as length or weight are not measured on all individuals caught by a survey but only on samples of the catch at some stations. It is assumed that the fish in the samples are chosen randomly from the fish in the catch. The best estimate of the average length or weight (etc.) at the station comes directly from the catch sample without raising. However, ratio estimates require all fish caught at each station to be included in the calculations. Consequently, catch sampling may rule out the use of ratio estimates, or it may require some sort of approximate fix, e.g. by raising the frequency distribution in the sample to one for the catch. Of course, the statistical properties of the fix may be poor and difficult to establish leading possibly to bias of the station mean, or a downward bias of the variance causing excessive confidence in the mean.

Standard errors

Standard errors for the ratio estimator are available from (Pennington and Vølstad 1994) or (Thompson 1992). Standard errors for the mean of haul results weighted equally depend on whether the stations can be considered as random sampling locations over the stock distribution. If 'yes', the standard deviation uses the usual formula for simple random sampling. If 'no', a spatial surface could be fitted to the haul results and the standard error of the integral derived from the model. Geostatistical estimates may also be feasible using a variogram, and bootstrapping (Beare *et al.* 2002) is another possibility. Standard errors are more complicated when catch sampling has taken place but an analytic approach is offered by Cotter (1998).

For a review of the wide range of estimation methods including variances used for fish surveys, see section 2.6 in Anonymous (2004).

References

Anonymous (2004). Report of the workshop on survey design and analysis (WKSAD). International Council for the exploration of the sea report ICES CM 2004/B:07, Ref. D,G. 72pp.

Beare, D., Castro, J., Cotter, J., van Keeken, O., Kell, L., Laurec, A.*, et al.* (2002) Evaluation of research surveys in relation to management advice (EVARES). Final report. *DGXIV Fisheries, European Commission, Brussels.* **FISH/2001/02 - Lot 1**. Available from john.cotter@cefas.co.uk

Cotter, A.J.R. (1998) Method for estimating variability due to sampling of catches on a trawl survey. *Canadian Journal of Fisheries & Aquatic Sciences* **55**, 1607-1617.

Halliday, R.G. and Mohn, R. (2001) Proceedings of the Fisheries Management Studies working group, 8-11 January 2001., Canadian Science Advisory Secretariat, Proceedings Series.

Jennings, S. (2005) Indicators to support an ecosystem approach to fisheries. *Fish and Fisheries* **6**, 212-232.

Pennington, M. and Vølstad, J.H. (1994) Assessing the effect of intra haul correlation and variable density on estimates of population characteristics from marine surveys. *Biometrics* **50**, 725-732.

Thompson, S.K. (1992) *Sampling*, John Wiley & Sons, Inc., New York

3. Biological indicators

3.1 Introduction

Biological indicators are available to measure most aspects of the health of a fish stock in addition to its numerical abundance. These include growth, age composition, fecundity, recruitment, and total mortality. Such indicators are 'state' indicators in the 'pressure-state-response' (PSR) system for classifying indicators (Jennings 2005). A selection of commonly considered biological indicators is presented here according to a format given by Halliday and Mohn (2001) for which see Appendix 1 of this report. The advice given for each indicator comes from discussions among the participants in the FISBOAT project. It is not intended to replace careful consideration of the relevance and value of each indicator to a particular stock.

3.2 Intrinsic population growth rate, r

Description

The slope of log total abundance against time.

Stock attribute

Numerical abundance summed over all ages.

Derivation

The indicator is Lotka's intrinsic population growth rate, *r*. It has been suggested for fish populations by (Quinn and Szarzi 1993).

Reference points

Taking $r = 0$ as the target reference point assumes that, without any noticeable impact of fishing, the population would be stable in the long term, even though it varies from year to year.

Interpretability

If *r* is significantly lower (respectively higher) than 0, the population is decreasing (increasing). The expected and undesired effect of fishing is to decrease *r*, although many other factors might have the same effect. A long term decline in *r* suggests that both recruitment and standing stock numbers are declining and implies that remedial action is necessary.

Measurability

r is readily and directly estimable from the time-series of abundance indices produced by a survey for all ages combined. The model is $N_t = N_{t-1}\exp(r)$. It can be estimated by fitting the mixed model, $\log N_t = \beta_0 + rt + \omega(t) + e_t$ where $\omega(t) \sim Normal(0, \sigma^2)$ represents the year to year variance, and $e_t \sim Normal(0, \sigma_e^{\ 2})$ represents random error. The estimate of *r* will depend on the time-window chosen since the linear slope measured by *r* is merely an average of fluctuations over time. The fitted linear slope acts as a smoothing function which may give a more stable indication of stock abundance trends than the time-series of raw abundance indices. Deviations from linearity could affect measurability adversely.

Sensitivity

Since *r* is a measure for all ages combined it is likely to be affected by large recruitment pulses, particularly if numbers of adults are low.

Example

Not available.

References

Quinn, T.J. and Szarzi, N.J. (1993) *Determination of sustained yield in Alaska's recreational fisheries.* International symposium on management strategies for exploited fish populations., Alaska sea grant college program, University of Alaska, Fairbanks, Alaska.

3.3 Total mortality, Z

Description

The coefficient of total mortality averaged over a given age range, i.e. the average mortality rate between years $t-1$ and t of all individuals aged $a_{\min}$ to $a_{\max-1}$.

Stock attribute

Z measures mortality due to fishing and natural causes. It is also affected by loss or gain of fish from a survey area as a result of a net migration occurring in relation to age.

Derivation

Z is the sum of natural and fishing mortalities: $Z = F + M$. It comes from the exponential model of mortality in population dynamics. The coefficient of total mortality, *Z*, is defined by ${dN}/{dt} = -ZN$ where *N* is abundance, *t* is time, and *Z* is conventionally regarded as positive. This solves to $N_t = N_0 \exp(-Zt)$ from which (i) $\log N_t = \log N_0 - Zt$, or (ii) $Z = -\log(N_t / N_{t-1})$.

Reference points

Z during a period of acceptable fishing mortality.

Interpretability

Z has been suggested as a robust indicator for exploited populations (Die and Caddy 1997). Different *Z* over different age ranges could be caused by less than full selectivity of the survey gear for some ages, or by migrations related to age particularly if the survey only covers part of the known range of the stock (Cotter *et al.* 2007). Interpretation of *Z* requires that *M* be assumed constant if, as is usually the case, it is not known accurately.

Measurability

Z may be estimated by fitting linear regressions to log abundance indices by year-class over age, equation (i) above, the so-called year-class curve method (Cotter *et al.* 2007). Removal of the youngest, and possibly the oldest ages may be necessary to find a satisfactory linear fit. Standard errors are available from the linear modelling. Alternatively, *Z* may be estimated separately for each (age, year) to (age+1, year+1) using equation (ii) above (Beare *et al.* 2002), and averaged over consecutive ages.

Sensitivity

Changes in *Z* over time are only likely to be discerned by surveys when commercial fishing effort changes substantially (Cotter, 2001). Changes in *Z* regionally resulting from net migration from one region to another as the fish grow older can be detected for plaice (Cotter *et al.* 2007). It is seldom possible to discern effects of changing *M* on *Z* except for year classes that are not susceptible to the fishery.

Example

See references.

References

Beare, D., Castro, J., Cotter, J., van Keeken, O., Kell, L., Laurec, A.*, et al.* (2002) Evaluation of research surveys in relation to management advice (EVARES). Final report. DGXIV Fisheries, European Commission, Brussels. *FISH/2001/02 - Lot 1*. Available from john.cotter@cefas.co.uk

Cotter, A.J.R. (2001) Intercalibration of North Sea International Bottom Trawl Surveys by fitting year-class curves. *ICES Journal of Marine Science* **58**, 622-632 [Erratum, Ibid. 58:1340].

Cotter, A.J.R., Mesnil, B. and Piet, G. (2007) Estimating stock parameters from trawl cpue-at-age series using year-class curves. *ICES Journal of Marine Science* **64**, 234-247.

Die, D.J. and Caddy, J.F. (1997) Sustainable yield indicators from biomass: are there appropriate reference points for use in tropical fisheries? *Fisheries Research* **32**, 69-79.

3.4 Numbers-at-length, numbers-at-age: NaL, NaA

Description

The length or age frequency distribution.

Stock attribute

Age structure and growth of the population. Relative numerical strengths of different year classes.

Derivation

Frequency distributions may be estimated for individual stations, strata, or for the survey as a whole.

Reference points

A sustainable stock should have a reasonable proportion of larger or older individuals capable of breeding, as well as to allow the commercial fishery to sustain itself when recruitment is poor. Protection of size or age classes just in the breeding category may not be sufficient if young fish tend not to be successful at breeding, or if the stock is poorly nourished such that individuals may not breed even though they have reached an age when they could.

Interpretability

A stock lacking large or old fish is likely to be over-fished and under-productive economically. However, similar effects could also occur due to predation or disease for example. Knowing indices of the abundance of predator-, and perhaps prey-species could be useful for interpretation of trends in frequency distributions. Increasing proportions of large or old fish implies better survival with age and may signal recovery of a stock.

Measurability

- NaA is harder to measure than NaL since otoliths have to be read.
- In general, RV surveys use gear with very small mesh so that most size and age groups will be fully selected. However, allowances must be made to frequency distributions if this cannot safely be assumed. Variability of the period when young fish settle to the bottom (and become vulnerable to a demersal trawl) relative to the survey period could affect estimated frequency distributions, as could the escape of large, fast-swimming individuals.
- The following comments follow from the note on estimation in the preliminary section of this manual. An age- or length-frequency distribution compiled from all fish of a species caught on a survey (or in a sampling stratum) will be most influenced by the frequency distributions prevailing at the stations where most fish were caught. [This estimate is to frequency distributions as the ratio estimate is to means.] This estimate would be preferred if the highest yielding stations are thought to provide the best indication of the frequency distribution for the whole stock.
- The alternative is to estimate a probability density for numbers-at-length or -at-age at each station where fish were caught, i.e. a histogram scaled to integrate to unity. Then the probability density for the stock (or a stratum) is obtained by averaging the station densities for each length or age class. [This estimate is to frequency distributions as the

station estimate is to means.] This estimate would be preferred if length frequencies differ noticeably between stations and no single station is thought to be more representative of the stock than another.

Sensitivity

In general, length and age composition of a stock change rapidly in response to fishing pressure, then numbers of older, larger fish remain low. Both indices are affected by pulses of recruitment, and this may be a more influential factor than slight changes of fishing when fishing pressure is high.

Example

Not available

References

3.5 Spawning Stock in number: SSN

Description

Mean catch in number of mature fish per tow (or per standard time or distance unit) realised during a survey.

Stock attribute

Abundance of fish with potential to breed. Sustainability of the stock.

Derivation

For trawl surveys, SSN is probably most simply computed in the same way as the usual survey abundance index but with numbers-mature at each fishing station substituted for numbers, or numbers-at-age.

Egg surveys directly provide an index of spawning stock biomass (SSB). If appropriate corrections are applied for mortality to the estimates of egg abundance, and if estimates of daily specific fecundity are available, then the index can be taken as an estimate of absolute SSB. Otherwise, it will only estimate relative SSB. SSB is converted to SSN by dividing by the mean weight of the mature fish in the population. The latter parameter, as for the fecundity parameters, is obtained by averaging the estimates from the individual fishing hauls, using weighting factors proportional to the egg abundance divided by the mean weight of the anchovies in the haul.

Reference points

Possibly, lowest historically observed estimates known to sustain a satisfactory recruitment.

Interpretability

- The index will probably be most closely related to numbers of the younger age classes that are mature since their abundance will usually be much higher than that of larger, older fish, depending on rates of total mortality, and annual recruitments.
- Changes in SSN are likely to be due to fishing but could also be caused by natural events.
- Further understanding of the health of a stock may be acquired by looking at indices for first-time and for replicate spawners separately: stocks subject to high *F* or high *M* generally show higher proportions of first-time spawners.

Measurability

Accurate assessment of maturity is crucial (Kjesbu *et al.* 2003; Murua *et al.* 2003) particularly as the proportion mature is likely to be most hard to estimate accurately for the most numerous age classes just coming into maturity. See also www.ices.dk/datacentre/datras/NSIBTSmanualRevVIIdraft.pdf. Careful standardisation of the maturity assessments across years and across survey crews is essential to avoid step changes in time or space purely as a result of inconsistent technique. Objective histological methods can be helpful for quality control of maturity assessments based on external morphology. The seasonal timing of the survey is also very important. Summer surveys can be especially poor for estimating maturity of species that spawn in winter or spring unless histological criteria are used to hindcast or predict maturation in the previous or next spawning season.

Sensitivity

SSN indices are likely to reflect changes in the age composition and reproductive potential of a stock over time, though precision will depend on standardisation of the techniques of maturity staging.

Example

Not available.

References :

Kjesbu, O.S., Hunter, J.R. and Witthames, P.R. (2003) Report of the working group on modern approaches to assess maturity and fecundity of warm- and cold-water fish and squids. *Institute of Marine Research*, Bergen, Norway. **Fisken og havet, 12**, 140 pp.

Murua, H., Kraus, G., Saborido-Rey, F., Witthames, P.R., Thorsen, A. and Junquera, S. (2003) Procedures to estimate fecundity of marine fish species in relation to their reproductive strategy. *Journal of northwest Atlantic Fisheries Science* **33**, 33-54.

3.6 Length statistics: Lbar, L25, L50, L75

Description

Mean (Lbar) or percentiles (L25, etc) of fish lengths found in survey catches.

Stock attribute

Growth, length frequency distribution of stock, recruitment success annually.

Derivation

Descriptive statistics for the length frequency distribution.

Reference points

Options are:

- Length at 50% maturity.
- Lengths associated with biological events such as migrations.
- Lengths associated with the fishery, e.g. length at 50% selectivity, minimum landing size (MLS).
- Historic values when the stock was deemed to be at a satisfactory level.

Interpretability

- In general, high fishing intensity reduces the relative abundance of large compared to small fish for two reasons: (a) large individuals are mostly older and therefore have had more exposure to fishing gear, and (b) because trawlers tend to catch large fish more effectively than small.
- The L25, L50, and L75 percentiles characterise the smaller, middle sized, and older fish, respectively, and are therefore expected to respond differently to recruitment pulses, growth factors, and to changes in abundance and spawning stock biomass.
- A short selection of references relevant to interpretation of fish length are by Ault et al. (2005), Kvamme and Froysa (2004), Deriso and Parma (1988), Piet and Jennings (2004), Jennings et al. (1999), and Rochet and Trenkel (2003).
- Lbar for mixtures of species has been shown to decrease in exploited communities (Jennings et al. 1999).

Measurability

Please refer to the note on estimation in the Introduction to Biological indicators. Care should be taken to standardise the bodily extremities to be used to measure each species, particularly across national surveys, and across years (Beckett 1983). Estimation of Lbar with omission of fish shorter than a given length, e.g. the length best separating 0 and 1-groups, or the MLS, is likely to improve its precision considerably when annual recruitments are very variable.

Sensitivity

Lbar is most influenced by the smallest, youngest year classes included in the estimate because they are usually the most numerous. Use of a ratio estimator, rather than a station estimator (see section on Estimation of indicators from surveys, above) would enhance this effect. It follows that entry of a new year class into the estimate could alter Lbar over only one year if that year class had been affected by a step-change of conditions affecting growth or abundance. However, the full effect on Lbar would not be manifest until all living ages had experienced the change. L25 is affected by the number and size of young fish; thus low recruitment of young, small fish would be expected to cause L25 to increase in the same year, and vice versa. L50 would be expected to respond in a similar way to Lbar. Fish longer than L75 are likely to belong to several age classes. Therefore, L75 is likely to decrease gradually over years in response to selective removal of predominantly large individuals by fishing faster than they are replaced by growth.

Example

Fig. 3.6.1 illustrates length-based indicator series derived from the EVHOE demersal trawl survey of the Bay of Biscay, one for anchovy and one for hake. No trends are evident in the Lbar or L quantile series for either species but, curiously, a linear decline in length variance could be fitted to the variance of length for hake. The explanation for this is not known.

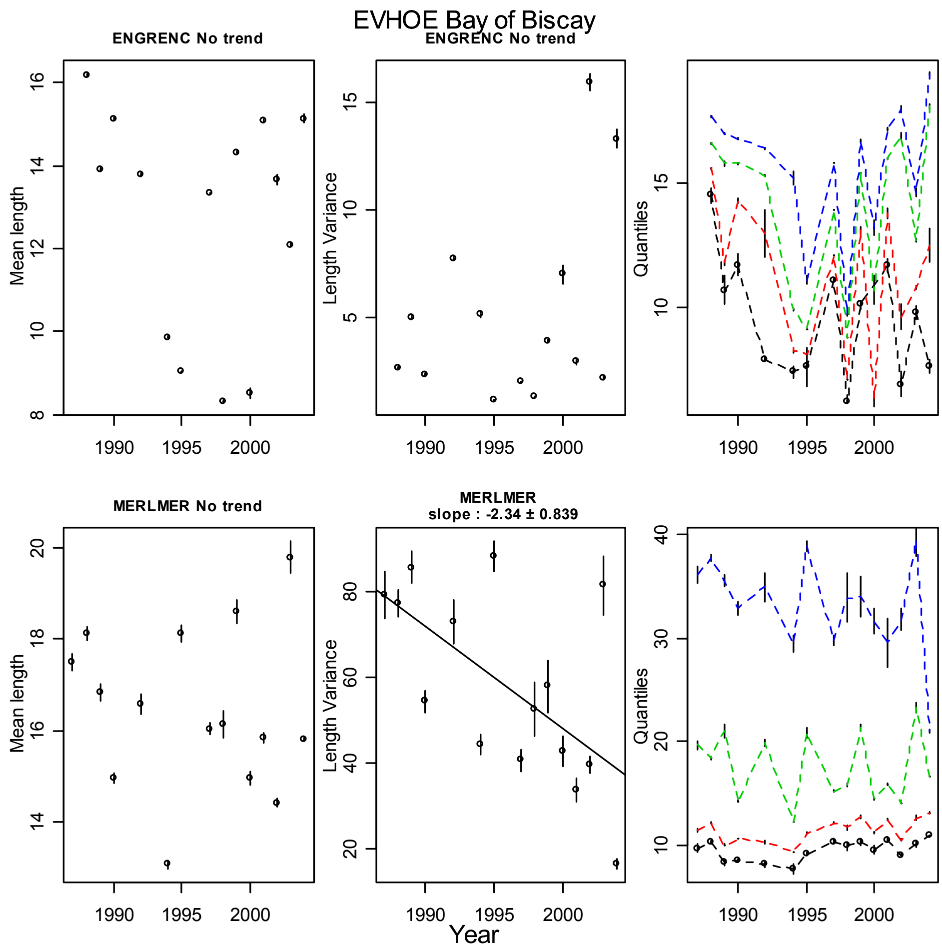


Figure 3.6.1: Time series of indicators for Bay of Biscay anchovy (code= ENGRENC) and hake (code= MERLMER) with long term trends. Data from the IFREMER EVHOE demersal trawl survey of the Bay of Biscay.

References

Ault, J. S., Smith, S. G., and Bohnsack, J. A. 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES Journal of Marine Science, 62: 417-423.

Beckett, J.S. 1983. Standards used for the sampling of commercial catch under ICNAF/NAFO. In *Sampling commercial catches of marine fish and invertebrates.* Ed. Doubleday, W.G. and Rivard, D. Canadian Special Publications in Fisheries and Aquatic Science, 66: 251-254.

Deriso, R.B. & A.M. Parma. 1988. Dynamics of age and size for a stochastic population model. Can. J. Fish. Aquat. Sci. 45: 1054-1068.

Kvamme, C. & K.G. Froysa. 2004. Assessing the effects on stocks of selectivity changes in a fishery. Fish. Res. 69: 283-292.

Jennings, S., S.P.R. Greenstreet & J.D. Reynolds. 1999 Structural change in an exploited fish community: a consequence of differential fishing effects on species with constrasting life histories. J. Anim. Ecol. 68: 617-627.

Piet, G. J., and Jennings, S. 2005. Response of potential fish community indicators to fishing. ICES Journal of Marine Science, 62: 214-225.

Rochet, M.J. & V. M. Trenkel. 2003. Which community indicators can measure the impact of fishing? A review and proposals. Can. J. Fish. Aquat. Sci. 60: 86–99.

3.7 Survey catch weight: W

Description

The total weight of one (or more) species caught on a survey with constant fishing effort.

Stock attribute

Stock biomass, size composition.

Derivation

Many surveys routinely weigh catches by species. If such data are not available, weights may be estimated for each species using the allometric formula: $W = \sum_l n_l a l^b$ where l is length class, n_l is the numbers caught during the entire survey in that length class, and a and b are constants for the species, the first being a scaling factor for units, the second being a factor relating to change of shape with increasing size. a and b are determined by plotting $\log(W_l / n_l)$ against $\log l$ using a sample of weighed and measured fish – more details are given for the Condition index, below. The estimates of a and b should be up-to-date. An even simpler weight index derived from length data may be adequate for some species; i.e. assume that $W \propto \sum_l n_l l^3$. In this formula, $b = 3$ implies no change of shape with size.

Reference points

A possibility is to use historical estimates of total survey weight when the stock was considered to display a satisfactory age composition, i.e. one having sufficient mature age groups present to provide resilience to occasional poor recruitment or temporary, heavy fishing pressure.

Interpretability

- Condition of the fish (weight/length) may be important for interpreting W but only if fish weights are measured directly, rather than being estimated from lengths.
- The quality of information that W provides about the weight of fish in the stock depends partly on the relative catchabilities of different sized fish by the survey; e.g. catchability of large individuals of strongly swimming species may be low on surveys with short tows.
- Calculation of W might omit fish shorter than the minimum landing size (MLS) for the fishery so that W becomes of direct relevance to the legal yield of the fishery. It might then be serviceable as guidance for a harvest control rule.
- Standardising W as a total weight for the entire survey area is proposed because high yielding stations then have most influence on the result. This could make the index of immediate relevance to commercial fishers who tend to target such localities. On the other hand, comparison of the survey mean weight per station with the stock mean weight per station (see section on Estimation of indicators from surveys, above) while also noting any changes in the estimated domain of the stock may be more relevant for ecological studies.
- W could also be translated into a mean individual weight. This would be akin to a condition index (see C below) but without allowing for body length. The ratio and the station estimator of the stock mean individual weight might then have different

interpretations due to the different weightings given to stations yielding different numbers of fish (see section on Estimation of indicators from surveys, above). The ratio estimator could be more relevant for a fishery that focuses on aggregations of fish, while the station estimator could be more relevant for regional, ecological studies.

Measurability

W can be expected to vary a lot for migratory stocks that are not satisfactorily enclosed by the survey domain, or whose distribution in relation to fishing stations varies with abundance. Variation from year to year of stations fished, or in the duration of tows should be allowed for by raising to a standard level of effort equivalent to the complete, standard survey while recognising that bias might have been caused by omission of certain stations. The season of the survey should be kept constant because *W* will be affected by seasonal changes in availability of food, and by enlargement of gonads for breeding. A constant selectivity of the survey trawl and fishing method with respect to size is crucial. For some ground-loving species, e.g. certain flatfish, distance towed over the ground may be more relevant as a standardising measure for survey CPUE than the duration of the tow in minutes.

Sensitivity

A merit of *W* is that it is less influenced by varying recruitment from year to year than raw length indices such as Lbar because young fish, although very numerous, are very small, so their abundance and weight tend to cancel in the overall index. *W* would be influenced by changing abundance and average condition of individuals.

Example

Heessen et al. (1997) state that the International Bottom Trawl Survey of the North Sea and Skagerrak covered the whole of the North Sea, Skagerrak and Kattegat from 1974 onwards. Catch-at-length data for quarter 1 were obtained for cod from the ICES secretariat and average catch weight per fishing station per rectangle was estimated using a ratio estimator for each of the 138 rectangles in the standard area used to estimate the abundance index for cod. Weight indices were estimated using $W \propto \sum_{l} n_{l.} l^3$. The results were summed over all rectangles to estimate *W* for the whole survey. The time series from 1974 to 2005 is shown in fig. 3.7.1, below, together with the numbers of rectangles where positive catches of cod were taken. The obvious downward trend in *W* from approximately 1980 is consistent with the well known decline in the North Sea cod stock in recent years. Variable results in the 1970s may possibly have been due to inconsistent survey practices in the early days of this survey. However, no attempt has been made here to remove inconsistencies at any time in the series. See Heessen et al. (1997) for details of inconsistencies over the period.

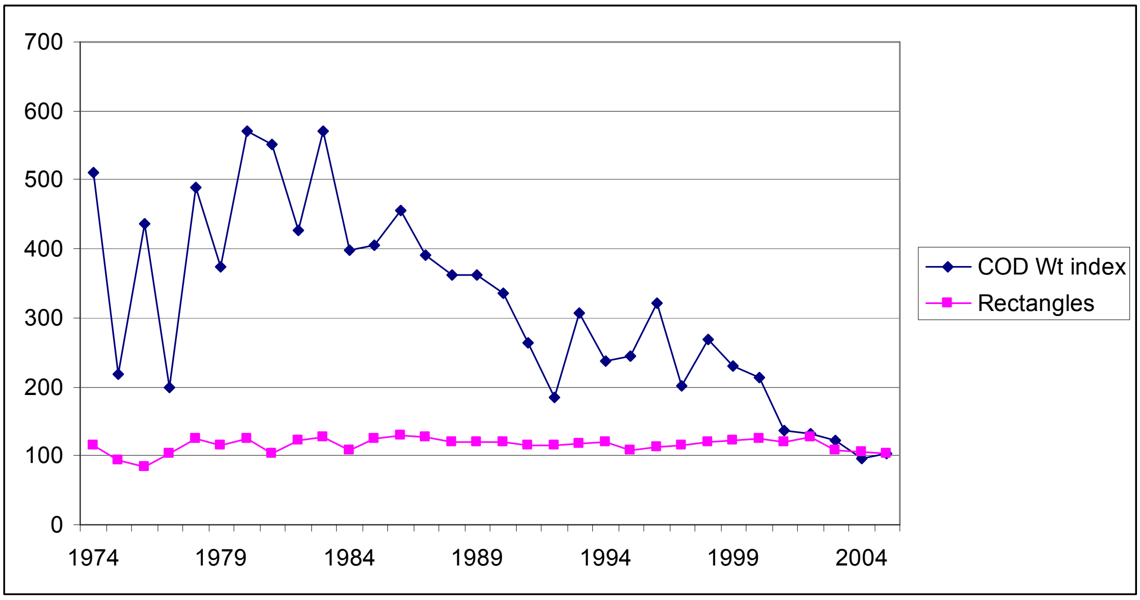


Figure 3.7.1. International Bottom Trawl Survey, quarter 1: cod weight indices, W, computed from numbers-at-length by ICES rectangle assuming isometric growth. Data supplied to FISBOAT by ICES secretariat.

References

Heessen, H.J.L. (1997) The International Bottom Trawl Survey in the North Sea, the Skagerrak and Kattegat. *International Council for the Exploration of the Sea.* ICES CM 1997/Y:31, 25 pp.

3.8 Condition: C

Description

Condition refers to average body weight for a given body size. It is often normalised as Fulton's condition index (Anderson and Neumann 1996), W/kL^3 , where k is chosen for appropriate scaling.

Stock attribute

Nutritional status of individuals, reproductive fitness.

Derivation

C may be estimated by a in the allometric equation $W = \sum_l n_l a l^b$ where l is length class, n_l is the numbers caught in that length class, a measures condition as well as serving to scale the units of measurement, and b represents the changing shape of the species with increasing length. Partitioning this equation by length gives $W_l = n_l a l^b$. Taking logs and substituting C for a gives $\log(W_l/n_l) = \log C + b \log l$. Therefore a regression of log (average weight per individual at length l) on $\log l$ will allow estimation of $\log C$ as the intercept and b as the slope. In many studies, b is standardised at 3, as for the Fulton index, implying growth without change of shape.

Reference points

Comparison with historic data might reveal when problems are occurring.

Interpretability

- Low condition implies too much competition for available food. This in turn implies that some mature individuals may not mature reproductively for the coming spawning season, or that their fecundity may be reduced by follicular atresia (Thoresen *et al.* 2006; Kennedy *et al.* in press). Low condition can also increase the age of first maturity, and possibly increase natural mortality of post spawning individuals.
- Condition varies in males and females according to season, for example after a winter fast, and over the reproductive cycle especially for species with a capital spawning strategy (Stearns 1992). See also (Lambert *et al.* 2003).
- Low condition implies reduced incomes for fishers because they will not attain high prices for the fish they catch.

Measurability

The regression could be carried out at each station and the estimated C averaged over all stations; i.e. a station-based estimator, giving equal weight to each station yielding fish. Alternatively, the regression could be carried out for all stations at once; this is a regression analogue of a ratio estimator because the results at each station are automatically weighted in relation to the number of fish caught at each length. The weights assigned to each station will therefore vary from length class to length class. Please refer to the section on Estimationof indicators from surveys, above.

Standard errors for a and b estimated from a regression are likely to over-estimate confidence in the result if the individual fish in the sample were not independently and randomly selected from the stock, as they would not be if taken with a trawl survey for example. This should be kept in mind when making comparisons from year to year. A spatial model could be a good way to summarise station-by-station results because condition may vary with location, particularly with latitude if accompanied by changing temperatures. Kriging is another possibility.

Sensitivity

Not known.

Example

Not available.

References

Anderson, R.O. and Neumann, R.M. (1996) Length, weight, and associated structural indices. In: *Fisheries Techniques*. ed Murphy, B.R. and Willis, D.W. 2nd edn, American Fisheries Society**:** 447-482.

Kennedy, J., Witthames, P.R. and Nash, R.D.M. (in press) The concept of fecundity regulation in plaice Pleuronectes platessa L. tested on three Irish Sea spawning populations. *Canadian Journal of Fisheries & Aquatic Sciences*.

Lambert, Y., Yaragina, N.A., Kraus, G., Marteinsdottir, G. and Wright, P.J. (2003) Using environmental and biological indices as proxies for egg and larval production of marine fish. *Journal of northwest Atlantic Fishery Science* **33**, 115-159.

Thoresen, A., Marshall, C.T. and Kjesbu, O.S. (2006) Comparison of various potential fecundity models for north-east Arctic cod Gadus morhua L. using oocyte diameter as a standardizing factor. *Journal of Fish Biology* **69**, 1709-1730.

Stearns, S.C. (1992) *The evolution of life histories.*, Oxford University Press, Oxford, UK

3.9 Gonadosomatic index: GSI

Description

The gonadosomatic index (GSI) is the ratio of gonad weight to body weight.

Stock attribute

Nutritional status, reproductive fitness.

Derivation

Average gonad weight/body weight, or gonad weight/length^3.

Reference points

Comparison with historic data might reveal when problems are occurring with regard to stock reproductive potential.

Interpretability

- GSI depends on fish size and will obviously be low for immature individuals. It will also increase rapidly towards the start of the spawning season. It is likely to be highest after ovulation when the first batch of eggs is ready for spawning, and lowest just after spawning when fish have not only lost their reproductive material but are also likely to show low bodily condition.
- GSI may vary with location, particularly with latitude if accompanied by changing temperatures. Spatial modelling or Kriging could be helpful for seeing these effects.
- A low GSI in fish of mature age at the start of the spawning season may imply skipped spawning (Rideout *et al.* 2005), suggesting too much competition for available food, as well as low reproductive success in the coming spawning season, either through lack of fertile adults or through reduced egg production.

Measurability

Surveys intended to estimate GSI should be timed to coincide with the onset of spawning, or, if that is not possible, to avoid the post-spawning period. They should occur in the same season each year. At least a sample of fish from every fishing station should be measured individually for length, body weight, and gonad weight. Stratification by depth bands may be helpful for estimation if GSI is related to depth, as it is for some species (Rijnsdorp 1989). A cut-off length for exclusion of immature individuals would save much pointless disection work on deck.

Weighing gonads requires that they be removed from the fish and weighed on a balance capable of resolving down to 1% of the gonad mass. Motion compensated balances capable of resolving to 0.1g are required for small fish such as sprat or anchovy weighed at sea. An advantage of GSI over maturity indices is that it does not require accurate maturity staging.

Sensitivity

The GSI is very sensitive to the maturity stage and the timing of the spawning cycle.

Example

Not available.

References

Rideout, R.M., Rose, G.A. and Burton, M.P.M. (2005) Skipped spawning in female iteroparous fishes. *Fish and Fisheries* **6**, 50-72.

Rijnsdorp, A.D. (1989) Maturation of male and female North Sea plaice (Pleuronectes platessa L.). *Journal du Conseil international Exploration du Mer* **46**, 35-51.

3.10 Length and age at maturity: LaM50, AaM50

Description

The length or age at which 50% of the individuals in a fish stock are estimated to have reached reproductive maturity.

Stock attribute

Reproductive capability, spawning stock biomass, nutritional status.

Derivation

The estimated proportion mature at each length is plotted against length and the length at which 50% of individuals are mature is the LaM50. Fitting a logistic or other function may be the best way to estimate this statistic using reaction norms (Heino *et al.* 2002). If otoliths are removed at the same time as gonads are examined, and subsequently read without breaking the link with the observation of maturity, the AaM50 can be estimated similarly with a plot over age.

Reference points

Reaction norms are used to describe the phenotypic range of LaM50 and AaM50, and to identify evolutionary selection of alleles favouring early maturation. They may already be in the genetic structure of the population, or they may be introduced by genetic drift and/or mutation (Heino *et al.* 2002).

Interpretability

- AaM50 has been found to decrease under the effect of fishing (Trippel 1995; Rochet et al. 2000). If individual growth remains similar under the impact of fishing, LaM50 will decrease in a similar manner. Compensatory growth might, to some degree, reduce the impact of fishing on the observed reduction in length at maturity but strong signals should still be detectable. The reaction norm method is used to investigate evidence for evolution of maturation stage with respect to size and age after several generations of high fishing mortality.
- LaM50 and AaM50 may both vary with location, particularly with latitude if accompanied by changing temperatures.

Measurability

Please refer to Measurability of Spawning stock number (SSN) for comments on estimating maturity and the timing of surveys. Since maturity staging requires that fish be opened and the gonads examined carefully, it is time-consuming on deck.

Sensitivity

Problems with achieving consistent maturity staging from year to year when using only external morphology may seriously reduce the sensitivity of LaM50 and AaM50 to fishery and environmental factors. Better sensitivity could be expected if histological examinations are carried out for each fish but this is obviously much more time-consuming.

Example

Not available.

References

Heino, M., Deickmann, U. and Godø, O.R. (2002) Measuring probabilistic reaction norms for age and size at maturation. *Evolution* **56**, 669-678.

IBTS w.g. (revision VII in draft in 2007). Manual for the international bottom trawl surveys. Appendices VII Finfish maturity key, and VIII Four stage maturity key for skates and rays (Rajidae). www.ices.dk/datacentre/datras/NSIBTSmanualRevVIIdraft.pdf

Rochet, M-J., P.A. Cornillon, R. Sabatier & D. Pontier. 2000 Comparative analysis of phylogenetic and fishing effects in life history patterns of Teleost fishes. Oikos 91: 255-270.

Trippel, E.A. 1995 Age at maturity as a stress indicator in fisheries. Bioscience 45: 759-771.

4. Spatial indicators

4.1 Introduction

Spatial indicators are statistics aimed at describing and summarizing the spatial distribution of populations, in terms of location, fish density or possibly an environmental variable e.g. depth. A list of 10 geostatistical indices are here proposed (Woillez et al., 2007) to characterise occupation, aggregation, location, dispersion, correlation and overlap. These notions are somewhat related, e.g., aggregation, dispersion and occupation, and formal relationships exist between indices. The centre of gravity of a population with a measure of dispersion around it had been proposed already (Swain and Sinclair 1994, Atkinson et al. 1997, Bez and Rivoirard 2001). The occupation and aggregation indices are not truly spatial in the sense that they are sensitive to the histogram and not to the spatial location of values. Various indices to characterise aggregation have been suggested (area coverage: Swain and Sinclair 1994, Gini index: Myers and Cadigan 1995, spatial selectivity index: Petitgas 1998) which all relate to the area coverage of highest values. But the spreading index is more general in the sense that the amount of zeroes do not affect this index. Therefore in the calculation of the spreading index the delineation of the data positive domain is not necessary. The spatial indices are useful in characterising the spatial organisation of the life cycle. It can be evidenced that young immature fish, young matures and older matures differ in some aspects of their spatial distributions, in particular for location, aggregation and dispersion (Woillez et al., 2007). The spatial indices have the potential to be used in a monitoring system so as to detect changes in the spatial distributions, which could be helpful in relating the spatial distribution properties of fish stocks to their dynamics, climate change or habitat conservation.

When selecting spatial indicators for this manual, care has been taken to avoid the problem of zero density values by excluding statistics that would depend on the inclusion or exclusion of zero density values according to their belonging to a more or less arbitrarily delineated domain. For instance, the mean of the density values within a given domain is not considered here, nor the variance or the Gini index of these values to measure their statistical dispersion. On the contrary, the contribution of the zero density values is zero in the statistics that have been selected here. For instance, the center of gravity, or mean location, of a population will depend on whether the density value at a sampled location is zero or not, but if it is zero, its numerical contribution to the center of gravity will be zero.

In particular the contribution of zero density values is zero in all statistics based on the individuals of the population: e.g. the mean location of the population, which is the mean location of the individuals that constitute this population (in such a case the statistics are weighted by the fish density). By contrast, the Positive Area looks at where the fish density values are strictly positive but does not depend on the level of these density values, that is, on each individual.

Some of the selected statistics (e.g. the center of gravity) would change if the fish density values were permuted between sampled locations (even assuming a regular sample grid). The other statistics would be unchanged: for instance the Positive Area measures the domain covered by the non-zero density values, not its shape, and it would be unchanged by permuting density values. Similarly the Spreading Area or the Equivalent Area will depend on the histogram of density values, not on their location (at least assuming a regular sample grid). As a consequence these statistics are not dependent on the large-scale spatial structure. However they do depend on the fine scale structure through the 'support', i.e. the surface in size and geometry (e.g. the trawled area) on which each fish density is measured.

Note: references for the following indicators are collected together at the end of section 4.

4.2 Centre of gravity: CG

Description

The center of gravity is the mean location of the population, that is, the mean of the location of its individuals (Bez, 1997).

Stock attribute

Mean location of the population.

Derivation

Let x be a point in two-dimensional space (short for the usual two-dimension notation (x, y)), and $z(x)$ be the density of population at location x. Then, the total abundance of the population (Q) is calculated from:

$$Q = \int z(x)\mathrm{d}x\,,$$

and the probability density function of the location $\underline{x}$ of a random individual is $z(x)/Q$. The centre of gravity (CG) is:

$$\mathrm{CG} = E(\underline{x}) = \int x\frac{z(x)}{Q}\mathrm{d}x = \frac{\int xz(x)\mathrm{d}x}{\int z(x)\mathrm{d}x}\,,$$

In practice, this statistic is estimated from the data through discrete summations over sample locations. In the case of irregular sampling, areas of influence around samples are used as weighting factors (**Figure 4.2.1**). Practically, from sample values z_i at locations x_i, with areas of influence s_i, we have:

$$\mathrm{CG} = \frac{\sum_{i=1}^{N} x_i s_i z_i}{\sum_{i=1}^{N} s_i z_i}\,,$$

The area of influence of a sample location is defined as the area made up of the points in space that are closer to this sample than to others. It can be evaluated by overlaying a very fine regular grid and counting the grid points closer to the sample. Known or supposed boundaries (e.g. land, a limit distance of influence from a sample location) of the sampled population may be used.

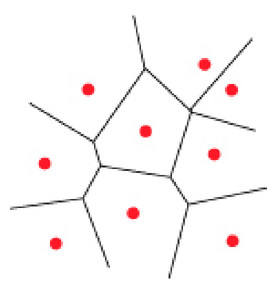


Figure 4.2.1. Areas of influence (delimited by the black lines) of sample points (in red).

Reference points

CG during a period of acceptable state of the population.

Interpretability

The CG indicates the mean location of the surveyed population. Note that fish may not be present at the CG location (which may be on land, e.g. within an island). Note also that part of the population may be not represented, when not covered geographically by the sampling. To check that movements of CG when following a series are not due to changes in the sampling design (e.g. due to bad weather), the CG of the sample locations (that is, unweighted by fish density) can also be produced.

Measurability

The estimated CG is sensitive to high fish density values. It may differ from the true unknown CG, particularly when high density values exist (whether sampled or not).

Sensitivity

Despite the possible difference between true and estimated CG, a shift during a series is likely to represent an actual shift, when it is gradual. On the other hand, an eccentric estimated CG requires a visual inspection of the fish density to detect the causes (e.g. unusual presence of high density values in some remote area, or disappearance of usually high values in some region; see the indicators Inertia and Number Of Patches).

Examples

Several authors have used the center of gravity, also referred to as the distributional centroïd or as the center of an ellipse to describe the distribution of a population or a life stage of a population such as Walleye pollock eggs and larvae (Kendall and Picquelle, 1990), Pacific hake larvae (Hollowed, 1992), cod off Newfoundland (Atkinson et al., 1997), yellowtail flounder off the Grand Bank (Brodie et al., 1998) or European hake in the Bay of Biscay, eggs and larvae (Alvarez et al., 2001) as well as fish at age (Woillez et al., 2007). For example, the CGs have been used to describe the distribution of the strong year classes on the Pacific hake late stage larvae and also the systematical shift towards the south east of cod off Newfoundland from 1987 to 1993. Distributions have been also described temporally along the season, e.g. for eggs and larvae of the European hake in the Bay of Biscay and of Walleye Pollock.

4.3 Inertia: I

Description

The inertia is the variance of the location of the individuals of the population, that is, the mean square distance between an individual fish and the centre of gravity of the population (Bez, 1997).

Stock attribute

It describes the dispersion of the population around its centre of gravity.

Derivation

With the notations used for CG, the inertia (I) is

$$\mathrm{I} = \mathrm{Var}(\underline{x}) = \frac{\int (x - \mathrm{CG})^2 z(x)\mathrm{d}x}{\int z(x)\mathrm{d}x}.$$

and is estimated as:

$$\mathrm{I} = \frac{\sum_{i=1}^{N} (x_i - CG)^2 s_i z_i}{\sum_{i=1}^{N} s_i z_i}.$$

Reference points

I during a period when the population was in an acceptable state.

Interpretability

The inertia I indicates how dispersed the population is around its center of gravity.

Measurability

I is sensitive to high density values. I is homogeneous to square nautical miles (in 2D). The square root of I, that is, the root mean square distance between individuals and their CG, may be preferred, being homogeneous to nautical miles.

Sensitivity

An increase in I, for instance, indicates a population more dispersed around its CG, i.e. high density values are more scattered. While the population is then scattered over a larger region, the actual area covered by the population may be smaller (see the different Area indicators).

Examples

Most of the authors cited for the CGs also described the studied population in terms of inertia, or they refer to the size of its graphical representation through an ellipse (see next section on anisotropy). Brodie et al. (1998) showed a drop in the area of the ellipse for the yellowtail flounder of the Grand bank in late 1980 after a period of stability. Atkinson et al. (1997) showed that a shift of the CG is accompanied by a decrease of the size of the ellipse, i.e. the inertia. For the European hake in the Bay of Biscay, concerning the eggs, the size of the ellipses increased somewhat from February to May in both directions, N-S and W-E (Alvarez et al., 2001), while the inertia increased with the age for fish stage (Woillez et al., 2007).

4.4 Anisotropy and isotropy: An, Is

Description

When the dispersion of the population around its center of gravity is the same along every direction, the spatial distribution is said to be isotropic. In general, the dispersion of a population around its centre of gravity is not identical in every direction of space: there is an anisotropy. The root mean square distance to the center of gravity is maximal along the first principal axis, and minimal along the second principal axis, orthogonal to the first one (in 2D). The anisotropy index is taken as the ratio (>= 1) between these distances, and the isotropy index as the inverse ratio (<= 1).

Stock attribute

Anisotropy measures the elongation of the spatial distribution of the population.

Derivation

In two dimensions, the total inertia of a population can be decomposed on its two principal axes, orthogonal to each other, explaining respectively the maximum and the minimum of the inertia. These two principal axes and their inertia can be obtained as the eigen vectors and values of a principal component analysis of the coordinates of the individuals of the population (i.e. the coordinates of the samples weighted by the fish densities) (Bez, 1997). The square root of the inertia along a given axis (or root mean square distance to CG) gives the standard deviation of the projection of the location of the population along that axis. These can be represented conveniently on a map with a cross depicting the two principal directions (**Figure** 4.4.1), or with an ellipse (with area proportional to the total inertia). The anisotropy index (>= 1) is the square root ratio between the maximum and the minimum of the inertia. Similarly, an index of isotropy can be defined as the inverse of anisotropy, ranging more conveniently from 0 to 1:

$$\text{Isotropy} = \sqrt{\frac{I\min}{I\max}} \text{ and } \text{Anisotropy} = \sqrt{\frac{I\max}{I\min}}.$$

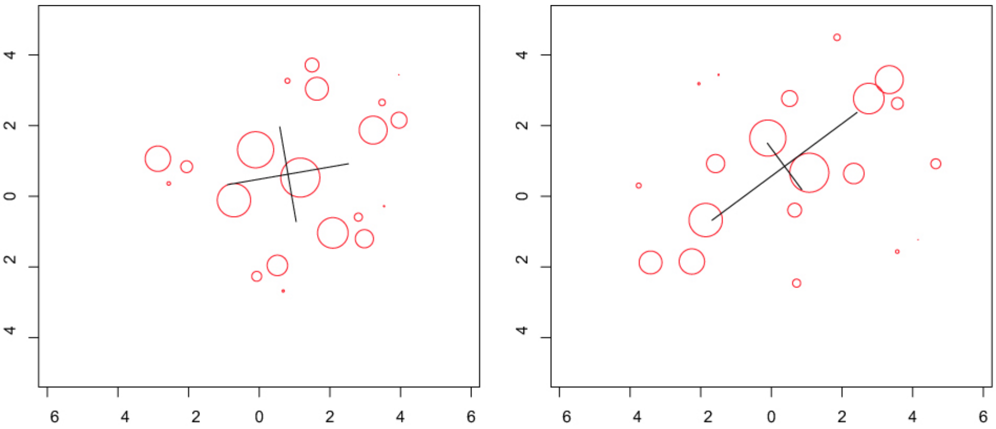


Figure 4.4.1. Two examples of spatially distributed data sets, with anisotropy being more marked in the second case. The black cross is located on the center of gravity, from which it represents the square root of inertia along the two principal directions.

Reference points

The indicators during a period of acceptable state of the population.

Interpretability

The anisotropy index gives roughly the elongation of the spatial population. It does not take into account the actual shape of the distribution, which may be different from being elliptical or may be constituted by different patches.

Measurability

The anisotropy and isotropy indices are equivalent, being the inverse of each other. However, since the anisotropy index is unbounded above 1, the isotropy index is more robust and may be more conveniently used, e.g. in correlation or regression analyses.

In case of isotropy, that is, when the anisotropy and isotropy indices approach 1, the principal axes, orthogonal to each other, become arbitrary.

Sensitivity

Sudden changes in anisotropy index may be due to the disappearance or, on the contrary, the appearance, of patches of fish in some areas.

Examples

Few authors have discussed the anisotropy even if it was available in the representation mode. In Alvarez et al., 2001, the direction of the principal axis of the hake egg distribution in the Bay of Biscay corresponds to that of the shelf break, i.e. NW-SE, along the whole sampling period. Woillez et al., 2007, completes the description for the fish stage, showing a preferential direction, more marked for age 0 and age 5+. The direction for age group 0 corresponds roughly to muddy sediment off Brittany. For ages 4 and 5+, the direction corresponds to the shelf edge, where older hake are mainly concentrated. For the intermediate ages, the population is still anisotropic, probably because of the general shape of the continental shelf, but the anisotropy is less marked.

4.5 Global index of collocation: GIC

Description

The global index of collocation looks at how geographically distinct two populations are by comparing the distance between their CGs and the mean distance between individual fish taken at random and independently from each population (Bez and Rivoirard, 2000).

Stock attribute

Overlap of two spatial populations.

Derivation

Let us consider two populations with densities $z_1(x)$ and $z_2(x)$ at point x, with ΔCG being the distance between their centers of gravity and I_{z_1} and I_{z_2}, their respective inertias (

Figure 4.5.1). Then the mean square distance between individuals taken at random and independently from each population is $\Delta CG^2 + I_{z_1} + I_{z_2}$. The global index of collocation (GIC) is:

$$GIC = 1 - \frac{\Delta CG^2}{\Delta CG^2 + I_{z_1} + I_{z_2}}$$

or 1 if $\Delta CG^2 = I_{Z_1} = I_{Z_2} = 0$. The spatial index ranges between 0, in the extreme case where each population is concentrated on a single but different location (inertia = 0), and 1, where the two CGs coincide.

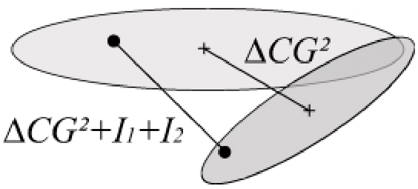


Figure 4.5.1. Collocation of two spatial populations, represented by two ellipses showing their center of gravity and their inertia, is measured with the GIC through specific distances.

Reference points

GIC between acceptable states of the populations.

Interpretability

Collocation is considered here in a global meaning, the populations being, grossly, in the same place. This is not to say that the two populations are present at the same locations. A spatial population that would be distributed all around a first one, with the same CG, would give a GIC equal to 1, even if not overlapping the first population.

Local overlapping between two populations would rather be addressed using the local index of collocation, that is, the noncentered correlation between their fish densities.

Measurability

Alternative indices, also between 0 and 1, may be sqrt(GIC) for collocation, or sqrt(1-GIC) for separation (ratio between distance between the CGs and distance between individuals from the two populations).

Sensitivity

Unusually high GIC requires inspection of the fish density data.

Examples

In Bez and Rivoirard, 2000, global and local collocation indices have been measured on pelagic species in the Bay of Biscay. Local collocation appears very small between mackerel and the other species (anchovy, sardine and horse mackerel). In Woillez et al., 2007, GIC is used to detect outliers in the age time series of the hake population in the Bay of Biscay. The year 2000 appears to be particular for age 0.

References

See the main references of the topics about the spatial indicators.

4.6 Number of spatial patches: NOP

Description

A spatial population of fish may be distributed into several spatial patches, with size much larger than a fish school. An algorithm has been written to identify patches (Woillez et al., 2007) by attributing each sample to the nearest patch, with respect to a maximal threshold distance to its CG. The Number of Patches then includes all patches that include more than a given part (e.g. 10%) of the total abundance.

Stock attribute

Patchiness.

Derivation

The algorithm starts from the value displaying the maximum density $z(x)$, and considers every other sample in decreasing order of density (**Figure** 2.6.1). The maximum value initiates the first patch (1). Then, the current sample value is attributed to the nearest patch, if the distance to its CG is smaller than the threshold distance *dlim* (2). Otherwise, the current sample value defines a new patch (3). Spatial patches whose abundance is >10% of overall abundance are retained (N). The summary index is then the number of patches.

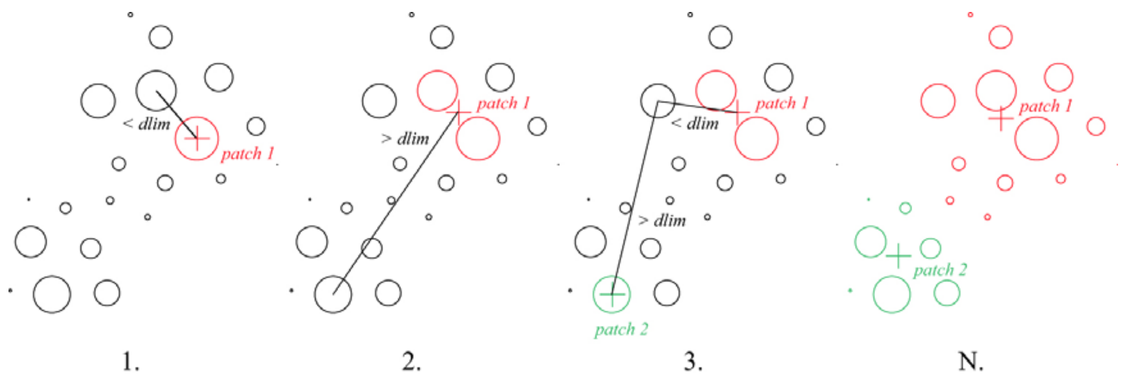


Figure 2.6.1. Main steps of the algorithm used to determine the number of patches of a spatial population.

Reference points

Number of patches during a period of acceptable state of the population.

Interpretability

The identification of patches is dependent on the threshold distance, typically some fraction of the diameter of the sampled domain, chosen by the user.

Measurability

The Number of patches is very sensitive to the location of the highest fish density values, but this makes sense.

Sensitivity

In a series, the location of patches is likely to present some stability. Hence a change in the number of patches is likely to reveal the disappearance, or the appearance of fish in some areas.

Examples

In Woillez et al., 2007, the Number of Patches have been illustrated on the hake population in the Bay of Biscay. According to ages, it increases slightly up to age 3 then decreases. Disappearance of patches has been observed and localised for age 0 hake, in particular for the year 2000.

References

See the main references of the topics about the spatial indicators.

4.7 Positive area: PA

Description

The positive area is the measure, in square nautical miles, of the space occupied by fish densities strictly above zero (Woillez et al., 2007).

Stock attribute

Area of presence, in square nautical miles, occupied by the stock, even with a low density.

Derivation

The positive area is estimated from data as the sum of the areas of influence around samples where there are fish densities >0 (Fig. 4.7.1):

$$\mathrm{PA} = \sum_i s_i 1_{z_i > 0} .$$

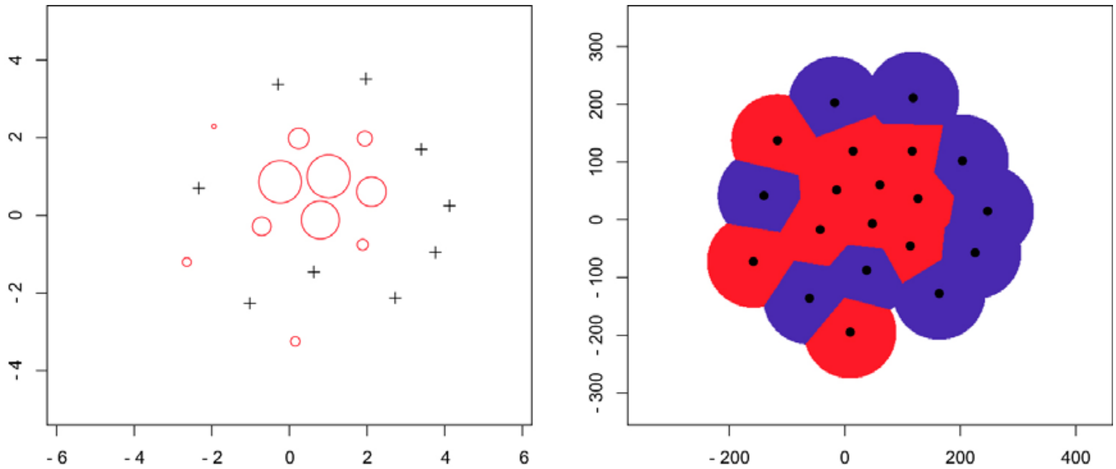


Figure 4.7.1. Bubble plot of the sample values and corresponding positive area shaded in red (with a limit to the area of influence of each sample).

Reference points

Positive Area during a period of acceptable state of the population.

Interpretability

The positive area measures the area of effective presence, in square nautical miles. It does not include zero density areas possibly existing between positive density areas, and it may correspond to a small fraction of the geographical envelope of fish presence, in particular when the dispersion (inertia) is high.

Measurability

Zero values of density make no contribution to the positive area. However, the positive area is sensitive to the low values of density, because the contribution of a small density value is the same as that of a high density value.

Sensitivity

Changes in the Positive Area may reveal changes in the way the population occupies the space, including the small fish density values that are usually numerous even though they contribute poorly to the global abundance.

Examples

In Woillez et al., 2007, Positive Area of the hake population in the Bay of Biscay was relatively stable until age 3 then dropped. It was shown also that whereas Positive Area decreased with age, inertia increased with age, the older hake occupying a smaller but more dispersed area.

References

See the main references of the topics about the spatial indicators.

4.8 Spreading area: SA

Description

The spreading area is a measure, in square nautical miles, of how the population is distributed in space, taking into account the variations in fish density (Woillez et al., 2007).

Stock attribute

A measure of the area occupied by the stock, based on how the abundance is spreading in space.

Derivation

Let T be the cumulative area occupied by the density values, ranked in decreasing order; let Q(T) be the corresponding cumulative abundance, and Q be the overall abundance. The SA (expressed in square nautical miles) is then simply defined as twice the area below the curve expressing (Q–Q(T))/Q as a function of T (Fig. 4.8.1):

$$\mathrm{SA} = 2\int \frac{Q - Q(T)}{Q} \mathrm{d}T$$

As (Q– Q(T))/Q decreases from 1 to 0 and is convex, the SA is less than the PA. It equals the PA when the population is evenly spread with a constant density.

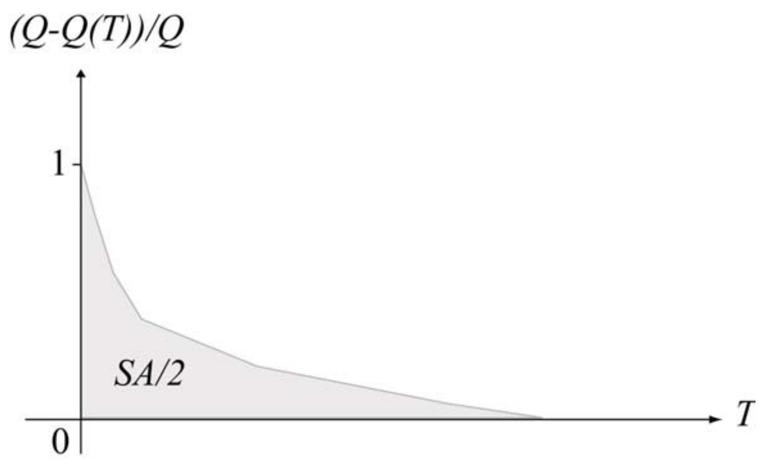


Figure 4.8.1. SA is defined as twice the area below the curve expressing (Q-Q(T))/Q as a function of T.

The curve above is a derivation of the Lorenz curve representing the histogram of fish density values, but having the advantage of receiving no contribution from zero density values. The spreading area can be related to the area occupied by the positive fish density values and their Gini index of dispersion G_0 through $\frac{SA}{PA} + G_0 = 1$ (Woillez et al. 2007).

Reference points

Spreading Area during a period of acceptable state of the population.

Interpretability

A spatial abundance is generally distributed in space into highly varying fish density values, spreading over its positive area. The spreading area index has been designed to describe this spreading, or equivalently the lack of aggregation or variation, while satisfying the condition of having no contribution from zero density values. Despite its name, the spreading area depends exclusively on the amount and histogram of positive fish density values.

Measurability

Zero values of density make no contribution to the spreading area. The spreading area depends on the variation in density values (and not on the absolute abundance) and is much less sensitive to low values of density than the positive area.

Sensitivity

Changes in SA are likely to reveal changes in the way the abundance is split into low and high density values.

Examples

In Woillez et al., 2007, the hake population of the Bay of Biscay has been described through the SA. This showed a better spread of the 3 year-old hake. In addition, a decrease of SA through the time series was detected for hake age 4 and 5+.

4.9 Equivalent area: EA

Description

The Equivalent Area represents the area, in square nautical miles, that would be covered by the population if all individuals had the same density, equal to the mean density per individual (Bez and Rivoirard, 2001).

Stock attribute

An individual-based measure of the area occupied by the stock, in square nautical miles.

Derivation

The transitive geostatistical approach (Matheron, 1971) can be used to describe the spatial distribution of a fish population when it includes a few large values of density, and when it is difficult to delimit a domain with homogeneous variations. The spatial structure is then represented by a (transitive) covariogram, a function of the distance between two locations:

$$g(h) = \int z(x) z(x+h) \mathrm{d}h \,.$$

Here, the equivalent area (EA) is defined as the integral range of the covariogram:

$$\mathrm{EA} = \frac{\int g(h)\mathrm{d}h}{g(0)} = \frac{Q^2}{g(0)} = \frac{Q^2}{\int z(x)^2 \mathrm{d}x} = \frac{(\int z(x)\mathrm{d}x)^2}{\int z(x)^2 \mathrm{d}x} \,.$$

It can also be written:

$$\mathrm{EA} = \frac{Q}{\int z(x) \frac{z(x)}{Q} \mathrm{d}x} \,.$$

It represents the area that would be covered by the population if all individuals had the same density, equal to the mean density per individual (the denominator in the previous equation (Fig. 4.9.1)).

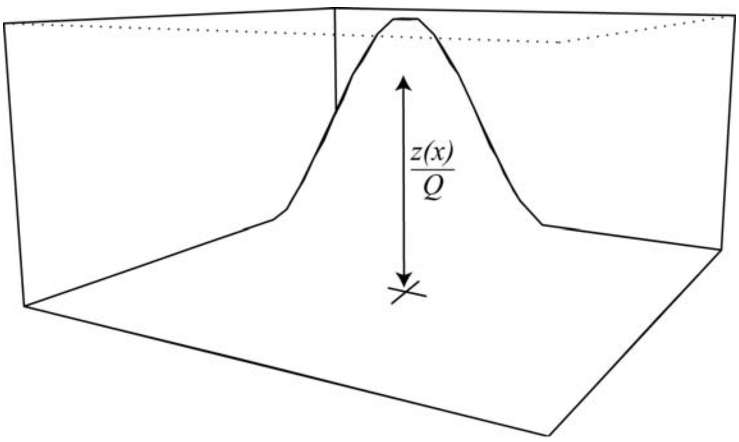


Figure 4.9.1. The probability density function for a random individual to be at x is given by z(x)/Q.

Practically, in the discrete case with sample values z_i and areas of influence s_i, it gives:

$$\mathrm{EA} = \left(\sum_{i=1}^{N} s_i z_i \right)^2 / \sum_{i=1}^{N} s_i z_i^{\ 2} .$$

The EA ranges from 0 to the PA. It would be equal to the PA if all strictly positive values of density were the same. The EA can be related to the area occupied by the positive fish density values and their coefficient of variation CV_0 through $\frac{PA}{EA} = 1 + CV_0^2$. The EA and SA are related through inequalities, in particular $EA \le \frac{9}{8} SA$ (Woillez et al. 2007).

Reference points

Equivalent Area during a period of acceptable state of the population.

Interpretability

The positive area describes the area of presence of fish, with a low density value being equivalent to a high one. The spreading area describes the area occupied by the stock, taking into account the variations in fish density. Now, the equivalent area is still another way to do this, while being individual-based (it gives the same weight to each individual, that is, the weight of a sample is proportional to its fish density). Just like the spreading area, the equivalent area depends exclusively on the amount and histogram of positive fish density values.

Measurability

The equivalent area is independent of the absolute abundance. Being individual-based, it is very sensitive to the highest density values. The inverse of the equivalent area can be considered as an index of aggregation (Bez and Rivoirard, 2001).

Sensitivity

Changes in EA are likely to reveal changes in the contribution of high density values to the total abundance.

Examples

In Woillez et al., 2007, the Equivalent Area on the hake population of the Bay of Biscay was shown to be larger for hake aged 3 years.

4.10 Microstructure index: MI

Description

The Microstructure Index (Woillez et al., 2007) measures the relative importance of structural components having a scale smaller than the sample lag (including random noise).

Stock attribute

The fine-scale variability of the fish density surface.

Derivation

The microstructure index (MI) is taken as the relative decrease of the transitive covariogram (Matheron, 1971; Bez et al., 1997) between distance zero and a distance h_0 chosen to represent the mean lag between samples (Fig. 4.10.1):

$$MI = \frac{(g(0) - g(h0))}{g(0)}$$

It lies between 0 and 1. Values close to 0 correspond to a very regular, well-structured density surface, and values close to 1 correspond to a highly irregular, poorly structured, density surface.

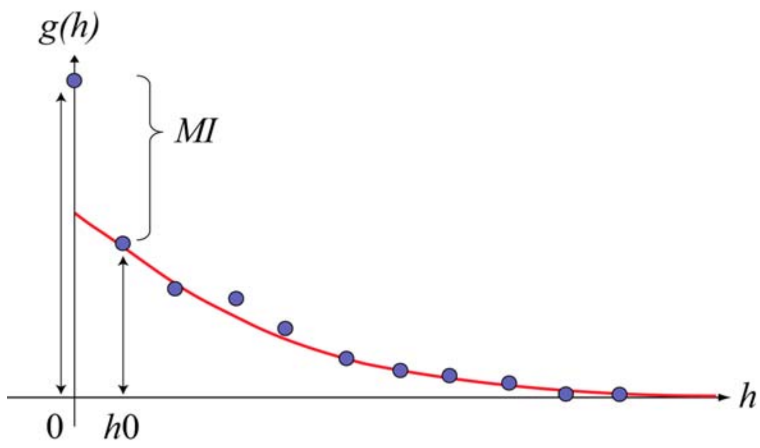


Figure 4.10.1. Real (dashed line), experimental (blue points) and modelled covariogram (red line) with the representation of the microstructure index.

Reference points

Microstructure Index during a period of acceptable state of the population.

Interpretability

The Microstructure Index does not make the distinction between spatial variability with a range less than the chosen lag but positive, and purely random variability (e.g. due to noise or sampling error).

Measurability

The Microstructure Index, as obtained from the transitive covariogram, is very sensitive to high fish density values (but it is more robust than its equivalent feature that would be obtained from the more traditional variogram or covariance).

Sensitivity

A high Microstructure Index is likely to correspond to fine-scale aggregations.

Examples

In Woillez et al., 2007, Microstructure Index has been followed through age and time on the hake population in the Bay of Biscay. It showed a relative stability for the younger ages, then it rose markedly from age 4.

4.11 References for spatial indicators

Alvarez P., Motos L., Uriarte A., and Egana J. 2001. Spatial and temporal distribution of European hake, Merluccius merluccius (L.), eggs and larvae in relation to hydrographical conditions in the Bay of Biscay. Fisheries Research, 50: 111–128.

Atkinson, D. B., Rose, G. A., Murphy, E. F., and Bishop, C. A. 1997. Distribution and abundance of northern cod, 1981–1993. Canadian Journal of Fish and Aquatic Sciences, 54 (Suppl. 1): 132–138.

Bez, N. 1997. Statistiques individuelles et géostatistique transitive en écologie halieutique. Thèse de Docteur en Géostatistique, Ecole Nationale Supérieure des Mines de Paris, France. 276 pp.

Bez, N., and Rivoirard, J. 2000. Indices of collocation between populations. In Report of a Workshop on the Use of Continuous Underway Fish Egg Sampler (CUFES) for Mapping Spawning Habitat of Pelagic Fish, pp. 48–52. Ed. by D. M. Checkley, J. R. Hunter, L. Motos, and C. D. von der Lingen. GLOBEC Report, 14.

Bez, N., and Rivoirard, J. 2001. Transitive geostatistics to characterize spatial aggregation with diffuse limits: an application on mackerel ichthyoplankton. Fish. Res. 50: 41-58.

Bez, N., Rivoirard, J., Guiblin, Ph., and Walsh, M., 1997. Covariogram and related tools for structural analysis of fish survey data. E.Y. Baafi and N.A. Schofield (eds), Geostatistics Wollongong`96, Volume 2, 1316-1327.

Brodie, W. B., Walsh, S. J., and Atkinson, D. B. 1998. The effect of stock abundance on range contraction of yellowtail flounder (*Pleuronectes ferruginea*) on the Grand Bank of Newfoundland in the Northwest Atlantic from 1975 to 1995. Journal of Sea Research, 39: 139–152.

Hollowed, A. B. 1992. Spatial and temporal distribution of Pacific hake, *Merluccius productus*, larvae and estimates of survival during early life stages. CalCOFI Report, 33: 100–123.

Kendall, A. W., and Picquelle, S. J. 1989. Egg and larva distributions of walleye pollock *Theragra chalcogramma* in Shelikof Strait, Gulf of Alaska. Fishery Bulletin US, 88: 133–154.

Matheron, G. 1971. The theory of regionalized variables and its applications. Les cahiers du Centre de Morphologie Mathématique, Fascicule 5. Ecole Nationale Supérieure des Mines de Paris, Paris. 212 pp.

Myers, R. and Cadigan, N. (1995) Was an increase in natural mortality responsible for the collapse of northern cod? Can. J. Fish. Aquat. Sci. 52: 1274-1285.

Petitgas P (1998) Biomass dependent dynamics of fish spatial distributions characterized by geostatistical aggregation curves. ICES Journal of Marine Science 55: 443-453

Swain D, Sinclair, A (1994) Fish distribution and catchability: what is the appropriate measure of distribution? Can. J. Fish. Aquat. Sci. 51:1046-1054

Woillez, M., Poulard, J-C., Rivoirard, J., Petitgas, P., and Bez, N. 2007. Indices for capturing spatial patterns and their evolution in time, with application to European hake (*Merluccius merluccius*) in the Bay of Biscay. ICES Journal of Marine Science, 64: 537–550.

5. Methods

5.1 Introduction

The following methods for integrating and interpreting time-series of indicators as estimated from research vessel surveys were developed and discussed during the FISBOAT project. In general, they attempt to ease the problem of deciding whether an indicator or a suite of indicators is signalling, or will signal, a change in the state of a fish stock that calls for an adjustment to the controls on a fishery.

Section 5.2 by Lembo et al. presents an age-length based simulation model, Aladym, for predicting the effects of different fishing pressures on a single population of fish. Section 5.3, also by Lembo et al. describes how a Monte Carlo approach can be used with Aladym in order to help create reference points for selected indicators.

Section 5.4 by Trenkel et al. points out the importance of relating current trends in indicator values with a previous reference period when indicator values were agreed to be either satisfactory or not. They offer a simple system for integrating different types of biological and fishery information provided by indicator time series, and by other sources if available, in a way that can be discussed by stock managers, stake-holders, and scientists in order to decide what, if any, measures should be taken to control the fishery. This would often form part of an adaptive management scheme.

Mesnil and Petitgas, section 5.5, describe how the quality control schemes that originated in manufacturing industry can be adapted to monitor fishery and environmental qualities derived from time series of indicators. The CUSUM control-chart method offers considerable potential for rapid detection of changes of state.

Bogaards et al., section 5.6, describe a simple, hypothesis-testing approach for deciding how long will be needed before an indicator series is expected to reveal a given linear trend. This could be helpful for deciding whether a survey is sufficiently sensitive to detect a response to new controls on a fishery within a reasonable time frame.

In section 5.7, Trenkel proposes a GAM and bootstrap-based solution for the long standing problem of deciding whether recent changes in an important time-series represent a valid signal, or just sampling noise. Such methods can greatly assist the provision of rapid, confident advice for managing a fishery.

In sections 5.8 and 5.9, Petitgas describes and illustrates two complementary approaches to making a single assessment from multiple time-series of indicators. In the first part, the well-known 'traffic light' method is set out. In the second, multivariate methods based on principal components and multi-factorial analysis (MFA) are described and illustrated with an example. The multivariate analysis is likely to provide complementary interpretation of results in the traffic light table.

Petitgas and Poulard in section 5.10 describe a multivariate statistical method for visualising groupings of indicator variables in space and time. They applied this method during the FISBOAT project to examine the changing spatial distributions of fish with age, as signalled by the spatial indicators described in the first part of this manual.

In section 5.12, Woillez and Rivoirard describe a multivariate statistical method applicable to parallel time series of values for many indicators when continuity of some trend in time is of interest, possibly in relation to a study of cause and effect. MAFs are linear combinations of variables that are conceptually similar to principal components (PCs) but, whereas a fitted series of PCs explains independent components of variance the magnitude of which decreases from first to last, a series of MAFs explains independent components of autocorrelation, the first of which displays the highest, and the last of which displays the lowest autocorrelation at lag 1 observation in the time series. Thus MAFs offer a way of finding the combination of variables that present maximal continuity in time.

Finally, in section 5.13, Cotter reviews and illustrates with example analyses a collection of nonparametric statistical methods that allow assessment of indicator trends whilst avoiding the assumptions and other uncertainties of modelling. Suggestions are made for improving the objectivity of statistical inference.

Several of the methods use especially written software. The FISBOAT project team agreed at the outset to use the R programming language because it is freely available from http://www.r-project.org/, it is highly versatile and because, by doing so, the portability of software and ideas is maximised. The software can be downloaded freely from http://www.ifremer.fr/drvecohal/fisboat/index.htm. Data sets and spreadsheets referred to in the following sections should also be available from this site.

5.2 Simulating population dynamics.

5.2.1 ALADYM(v 08)

Lembo G., A. Abella, F. Fiorentino, S. Martino and M.T. Spedicato
COISPA Tecnologia & Ricerca, Bari, Italy

Introduction

ALADYM (Age-Length Based Dynamic Model) is an age-length based simulation model developed within the conceptual framework of dynamic pool models, following the predictive Thompson & Bell (1934) approach. The model is designed to predict, through simulations, the effects of different fishing pressure scenarios on a single population, in terms of different metrics and indicators. Removals are simulated on the basis of the total mortality rate modulated using harvesting pattern and a fishing activity coefficient. *Aladym* can work in absence of fishery-dependent data, although its predictive capability of real catch levels can be verified using information on commercial catches or fishing activity per month.

From the *Aladym* core model three complementary, but independent, tools have been derived:

- the quasi-deterministic dynamic tool named *Aladym-r*;
- the tuning tool *Aladym-z*;
- the stochastic dynamic tool named *Aladym-q*.

The core *Aladym* model is described in this chapter together with *Aladym-r* and *Aladym–z*, while *Aladym-q* is described in the following one.

General assumptions

The basic assumptions of the model are:

- natural mortality as estimated reflects the rate of decline of a population from all causes excluding fishing;
- total mortality *Z* reliably reflects the decline of ages/sizes in the population, including the effects of different fishing gears;
- the growth, the natural mortality, and the maturity parameters are assumed constant over time;
- given the small time interval (1 month) between cohorts the effect of the spreading of the lengths with respect to the ages can be neglected.

Derivations

The quasi-deterministic dynamic tool named Aladym-r

General framework

The model is designed to simulate population dynamics of a given species accounting for differences by sex in growth, maturity and mortality. All the quantities are calculated as vectors with a time step Δt (time slice=1 month). An operational framework of the *Aladym-r* model is shown in fig. 5.2.1. Step A) regards the *input and initialization.* In order to generate an unbiased initial population, the number of runs specified by the user (e.g. 100) is performed in this step, randomly varying the recruitment, the growth and the size-at-maturity parameters according to the values and distributions specified by the user. The user can choose among the following distribution type: log-normal, normal, gamma and uniform. For the parameter t_0 a uniform distribution is associated by default. Two populations are generated: the exploited (where total mortality is acting) and the unexploited one (where natural mortality only is acting).

The obtained initial populations enter in the *start loop* (or *seed run*) (step B in fig. 5.2.1), where the dynamics are formulated to follow the evolution of several cohorts over a monthly scale. Here the number of recruits entering in the population is generated from a stock-recruitment relationship. Alternatively, it is given as an input vector. In both cases, a uniform variability for the obtained number of recruits can be set by the user. The *start loop* runs for a number of years that is a multiple of the two sex life-spans. This step aims to eliminate the artefacts in the initial population due to the use of an equilibrium model in the initialization step. After this phase, the *simulation loop* starts and runs over the period required by the user (step C in fig. 5.2.1) generating the outputs (step D in fig. 5.2.1).

Model components

Growth

The growth process is modelled using a von Bertalanffy growth function:

$$L_{age} = L_{\infty} \cdot \left(1 - e^{-K\cdot(age - t_0)}\right).$$

For each age (time step Δt = 1 month) length is calculated using the input parameters L_{∞}, K and t_0. The average length in the time interval (t, $t+\Delta t$) is calculated as:

$$\overline{L}_{age} = L_{\infty} + \frac{(L_{age} - L_{age+\Delta t})}{K \cdot \Delta t}$$

The weight at average length, for each age, is calculated from the length-weight relationship in the form: $W_{age} = a\overline{L}_{age}^{b}$; with a and b as input parameters.

Population

The population dynamics is formulated following the simultaneous evolution of several cohorts at monthly scale through the exponential population decline model, both in absence (1) and in presence (2) of fishing mortality:

$$\frac{dN}{dt} = -MN \qquad (1)$$

$$\frac{dN}{dt} = -ZN \qquad (2)$$

used respectively in the form (3) and (4):

$$N_{(t+\Delta t),j} = N_{t,j} e^{-M_{t,j}\cdot\Delta t} \qquad (3)$$

$$N_{(t+\Delta t),j} = N_{t,j} e^{-(F_{t,j} + M_{t,j})\cdot\Delta t} \qquad (4)$$

where j indicates the cohort, t the time, Z, M and F the total, natural and fishing mortality respectively. (Notice that in any formula where j, *age* and t are present, it is assumed that *age* represents the age of the cohort j at time t).

Maturity

Maturity *Mat* is a function of the length L and is calculated following an ogive model (Quinn and Deriso, 1999):

$$Mat(L) = \frac{1}{1 + e^{-r(L - Lm50\%)}}$$

where r is the ogive slope and $L_{m50\%}$ is the length at which 50% of fish mature.
The proportion of mature fish at age is computed as:

$$Mat_{age} = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{m75\%} - L_{m25\%}}\right) \cdot \left(L_{m50\%} - \bar{L}_{age}\right)}}$$

where the maturity range $L_{m75\%}$-$L_{m25\%}$, is related to the ogive slope.

Biomass

The biomass (B_j) and the spawning stock biomass (SSB_j) of the cohort j at time t are respectively computed as:

$$B_{t,j} = N_{t,j} \cdot w_{age};$$
$$SSB_{t,j} = N_{t,j} \cdot w_{age} \cdot Mat_{age}$$

Analogously, the unexploited biomass (UB_j) and the unexploited spawning stock biomass ($USSB_j$) of the cohort j at time t are calculated as:

$$UB_{t,j} = UN_{t,j} \cdot w_{age};$$
$$USSB_{t,j} = UN_{t,j} \cdot w_{age} \cdot Mat_{age}$$

Initial recruitment and stock recruitment relationship

During the step A) (fig. 5.2.1) the initial number of individuals in the population are from estimates of recruitment independently obtained from e.g. trawl surveys or other sources. These numbers randomly selected for each of the e.g. 100 runs (see also the *general framework* paragraph) are used to initialize the population. Successively (step B and C in fig. 5.2.1), the number of individuals entering in the population can be a vector or is estimated from one of the following user selected stock-recruitment relationships:

Beverton & Holt (1957):
$$R = \frac{S}{(a + bS)};$$

Ricker (1954):
$$R = a \cdot S \cdot e^{(-bS)};$$

Shepherd (1982):
$R = a \cdot S / \left[1 + (S/c)^b\right]$;

Barrowman & Myers (2000):
$$R = \alpha \cdot \min(S, S^*) = \begin{cases} \alpha \cdot S \ \ if \ \ S < S^* \\ \alpha \cdot S^* \ if \ \ S \geq S^* \end{cases}$$

$$R = \begin{cases} \alpha \cdot S & if \ \ S \leq S^* \cdot (1-\delta) \\ \alpha \cdot (S - \dfrac{(S - S^* \cdot (1-\delta))^2}{4\delta \cdot S^*}) & if \ \ S^* \cdot (1-\delta) < S < S^* \cdot (1+\delta) \\ \alpha \cdot S^* & if \ \ S \geq S^* \cdot (1+\delta) \end{cases}$$

R and S represent the number of recruits and spawners respectively, whilst a, b, c, α, δ, S^* are the model's parameters. Uniformly distributed random variations can be applied by the user to the number of offspring (from the vector or stock-recruitment relationship).

The number of the events (on monthly basis) generating the offspring is an input of the model. The population of spawners generating the recruits is calculated by summing up the number of individuals of the different age classes of the different cohorts occurring in the population one or more (depending on the biological features of the species) months before the offsprings are produced. Thus this quantity is calculated as follows:
$SSN_t = \sum_j SSN_{t,j}$;
where $SSN_{t,j}$ represents the number of mature females at time *t,* of the cohort *j*:
$SSN_{t,j} = N_{t,j} \cdot Mat_{age}$.

Mortality
The natural mortality can be constant for each age/length, or a vector by age/length calculated outside the model and used as input. Alternatively, it is estimated inside the model from the Chen and Watanabe equations (1989):

$$M(t) = \begin{cases} \dfrac{K}{1 - e^{-K \cdot (t - t_0)}} & t \leq t_M \\ \dfrac{K}{a_0 + a_1 \cdot (t - t_M) + a_2 \cdot (t - t_M)^2} & t \geq t_M \end{cases}$$

where:

$$t_M = -\frac{1}{K} \ln\left|1 - e^{Kt_0}\right| + t_0$$
$$a_0 = 1 - e^{-K \cdot (t_M - t_0)}$$
$$a_1 = K \cdot e^{-K \cdot (t_M - t_0)}$$
$$a_2 = -\frac{1}{2} K^2 \cdot e^{-K \cdot (t_M - t_0)}$$

The two parameters of the Chen and Watanabe model are t_0 and K. The asymptotic length (L_∞) is not necessary, but t_0 cannot be equal to 0 (otherwise the parameter t_M cannot be defined). The quantities a_0, a_1, a_2 and t_M cannot be strictly considered as parameters of the model, as they depend on t_0 and K. The parameter t_M represents the age beyond which the contribution of the fish of a given cohort can be considered negligible. If parameters are consistent the relationship between age and natural mortality shows a "bath tube" shape.

The fishing mortality rate $F(L)$ is modelled for each cohort using the following general equation (Sparre and Venema, 1998):

$$F(L) = F_{max} \cdot S(\bar{L})$$

where F_{max} is the maximum fishing mortality and $S(\bar{L})$ the proportion of retained fish. In *Aladym* the fishing mortality rate is calculated as

$$F(L) = F_{max} \cdot S(\bar{L}) \cdot f_{act}$$

where maximum fishing mortality (F_{max}) is calculated as

$$F_{max} = QZ_{input} - M_{\min}$$

using the input values of QZ (a Z proxy) and where M_{min} represents the minimum value that the M vector assumes. As an alternative option, F_{max} can also be a user selected input to be set for each month. In addition, a fishing activity coefficient (f_{act}) is introduced in order to consider the possibility of a fishing ban or changes in fishing effort throughout time.
The value of QZ by sex can be assumed, as a first order approximation, numerically equal to the value of Z observed that is obtained from estimations outside the simulation model (e.g. from trawl-survey). A better approximation of QZ is obtained using the tool *Aladym-z* (see a later paragraph).

In the model, the probability of selection $S(\bar{L})$ of the cohort j is calculated at time t from one of the two following user-selected relationships:

$$S(\bar{L}) = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{75\%} - L_{25\%}}\right) \cdot \left(L_{50\%} - \bar{L}_{age}\right)}};$$

or

$$S(\bar{L}) = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{75\%} - L_{25\%}}\right) \cdot \left(L_{50\%} - \bar{L}_{age}\right)}} \cdot \frac{1}{1 + e^{\left(\frac{-2 \cdot Ln3}{D_{25\%} - D_{75\%}}\right) \cdot \left(D_{50\%} - \bar{L}_{age}\right)}};$$

where $L_{50\%}$, $L_{75\%}$ and $L_{25\%}$ are the selectivity parameters and $D_{50\%}$, $D_{25\%}$, $D_{75\%}$ the de-selection parameters of the model. The total mortality Z at time t for the cohort j is thus computed as

$$Z_{t,j} = F_{t,j} + M_{t,j}$$

that is the value acting on the population in the model computations.

The biomass of individuals of the cohort *j* at time *t* death for all causes ($BP_{t,j}$) is computed as

$$D_{t,j} = N_{t,j} - N_{t+\Delta t,j} = N_{t,j} \cdot \left(1 - e^{-Z_{t,j} \cdot \Delta t}\right),$$
$$BP_{t,j} = D_{t,j} \cdot w_{age};$$

while the biomass of those dead from all causes excluding fishing ($BND_{t,j}$) is computed as

$$BND_{t,j} = \frac{M_{t,j}}{Z_{t,j}} \cdot N_{t,j} \cdot (1 - e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t}) \cdot \mathrm{w}_{\mathrm{age}}.$$

Harvest control rules

The simulation approach can be used as a tool to convert survey biological information and relative assessment into quantitative HCRs. The options implemented in the simulation model are based on the following aspects: *QZ*, gear selectivity (size at first capture $L_{50\%}$ and selection range) and fishing activity (alone or in combination). These three are inputs that can be used to simulate different exploitation scenarios. The effects of HCRs (selectivity and fishing activity) are then analysed in terms of the sustainability of the population in the long-term. For example, the ratio between the mean spawning stock biomass and the mean unexploited spawning stock biomass (*SSB/USSB*, output) is also estimated for each harvesting scenario.

A vector of yield (*Y*) by time is also simulated, estimating the catch (*C*) according to the following general equation (Gulland, 1969):

$$C_{\Delta t} = \int_0^{\Delta t} F \cdot N_0 \cdot e^{-Z \cdot \tau} d\tau = \frac{F}{Z} N_0 \cdot \left(1 - e^{-Z \cdot \Delta t}\right)$$

where Δt is the time to which the catch is referred. Thus the catch (Yield) in the time interval (t, $t+\Delta t$) is computed in *Aladym* as (Sparre and Venema, 1998):

$$Y_{t,j} = \frac{F_{t,j}}{Z_{t,j}} \cdot N_{t,j} \cdot (1 - e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t}) \cdot \mathrm{w}_{\mathrm{age}}.$$

Software

Aladym is written in the R language and licensed as open source under GPL2. The data and parameters feeding the model can be easily entered using an excel data sheet. The results of the simulation are stored into three Export files (.din for inputs, .dou for outputs, .RData for the R workspace) and saved in the same directory where R is started using the basename of the input sheet.

To give an idea of the running time, *Aladym-r* requires about 25 seconds (assuming 40 years of start loop and 20 years of simulation) with a Intel (R) Pentium (R) personal computer with a processor of 1.70 GHz and 1 GB RAM. The tool *Aladym-z* requires about 2.6 hours (assuming 40 years of start loop and 20 years of simulation) with the same computer. The software can be downloaded from the Fisboat web-site, where also a detailed description of the input sheet for user help is available.

Inputs

Input parameters to the *Aladym-r* model are:

- von Bertalanffy growth parameters by sex with associated variability,
- length-weight relationship parameters by sex;
- maturity ogive parameters by sex ($L_{m50\%}$ and $L_{m25\%}$-$L_{m75\%}$ range);
- natural mortality by sex (a constant value or a vector);
- seed values (minimum, maximum, *ln*-mean and *ln*-standard deviation) of recruitment by sex;
- proportion of offsprings entering in the stock by month;
- stock-recruitment relationship parameters or a vector of recruit numbers by month both with associated variability;
- time elapsing from spawning to birth;
- sex-ratio (female/total) of offsprings;
- F_{max} by month (option 2) or from the model (option 1);
- *QZ* by sex;
- selection ogive parameters (2 options) of the gear used by the fleet ($L_{50\%}$ and $L_{25\%}$-$L_{75\%}$ range, $D_{50\%}$ in case of the selectivity option 2);
- fishing activity coefficient by month (0, in case of absence of fishing activity).

Outputs

The outputs automatically produced by the simulations of *Aladym-r* can be summarised in the following items:

- Export data file (.dou):
- exploited and unexploited population by sex, per month and age;
- exploited and unexploited biomass by sex, per month and age;
- exploited and unexploited population of females belonging to the spawning stock per month;
- total mortality Z calculated by the model for females, males and the whole population in each month and year of the simulation as follows (Sinclair, 2001):
- $$Z_t = \frac{1}{\Delta t}\ln\left(\frac{\sum_{j=1}^{\infty} N_{t,j}}{\sum_{j=2}^{\infty} N_{t+\Delta t,j}}\right);$$
- exploited and unexploited biomass per month;
- exploited and unexploited spawning stock biomass per month;
- ratio between exploited and the unexploited spawning stock biomass per month;
- average length and age of exploited and unexploited populations per month;
- average length and age of exploited and unexploited spawning populations per month;
- yield in tonnes per month;

- average length and age of catches per month;
- fishing mortality per month calculated as;

$$F_t = \frac{1}{\Delta t}\ln\left(\frac{\sum_{j=1}^{\infty} N_{t,j}}{\sum_{j=2}^{\infty} N_{t+\Delta t,j}^{F}}\right),$$

where $N_{t+\Delta t,j}^{F}$ is the number of survivors at the time $t+\Delta t$ under the hypothesis that only fishing mortality is acting;
- biomass of natural losses and total biological production per month.

Plots per year of the outputs listed from items 4 to 13 are also produced. Some other outputs are also made available to the user:
- average length at age and age by sex;
- natural mortality at age/length by sex;
- weight at age/length by sex;
- proportion of mature individuals at age/length by sex.

These outputs help the user to check the results obtained from the sub-models, in particular those related to the VBGF, the length-weight relationship, the natural mortality, and the maturity.

Practical guidelines

The *Aladym* core model does not make any fixed or hidden (from the user) assumption about the values of the parameters describing the behaviour of the equations on which the model itself is built. The user is allowed to (and needs to) input all the parameters involved: whilst this makes the model highly flexible in adapting to different species/environments it loads the user with the responsibility to validate each single value and to assess the coherence as a whole. Very few checks are foreseen at the moment to supervise the consistency of the data supplied: often it is a critical analysis of the results which spots such consistency. The checks guarantee the positivity of F_{max}, of length at t_0 and a sex ratio between 0 and 1.

The model is extensively based on a closed form solution to the dynamical equations it solves. Thus two key options, both related to the early phase, are available for tuning: the 'Multiplier of Life-span' which controls the amount of years that must be simulated in order to cancel the artefacts from the equilibrium model used to initialise the population; and the 'Number of Run for seed randomization' which sets the number of samples to be taken in order to derive the average values for the growth and population parameters. For both parameters the rule is: bigger is better, however the default values (1, 100) are a reasonable choice.

One of the parameters highly influencing the behaviour of the model is *QZ* which, however, does not have an immediate counterpart but can be naively associated to the total mortality *Z*. A specific tool (*Aladym-z*) has been developed which, starting from the observed values of Z and the description of the life and population traits, is able to calculate values of *QZ* which better approximate the given scenario. Starting from the *Z_observed*, *Aladym-z* iterates the model modifying, in each run, the amplitudes of the *QZ* waveforms. It stops when the Least Square convergence criteria are met.

Sensitivity

The extensive number of simulation runs performed has shown that the model behaviour is influenced by the consistency between the set of life-history parameters and the population dynamics. The model results are thus expected to be particularly sensitive to the stock-recruitment relationship and natural mortality.

Strengths/weaknesses

In *Aladym* the following points can be considered the strengths:

- the model is designed to work in the absence of fishery-dependent information;
- the model is built using separated components that give it enough flexibility to account for the use of different equations;
- the model allows the population dynamics to evolve in a very detailed time scale, thus permitting analysis of fluctuations within the year;
- the detailed time scale allows modelling the effects of the harvest control over the year;
- the model allows input of natural mortality varying by age/length, and is thus able to allow for species exploited at an early phase.

The following points can be considered as the weak ones:

- the model does not account for environmental changes, such as those related for example to temperature variations, or food availability;
- the life-history traits that are used for modelling the population dynamics (e.g. growth, natural mortality, maturity) are assumed stable along the time and not to be density dependent; only direct effects of the fishery on the population are considered;
- the model does not include components of spatial behaviour;
- harvesting scenarios based on the control of the total catches are not foreseen;
- the user should be aware of the range of validity of the sub-model parameters such as those related to the stock-recruitment relationships.

References

Beverton R.J.H., Holt S.J. (1957). On the dynamics of exploited fish populations. U.K. Min. Agric. Fish., *Fish. Invest*. (Ser. 2), 19: 533 pp.

Barrowman N.J., Myers R.A. (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. *Can. J. Fish Aquat. Sci*., 57: 665-676.

Chen S., Watanabe S. 1989. Age Dependence of Natural Mortality Coefficient in Fish Population Dynamics. *Nippon Suisan Gakkaishi*, 55(2): 205-208.

Gulland J.A. (1969). Manuel des methods d'évaluation des stocks d'animaux aquatiques. Première partie-Analyse des populations. *Manuels FAO de science halieutique*, N. 4 : 160 pp.

Quinn T.J.II, R.B. Deriso (1999). *Quantitative fish dynamics*. Oxford University Press, Oxford, U.K.: 542 pp.

Ricker W.E. (1954). Stock and recruitment. *J. Fish. Res. Bd. Can.*, 11: 559-623.

Shepherd J.G. (1982). A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. *Journal du Conseil, Conseil Internationale pour l'Exploration de la Mer*, 40 (1): 67-75.

Sinclair, A. F. (2001). Natural mortality of cod (*Gadus morhua*) in the Southern Gulf of St Lawrence. ICES Journal of Marine Science, 58: 1-10.

Sparre P., Venema S.C. (1998) - Introduction to tropical fish stock assessment. Part 1, manual. *FAO Fish. Techn. Pap*., 306 (1) Rev. 2: 407 pp.

Thompson W.F., Bell F.H. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. *Rep. Int. Fish. (Pacific Halibut) Comm.*, (8): 49 p.

Working documents

Lembo G, M.T. Spedicato (2006a). ALADYM – Age-Length based dynamic Model. *In*: Petitgas P. (coordinator). Fisboat Activity Report, Appendices: 113-114 p.

Lembo P. and M.T. Spedicato (2006b). ALADYM Age-Length Based Dynamic Model. MEDITS Coordination Meeting, Kavala, Greece, April 2006.

Lembo G., Spedicato M.T. (2006c). Red mullet assessment in the GSA 10 using Aladym simulation model. Sub-Committee Stock Assessment SAC-GFCM, FAO Headquarters, 11-14 September 2006. 6 pp.

Lembo G., Spedicato M.T., Abella A., Fiorentino F, Martino S. (2006d). Aladym description (version 06). Fisboat Meeting, IJmuiden 06-10 November 2006, 11 pp.

Lembo G., Spedicato M.T. (2006e). Aladym: un modello previsionale di valutazione e gestione per le risorse della pesca nel Mediterraneo. 5° Convegno Nazionale delle Scienze del Mare, promosso ed organizzato da CoNISMa in collaborazione con A.I.O.L., S.I.B.M. e S.It.E. Viareggio, 14-18 Novembre 2006. abstract.

Lembo G., Martino S., Abella A.J., Fiorentino F. and M.T. Spedicato. 2007. *ALADYM* (Age-Length Based Dynamic Model): a stochastic simulation tool to predict population dynamics and management scenarios using fishery-independent information. Scientific Advisory Committee-Sub-Committee Stock Assessment. Workshop on trawl survey based monitoring fishery system in the Mediterranean, Rome, Italy, 26-28 march 2007. 6 pp.

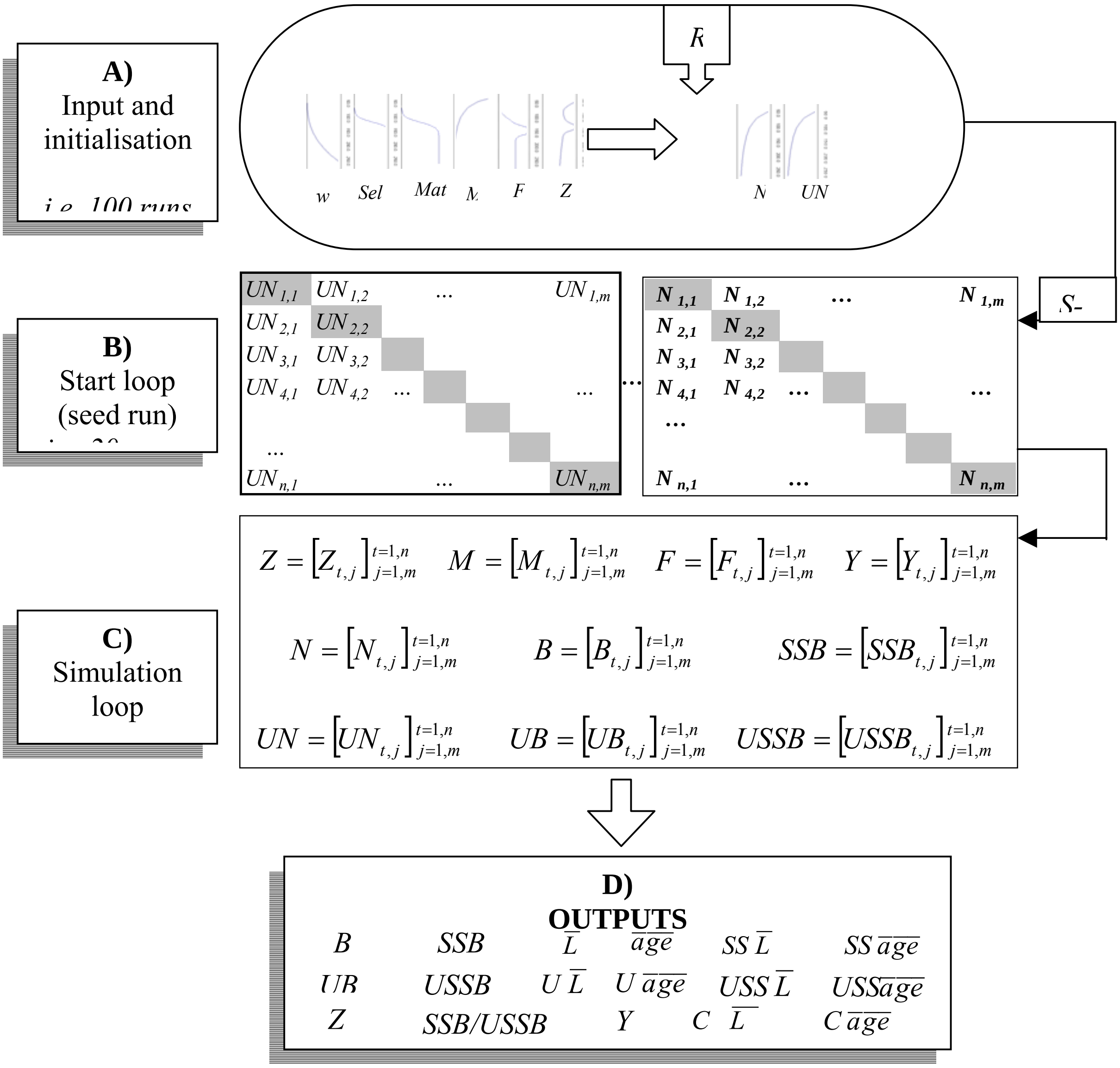


Figure 5.2.1. Scheme of the *Aladym-r* tool. *R*=recruitment; *w*=individual weight; *Sel*=selectivity; *Mat*=maturity; *M*=natural mortality; *F*=fishing mortality, *Z*=total mortality; *N*=exploited population, *UN*=unexploited population, *B*=exploited biomass, *SSB*=exploited spawning stock biomass, *UB*=unexploited biomass, *USSB*=unexploited spawning stock biomass, *S-R*=stock-recruitment relationship, *Y*=yield, *t*=time, *j*=cohort.

5.2.2 Estimating indicators and reference points

Lembo G., A. Abella, F. Fiorentino, S. Martino, and M.T. Spedicato
COISPA Tecnologia & Ricerca, Bari, Italy

Introduction

Aladym-q adds to the same mathematical model of *Aladym-r* the capability to deal with the stochastic representation of some input parameters, in order to evaluate the corresponding distribution functions of the output variables using a MonteCarlo approach. This feature aims to build a procedure to help identification of indicators and/or reference points, associating a confidence interval with them.

Derivation

The stochastic dynamic tool Aladym-q

The stochastic dynamic model defined as *Aladym-q* follows the same basic formulations as *Aladym-r.* The main difference consists in modelling the uncertainty of estimates related to the initial recruitment, growth and maturity traits of the population through stochastic processes. Moreover, a uniform distribution is applied to the number of recruits generated by the stock-recruitment relationship. In addition, probability distribution functions (*pdf*) selected by the user are applied to the growth parameters K and L_∞, and to the maturity parameters. This makes *Aladym-q* more adaptable for estimating the probability associated to metrics, indicators and reference points.

An operational framework of the *Aladym-q* is in fig. 5.3.1. The step AA) concerns the *input and initialization.* Given the parameters of the identified *pdfs* a first random realization is made in this step. Then the population evolves in the steps BB) and CC). These steps are reiterated for a number of realizations, sampling at each run a new set of parameters from the *pdfs*. In the output step *pdfs* and cumulative *pdfs* are generated, the latter calculated according the following general formulation:

$$f(X) = P(X < x) = \int_{-\infty}^{x} pdf(\chi)d\chi$$

Software

Aladym is written in the R language and licensed as open source under GPL2. The data and parameters feeding the model can be easily entered using the same excel data input sheet as *Aladym-r*. The differences regard the number of realizations to be performed (user selected and mandatory for *Aladym-q*) and the parameters of the *pdfs* associated with growth and maturity, that for *Aladym-q* operate also in the simulation loop. The results of the simulation are stored into three Export files (.din for inputs, .dou for outputs, .RData for the R workspace) and saved in the same directory where R is started using the basename of the input sheet.

To give an idea of the running time, using a Intel (R) Pentium (R) personal computer with a processor of 1.70 GHz and 1 GB RAM, *Aladym-q* might requires 572 seconds for 100 realizations, ~1.5 hours for 1000 realizations and about 17 hours for 10000 realizations (assuming 40 years of start loop and 20 years of simulation). The software can be downloaded from the Fisboat web-site, where also a detailed description of the input sheet for user help is available.

Inputs

As regards the *inputs*, besides those already mentioned for *Aladym-r*, *Aladym-q* requires:

- the number of realizations;
- the parameters of the *pdfs*.

Outputs

The *outputs* automatically produced by the simulations of *Aladym-q* can be summarised in the following items:

- Export data file (the quantities are related to each realization):
  1. exploited and unexploited biomass in tons per month;
  2. exploited and unexploited biomass of spawners in tons per month;
  3. ratio between exploited and unexploited spawning stock biomass per month;
  4. *Z* calculated by the model combined for sex per month and by sex per year;
  5. QZ (the input values) by sex;
  6. average length and age of exploited and unexploited populations per month;
  7. average length and age of exploited and unexploited spawner populations per month;
  8. F per month;
  9. yield in tons per month;
  10. average length and age of the catches per month;
  11. biomass of natural losses and total biological production in tons per month.
- Plots of the *pdfs* and the cumulative (*cpdfs*) are interactively produced per year for the same items listed above.
- Some other outputs are also made available to the user:
  12. average number of recruits at each realization;
  13. growth and maturity parameters by sex at each realization.

These outputs help the user to check the results from the sub-models related to the VBGF, the maturity, and the recruitment. In addition, they also allow the outputs at each realization to be tracked with the related key-inputs.

Practical guidelines

The same considerations that were developed for *Aladym-r* hold for *Aladym-q*. A new parameter is introduced for tuning the quality of the output pdfs: the number of realizations. This parameter should be set so as to account for a trade-off between the running time and the target confidence level. Experiments showed that values in the range from 1000 to 10000 give an error level varying from about 6-7 to ~1%. These confidence levels are well below the precision by which most of the input parameters are known.

As regards sensitivity and the strengths/weaknesses of the models, similar consideration as were developed for *Aladym-r* can be applied to *Aladym-q*, although the latter tool has the advantage of including stochastic effects in some of the key life-history traits. This stochasticity masks the effects due to uncertainty on the knowledge of input data and of their relationships.

References

See section 5.2: Simulating population dynamics. Aladym model.

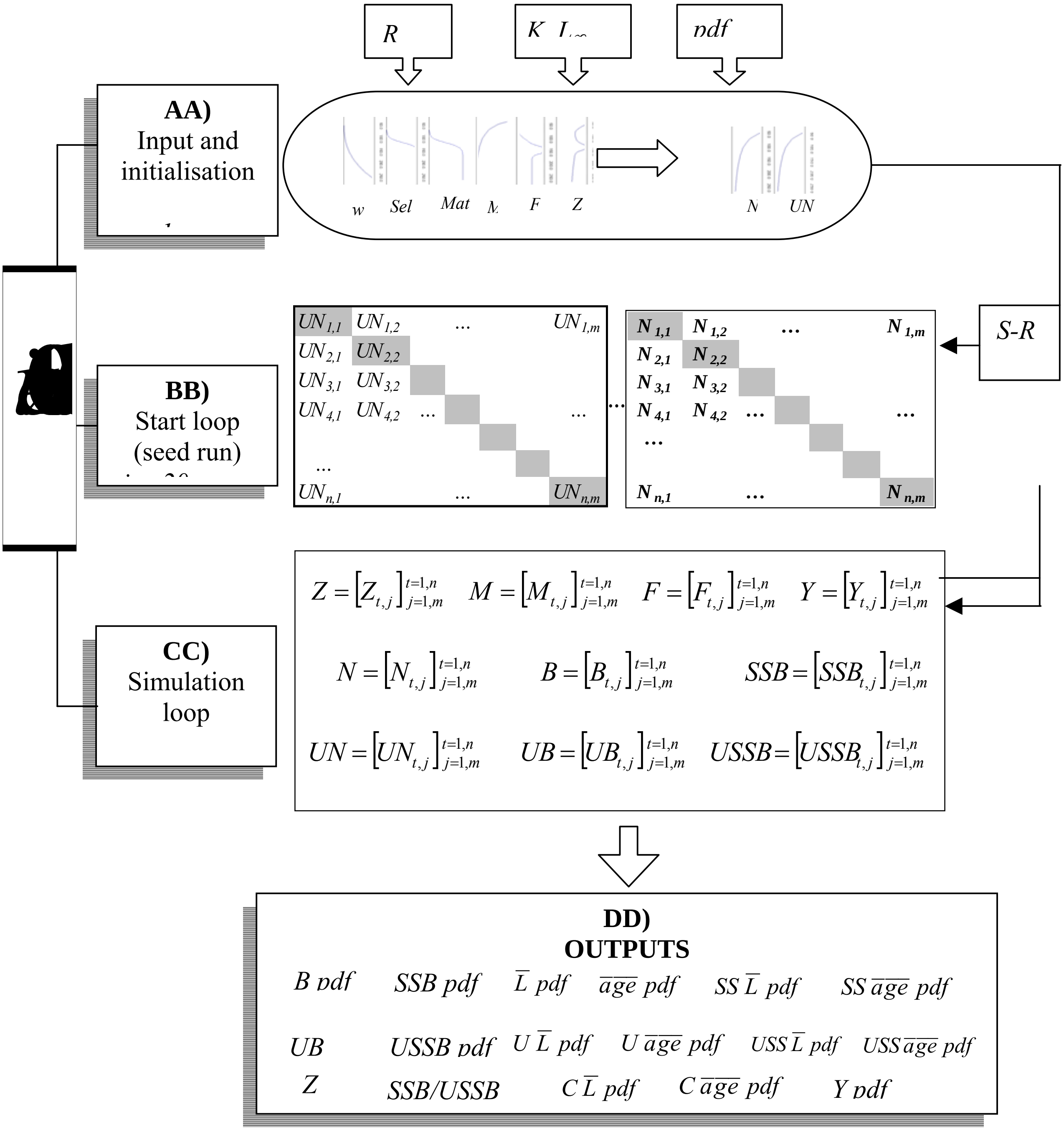


Figure 5.3.1. Scheme of the *Aladym-q* tool. *pdf*=probability distribution function; *K, L_∞* growth parameters, *R*=recruitment; *w*=individual weight; *Sel*=selectivity; *Mat*=maturity; *M*=natural mortality; *F*=fishing mortality, *Z*=total mortality; *N*=exploited population, *UN*=unexploited population, *B*=exploited biomass, *SSB*=exploited spawning stock biomass, *UB*=unexploited biomass, *USSB*=unexploited spawning stock biomass, *S-R*=stock-recruitment relationship; $\overline{L}$ = average length ; $\overline{age}$ = average age ; *SS*=exploited spawner's population; *USS*=unexploited spawner's population; *C*=capture in numbers; *Y*=yield, *t*=time, *j*=cohort.

5.3 Indicator time-series methods.

5.3.1 Nonparametric method for determining recent trends

Verena Trenkel
Ifremer, Nantes, France

Introduction

The most commonly used method for determining the direction of changes in estimated indicator time series is fitting linear models and then using the sign of the slope if it is significantly different from zero ((e.g.Trenkel and Rochet 2003). This method is reliable for determining long term time trends. However, it is less satisfactory for short term time trends. Trends might not be linear and inter-annual variability in estimated indicators can be strong enough to mask short term changes. Furthermore, if only the trend over the most recent years is considered, the overall evolution is not taken into account, for example whether the indicator values are among the lowest or highest of the available series or whether the indicator has a tendency to fluctuate randomly with a certain phase. Consider the fluctuations of plaice ln-abundance estimated for the Southern North Sea (Fig. 5.7.1, left panel). The population seems to have been fluctuating randomly over the course of the 23 years. So looking at this picture one would probably conclude that the recent years are not any different from the whole time series. However, depending on how many recent years are used for estimating a linear trend, a positive, negative or no trend will be found. Similarly for dab, although ln-abundance has been decreasing in the most recent years, overall the population levels remain well above that at the beginning of the series. So, it seems desirable to include the whole time series in the assessment of the dynamics of the most recent years.

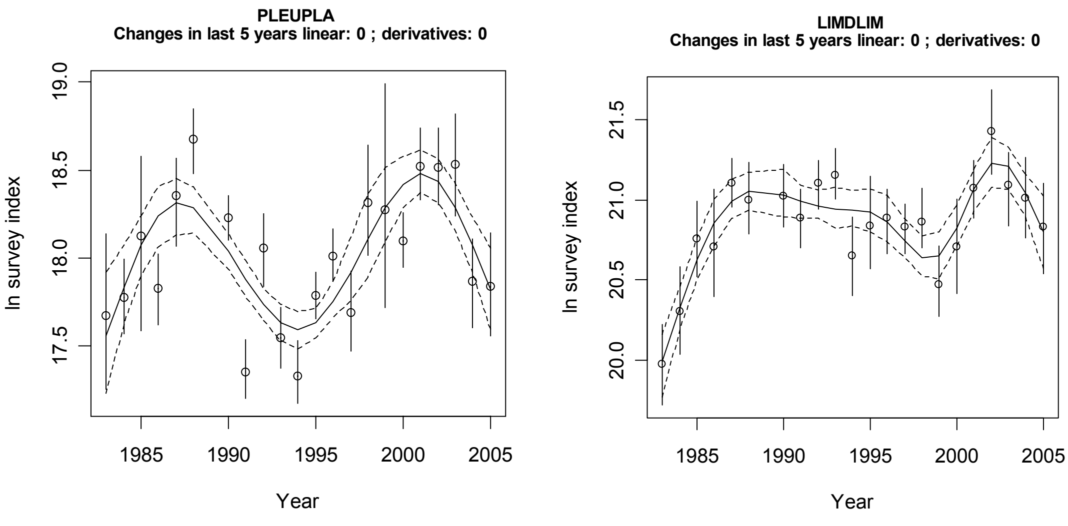


Figure 5.7.1. Time series of ln-abundance for plaice and dab in Southern North Sea based on IBTS data. The continuous line is generalised additive model (GAM) fit. The broken lines are 95% confidence bands for this fit based on a parametric bootstrap of annual indicator estimators.

In this document a method is proposed to estimate the direction of recent changes making use of first and second derivatives of smoothed indicator time series and the position of the most recent years with respect to the full time series. The first derivative, which is actually the local slope (tangent) at each point of a function, here each year, describes the dynamics of the indicator changes. In contrast, the second derivative describes the changes in the slope. A positive second derivative indicates that the slope is increasing while a negative second derivative means that the slope is decreasing. A location at which the second derivative is zero is called a change point as at this point the dynamics change from accelerating to decelerating, i.e. the slope gets smaller from this point onwards, or vice versa. The slope will be zero when either a maximum or minimum is reached. Figure 5.7.2 illustrates the signs of the first and second derivatives. The proposed method is described in details in the next section.

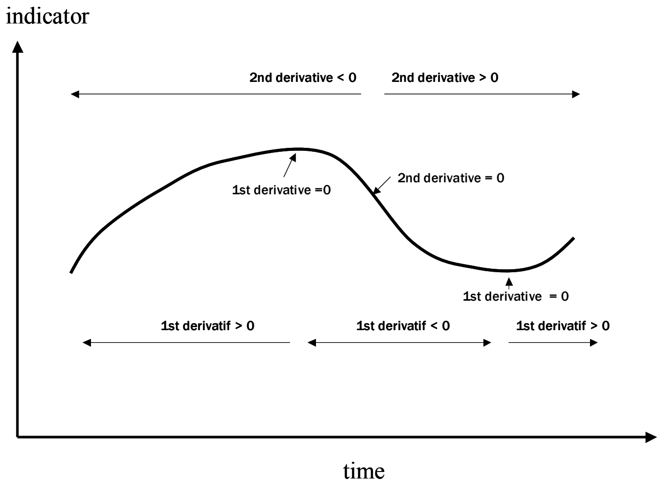


Figure 5.7.2. Diagram showing the signs of the first and second derivative of an indicator time series used for determining the direction of change.

Method

The proposed method consists of several steps which are

1. fit a generalised additive model to the time series in order to obtain a smoothed series;
2. calculate first and second derivatives for the smoothed time series for all years (including years with no data);
3. determine direction of change in recent years using a combination of criteria for the smoothed series as well as the first and second derivatives of the smoothed series.

To obtain smoothed indicator series, generalised additive models (GAM) are fitted with *year* as a cubic regression spline and automatic selection of the degrees of freedom (minimum 3) using the mgcv package in R (R development Core Team 2003) developed by Wood (2000).

As spline models are twice differentiable, first and second derivatives of the smooth series can be calculated for every year of the time series using an approximation based on first and second differences, as used by Fewster et al. (2000).

The first derivative of indicator *I* in year *t* is approximated by the first difference

$$\hat{I}'(t)=I\left(t+\frac{1}{2}\right)-I\left(t-\frac{1}{2}\right)$$

The second derivative is approximated by the sixth difference

$$\hat{I}''(t)=\frac{1}{180}\{2I(t+3)-27I(t+2)+270I(t+1)-490I(t)+270I(t-1)-27I(t-2)+2I(t-3)\}.$$

For the third (t=3) and two before last years the fourth difference is used to approximate the second derivative

$$\hat{I}''(t)=\frac{1}{12}\{-I(t+2)+16I(t+1)-30I(t)+16I(t-1)-I(t-2)\}$$

Finally, for the second (t=2) and one before last the second difference is calculated for the estimating the second derivative

$$\hat{I}''(t)=I(t+1)-2I(t)+I(t-1)$$

For obvious reasons the first and second derivatives cannot be estimated for the first and final year of the time series.

In order to determine whether the estimated first and second derivatives are significantly different from zero for a given year, i.e. $\hat{I}'(t)\neq 0$ and $\hat{I}''(t)\neq 0$ for i=1,...T, a parametric bootstrap is carried out. For this, indicator time series are created by resampling each data point (year) $I(t)^b \sim N(I(t),\sigma(t))$ from a normal distribution with, as mean, the estimated indicator value for year t, $I(t)$ and, as standard deviation, its estimated standard deviation $\sigma(t)$. A separate GAM is then fitted to each bootstrap series $I(1)^b ... I(T)^b$, b=1,...B, using the same degrees of freedom (degree of smoothness) as was found optimal for the original indicator time series. Subsequently, for each bootstrap sample, first and second derivatives are estimated by year. This provides the distribution of first and second derivatives for each year based on which the 2.5 and 97.5 percentiles are calculated. If the value zero is included in the interval formed by the 2.5 and 97.5 percentiles, which is actually a 95% confidence interval, the derivative of the given year is not significantly different from zero and the indicator variable for the derivative is set to zero, otherwise the sign of the derivative is either positive (1) or negative (-1) depending on whether the values within the confidence interval are all negative or positive. The result of this test is a time series of an indicator variable for the first derivative which is either 0, 1 or -1. Similarly for the second derivative.

In order to determine the direction of recent changes in indicator time series, the indicator variables with the signs of the first and second derivatives are combined in a decision rule (Table 5.7.1). In addition, the location of the minimum and maximum value in the time series is used in order to put the most recent years into the perspective of the whole time series. If the maximum is not found within the last three years and the annual slopes (first derivative) are predominantly negative and annual second derivatives are negative or zero in the last five years (no change point appears with sign of second derivative passing from –1 to +1), the direction of change is declared as recently decreasing. The second derivative is used to establish whether an improvement has already taken place most recently. Similarly for a recently increasing series, the minimum should not be within the last three years, the average of the annual slopes should be positive (apart from one year) and no change for a decreasing trend (sign of second derivatives positive) should have occured during the last five years. For all other cases there is no indication for a change. These decision rules are proposed based on empirical tests, however they are by no means prescriptive. The important point is the principle, i.e. the combination of different measures of the dynamics of a time series, minimum, slope and change points. In particular the time spans considered, which is five years for the first and second derivatives and three years for the location of the maximum and minimum, are easily adapted for a particular study.

Table 5.7.1. Decision rules used to determine direction of recent changes for an indicator time series based on first and second derivatives of smoothed indicator time series. y_m = number of years to consider for minimum/maximum; e.g. y_m= 3. y_c = number of years to consider for first and second derivatives, e.g. y_c= 5

Decrease	Increase
1. Maximum value before final y_m years AND 2. Signs of annual slopes for final y_c years negative or at most 0 for 1 year AND 3. Sign of annual second derivatives during final y_c years negative or zero (persistence of decrease)	1. Minimum value before the final y_m years AND 2. Signs of annual slopes during at least final y_c years positive or at most 0 for 1 year AND 3. Sign of annual second derivatives during final y_c years positive (persistence of increase)

For comparison purposes, linear time trends over the whole data series and the last five years are also calculated.

Example: cod in North Sea

As an example the method was applied to the indicator table for cod in the North Sea based on IBTS data. Fig. 5.7.3 gives the smoothed indicator time series. The direction of recent change assessed by two methods is indicated in the header of each figure. The proposed method is referred to as ‘derivatives’. Thus the diagnosis obtained with the proposed method is that ln-abundance and L50 maturity are decreasing, while mean length and the length quartiles are all increasing. Thus all signs points towards a deterioration of this cod stock. In contrast to the proposed method, linear time trends over the most recent five years were only signficant (α=0.05) for the ln-abundance time series.

The total mortality estimates Z and the length quartiles L25 and L75 in Fig. 5.7.3, are varying interannually more than seems plausible biologically. As a consequence the smooth function fits (cubic splines) might not be considered representative for the temporary evolution of these indicators and the resulting diagnoses might be considered unreliable. This example points out the need to carefully select the indicators and to evaluate their reliability before using them for any assessment purposes because the results obtained with the proposed method will entirely depend on the suitability of the GAM fits.

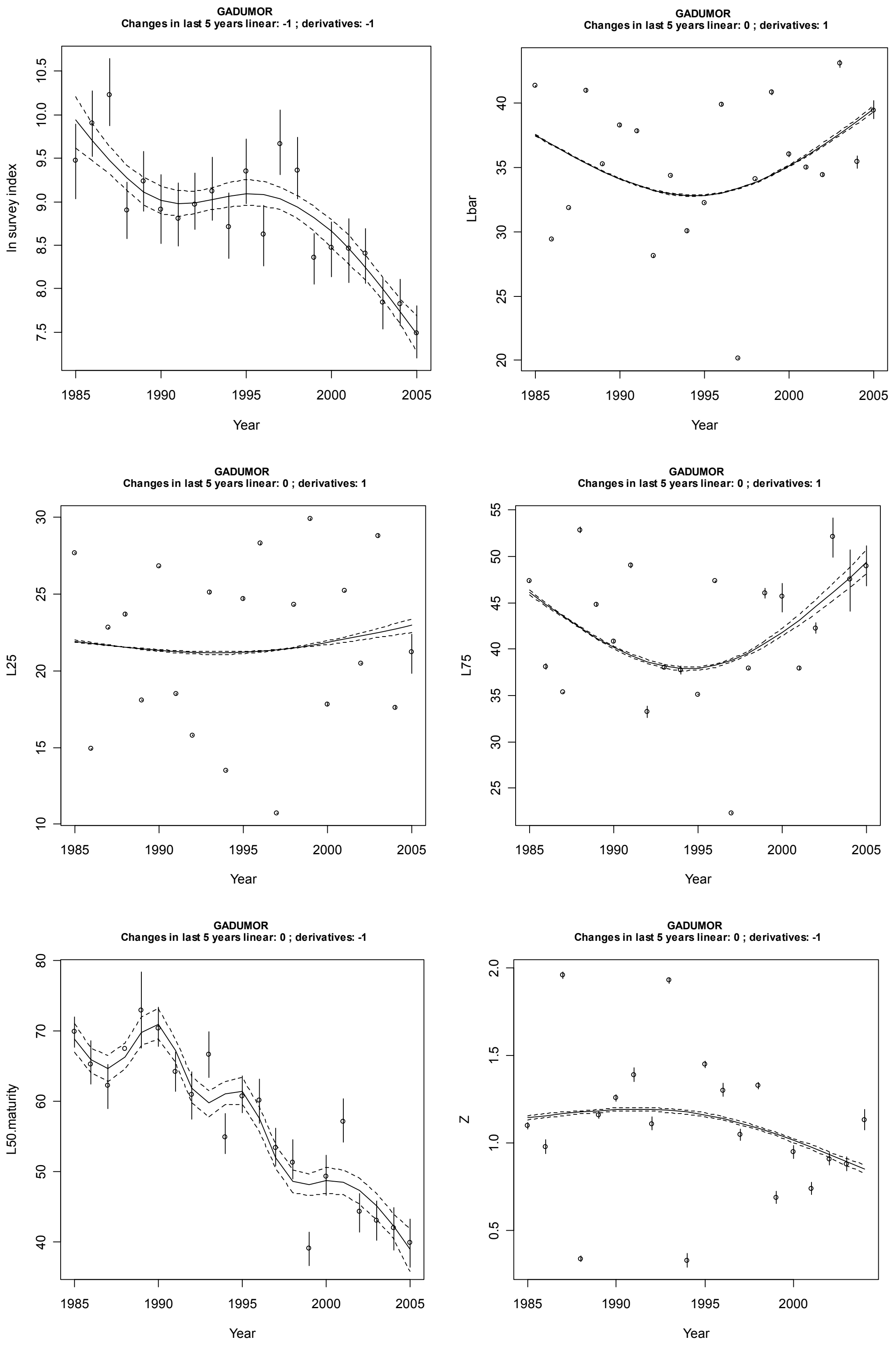


Figure 5.7.3. Indicator time series of North Sea cod with cubic spline model. Assessment of recent direction of changes in figure headers using the proposed method ('derivatives') and linear trend estimation for the final five years.

Software

The software for the non-linear estimation procedure based on first and second derivatives can be downloaded from http://www.ifremer.fr/drvecohal/fisboat/index.htm. To run it:

1. Copy the function "FunctionsTimeChangeDerivatives.R" and script "ScriptTimeChange.R" into same folder as the data table.
2. Edit line 10 in script replacing file name,
   e.g. indicest<-read.table("codNS_tab2_wp2A.txt",header=T,sep="\t",as.is=T)
3. Edit line 28 to select the time horizon for first and second derivatives
   e.g. lastn=5
4. Edit line 31 for time horizon for maximum and minimum values
   e.g. lastnmin=3
5. Run edited script.
6. Results are obtained as a table called Trendestimates.txt and as smoothed time series plots for each indicator

Example results for North sea cod

1. Trendestimates.txt

Area	**Species**	**Indicator**	**LinearSlope**	**PvalueAll**	**LinSlopeLastYears**	**PvalueLast**	**DiagnosLinearRecent**	**DiagnosNonLinearRecent**
NorthSea	GADUMOR	ln_survey_index	-0.067	0.00014	-0.219	0.04792	-1	-1
NorthSea	GADUMOR	Lbar	0.101	0.61485	0.996	0.47505	0	1
NorthSea	GADUMOR	L25	0.051	0.80099	-1.09	0.51169	0	1
NorthSea	GADUMOR	L75	0.161	0.55643	2.73	0.13226	0	1
NorthSea	GADUMOR	L50.maturity	-1.525	0	-3.657	0.06658	0	-1

Explanation of column names:
LinearSlope: Linear slope over whole time series
PvalueAll: p-value for linear slope over whole time series
LinSlopeLastYears: Linear slope for most recent years
PvalueLast: p-value for above linear slope
DiagnosLinearRecent: sign of slope for most recent years if significant (p-value <=0.05)
DiagnosNonLinearRecent: direction of change using proposed method (decrease =-1, increase=1, no change =0)

2. Figures
e.g. ln survey index.wmf (fig. 5.7.4).

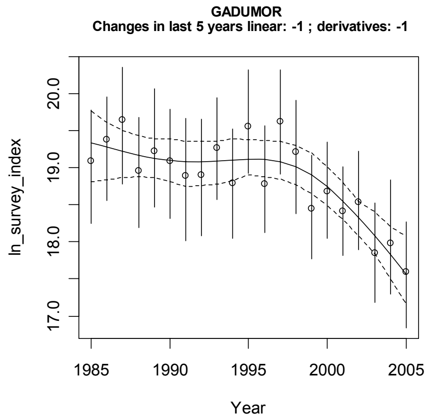


Figure 5.7.4. Ln-transformed survey time series with fitted smoothed model (GAM). Header provides direction of changes as found by fitting a linear slope over the 5 most recent years, and by the proposed method, referred to as 'derivatives'. This information is repeated in the table.

References

Fewster, R.M., Buckland, S.T., Siriwardena, G.M., Baillie, S.R. and Wilson, J.D. (2000) Analysis of population trends for farmland birds using generalized additive models. *Ecology (New York)* **81**, 1970-1984.

Trenkel, V.M. and Rochet, M.-J. (2003) Performance of indicators derived from abundance estimates for detecting the impact of fishing on a fish community. *Canadian Journal of Fisheries & Aquatic Sciences* **60**, 67-85.

Wood, S.N. (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society B* **62**, 413-428.

5.3.2 Assessing the power to detect future trends

Hans Bogaards, Charlotte Deerenberg, Gerjan Piet
Wageningen Institute for Marine Resources and Ecosystem Studies
IJmuiden, The Netherlands

Introduction

An assessment of the power of a survey to detect future trends in candidate indicators is obviously an important task before those indicators are selected for use in some kind of fishery management programme. Formally, the power of a statistical test is the probability of not making a type II error, where a type II error is defined as accepting the null hypothesis when the alternative is true. Within the context of a monitoring program, power can be interpreted as the probability that a particular trend change will be detected. This note describes a method for estimating the power of a survey to detect future, linear trends given standard modelling assumptions. Software in R available from the FISBOAT website is also described.

Method

Power calculation first requires a specification of the testing procedure to be used, together with the significance level α (the probability of making a type I error: rejecting the null hypothesis when it is true) of the test. Furthermore, it requires the definition of a null hypothesis H_0 and an alternative hypothesis H_1.

To provide a generic power calculator for the evaluation of candidate indicators, it is here assumed that the time series of an indicator can be described by a stochastic linear model with an additive normally distributed error term. Observations are assumed to be derived from annual sampling schemes. In the case of missing observations, the analysis is restricted to the longest consecutive stretch of non-missing values in the time series. The slope of the historic trend line is estimated by simple linear regression analysis, which implies that the residual variance is assumed to be constant and has no autocorrelation. The extent to which these assumptions are violated should be judged by the user. Visual inspection of the time series with its trend line and residuals is imperative.

The testing procedure concerns the slope of the trend line for a specified number of future years of follow-up. The test statistic *T* is defined as the difference between the observed slope of the future trend line and its anticipated value under H_0, divided by the standard error of the estimated slope parameter. If the variance of observations about the linear trend will remain constant, *T* will follow a non-central *t*-distribution, the non-centrality parameter being equal to the slope parameter under H_1 minus its value under H_0, divided by its standard deviation under H_0. From this, it follows that the power of a one-sided test is the probability that *T* is more extreme than some critical value *c*. The power of a two-sided test is the overall probability that *T* is more extreme than either *c* or –*c*. Critical values are determined by the significance level of the test. Typically, critical values in a two-sided test with $\alpha = 0.05$ correspond to the 2.5^{th} and 97.5^{th} quantiles of a central *t*-distribution.

Final comment

It is assumed that the residual variance is constant throughout, not only for the duration of the historic time series, but also for the interval over which the power of an indicator is to be evaluated. Transformations to stabilize the variance and to make the trend linear may be considered in advance, after which power should be evaluated on the transformed time series. In addition, the power of a particular indicator will be underestimated if intervention strategies tend to reduce its randomness, or overestimated if the opposite is true.

References

Gerrodette, T. 1987. A power analysis for detecting trends. *Ecology*, 68: 1364–1372.

Nicholson, M.D. and Jennings, S. 2004. Testing candidate indicators to support ecosystem-based management: the power of monitoring surveys to detect temporal trends in fish community metrics. *ICES Journal of Marine Science*, 61: 35–42.

Piet, G.J. and Jennings, S. 2005. Response of potential fish community indicators to fishing. *ICES Journal of Marine Science*, 62: 214–225.

Software

After the function templates have been downloaded from http://www.ifremer.fr/drvecohal/fisboat/index.htm and sent to R, the power of a candidate indicator can be obtained via a call to the function linear.trend(). This function requires at least two arguments:

file	tab-delimited text file to be used, e.g. “codNS_tab2_wp2a.txt”
var	candidate indicator to be evaluated, e.g. “L50.maturity”

If the candidate indicator has been calculated separately for each age-class (for example, the wp2a spatial indicators), a third argument is required:

age	age-class to be evaluated, e.g. “A4”

The name for *var* should correspond to the variable name as provided in the header of the requested text file, whereas the name for *age* should correspond to a value of the variable Age.

Additional arguments that may be set optionally by the user are:

dir	directory where *file* is located (default: working directory)
h0	slope under H_0 (default: continuation of trend line)
h1	slope under H_1 (default: stabilization of trend line)
alpha	significance level of the test (default: $\alpha = 0.05$)
horizon	interval over which power is to be evaluated (default: 25 years)

Arguments acting as character strings should be enclosed in quotation marks, e.g.

```
> linear.trend(file="codNS_tab2_wp2a.txt", var="L50.maturity", h1=0, alpha=0.01)
```

Power is calculated both for one-sided and two-sided tests. In one-sided tests, an increased slope parameter is anticipated if the historic trend line was decreasing and vice versa. By doing so, ecosystem-based management objectives are evaluated more efficiently than in two-sided tests, as the latter are more conservative.

Two graphs are output by the R script. The first shows the time series that is used, together with the best-fitting linear trend line, and residuals which are plotted separately. The second shows the results of the power calculations, for one-sided and two-sided tests. Figures are also printed to the screen, with oc1 and oc2 denoting one-sided and two-sided test results, respectively.

5.3.3 Statistical Process Control (SPC) schemes

Benoit Mesnil and Pierre Petitgas
IFREMER, Nantes

Introduction

Control charts are part of the statistical process control (SPC) tools routinely used over decades to monitor manufacturing processes and signal anomalies in performance. The process has some inherent variability and is said to be 'in-control' as long as it remains within acceptable bounds. If an anomaly occurs causing a deterioration in quality beyond the baseline variability the system is said to be 'out-of-control'. Control charts are graphical displays of some summary statistic of the observation data (e.g. an indicator) against the order index of the sample (e.g. time), together with reference 'marks' based on the in-control mean and variance, that are designed to detect whether a worrisome change in process output is indicated by the current data and a fix is required. Since there are costs associated with both false alarms and quality losses, the charts' parameters are tuned to achieve a desired trade-off between the risk of false alarm and the power to detect changes promptly.

The cumulated sum type of control chart in its 'decision interval' form (DI-Cusum) has been selected for this project because it is advocated in SPC textbooks (e.g. Montgomery, 1991; Hawkins & Olwell, 1998) for the type of data considered in the survey indicators context. Control charts can be designed to monitor changes in mean level (location charts) or in variance (scale charts) of process outputs; explanations are only given for location charts here. More details can be found in the literature digest made for this project by Mesnil & Petitgas (WD) and in the papers cited therein.

Derivation

Suppose a suite of observations (individual or group means) x_i collected at time $i = 1, \ldots m$ and assume that their in-control mean μ and standard deviation σ_x are known from a pilot study or for a reference period. In the following, it is considered that the data are first standardised through the transformation $z_i = (x_i-\mu)/\sigma_x$.

The decision-interval Cusum works by recursively accumulating positive and negative deviations separately with two statistics:

$$S_i^+ = \max\left[0, S_{i-1}^+ + z_i - k\right]$$

for positive deviations ('one-sided upper Cusum'), and

$$S_i^- = \min\left[0, S_{i-1}^- + z_i + k\right]$$

for negative deviations ('one-sided lower Cusum'), with starting values normally set as $S_0^+ = S_0^- = 0$. A Cusum chart is obtained by plotting these statistics against i.

The parameter k is usually called the *reference value*, or the *allowance*, and is related to the size of the smallest shift in the level of x that one is wishing to detect quickly. Note that deviations smaller than k are ignored in the recursions above. The decision rule is to declare an out-of-control state whenever S^+ exceeds the *decision interval h* or S^- falls below $-h$. The values chosen for the parameters h and k (in standard deviation units) determine the performance of the control chart; there is no theoretical objection against setting different h-k pairs for upper and lower Cusum's if changes in one direction matter more than in the other.

The performance of control charts is generally evaluated in terms of their run length. A run is the number of sampling events that elapse between the start of the monitoring and the first alarm. Run length is a random variable whose probability distribution depends on the process and the chart parameters, and it is its expectation – called Average Run Length (ARL) – that is commonly used as a summary measure of performance. In many instances the run length distribution is very broad and skewed, and it may be misleading to only consider its mean; the experts recommend to also look at other percentiles, whenever the distribution can be computed. The notation ARL(δ) is used to designate the ARL of an SPC scheme for detecting a change of size δ (in σ_x units) occurring in the process mean level. Thus, ARL(0) is the ARL of a scheme when the process actually stays in-control all the time (in-control, or IC ARL); yet, due to its inherent variability, an alarm may be raised by chance alone when the chart is updated with a new datum. In other words, ARL(0) is the average time until a false alarm is raised, which should ideally be large. Conversely, if the mean of the process distribution shifts from μ to $\mu+\delta$, due to an anomaly the chart should detect this quickly, implying a short ARL(δ). Chart parameters can be tuned to achieve the desired compromise, as explained in the guidelines below.

Software

Two R scripts have been developed to implement a Cusum monitoring scheme: *CusumTutorial.r* is generic, for exploring Cusum charts with 'free-format' time series vectors; *FBCusumCharts.R* is designed to automate the production of standard tables of results for the report ('traffic light template'). Both use a set of functions stored in the separate file *CusumFuncs.r* that must be sourced into the user's R workspace (on first use) as instructed in the scripts. The scripts are meant to be run in a stepwise fashion (highlight a line or a block and submit to R) and are amply commented to guide the user.

The top part of *FBCusumCharts.R* deals with each indicator in turn. Note that a logarithm transformation is applied to the Survey and Recruit indices (columns 5 and 6); the reference period for each case study is 'hard-wired' but can be edited if needed; an indication of an appropriate value for the allowance *k*, based on the mean deviation from the reference mean outside the reference period, is proposed but is not coded as a default value. Once the full set of indicators has been processed, the bottom part of the script gathers the individual *resnam.#* objects to produce the table of alarms (signed Cusum values above *h* or below *–h*) and the table of Cusum parameters and saves them to files.

This implementation includes functions to compute in-control or out-of-control ARLs and run length distributions of one-sided Cusum for normal data, adapted from a Fortran code by F.F. Gan (1993) found on the StatLib JQT archive. They have been checked against the values tabulated in various SPC textbooks and articles, and the results match very well. They need to be optimised for R, to speed up the computation of RL distributions which requires some patience at the moment.

Practical guidelines

A) Cusum design: tuning the chart parameters k *and* h

With fisheries survey data, we in general have to analyse time series of one or several indicators of population status (control variables). We have one value per indicator per year (individual data) with perhaps the precision on the indicator in each year. We distinguish 2 phases: Phase I for defining the in-control (IC) period; and Phase II for designing the Cusum to signal change from the in-control state with desired performance.

Phase I

The task in Phase I is to set the IC or reference process parameters μ and σ. It is a critical phase in that the values adopted for these parameters will condition the diagnostic that will be made later. Normally, this is an experimental phase where the state of the system is closely checked, many measurements are taken and scrutinised, to retain only those that can be safely assumed to correspond to a well behaved process. In our case, we will often start with existing data collected in the past, and Phase I will essentially consist in the definition/choice of an in-control (or reference) period, and using the data in the selected to years to estimate in-control parameters μ and σ. The IC period can be defined on various criteria, including an analysis of the times series. The IC period is best defined collectively, on expert knowledge, as the period in which the population was in a satisfactory state and/or showing satisfactory dynamics. For example, within the Fisboat project, the IC period was defined collectively during a workshop as the period when the indicator value showed "satisfactory" values with low variation (no obvious outlier). Thus, the IC period may not necessarily comprise consecutive years. Sensitivity to the IC period should be analysed and the IC period may be also re-defined *a posteriori*. This is consistent with the iterative and rejection procedures described in SPC textbooks.

Phase II

In Phase II the task is to design (or tune) the Cusum scheme to signal a specified deviation from the IC mean with a desired performance, i.e. this is where the chart parameters *k* (allowance) and *h* (interval) are determined. The choice of *k* is based on the magnitude of the shift δ in the mean that makes "a meaningful impact" on the system, driving it out of control. The value of *h* determines whether an alarm is raised or not (an alarm is triggered when the cusum plot crosses the horizontal line at *h*, or -*h* or +*h* for a two sided Cusum). The rationale for choosing *h* is primarily based on minimising the risk of false alarm, but the ability to promptly detect shifts that matter should also be preserved. Setting *h*, once *k* is chosen, involves Run Length considerations. A four-step procedure is suggested[1]:

1. Regarding *k*, if δ is the shift of interest (in sd units), there is broad support in the literature for setting *k* at half the value of that shift (formal demonstration in Chap. 6 of Hawkins & Olwell), and this rule can be safely adopted. The "meaningful" shift δ can be set after analysing the deviations from μ outside the IC period. For instance, the shift to be detected can be set to a percentile of these deviations or to their mean. For fisheries survey based population indicators, *k* will take in general a value between 0.5 and 1.5; too small values of *k* should be avoided (Hawkins & Olwell, p. 33).
2. Using tables or software with a zero value for the shift δ, search for an *h* that gives desirably large IC ARL(0) given *k*, and thus a low risk of false alarm. Larger values of *h* (and *k*) lead to larger ARLs.
3. Because the RL distribution may be quite skewed, consideration of the average RL alone may be misleading and, using the function `arldis.f` in the Fisboat R scripts, the full distribution of the in-control RL should be checked. For example, if you choose *k* and *h* to aim for a "large" IC ARL of 100, and observe a "small" value of 10 samples or less for the 25th percentile, it is likely in the actual application of the scheme that more false alarms will occur than the large ARL(0) makes you think. If so increase *h*.

[1] Reminder: in all this, we assume the indicator series have first been standardised.

4. The h value arrived at in the previous steps may have made you reasonably content with the false alarm risk, but you then need to ascertain that the scheme is able to quickly detect the shifts you are interested in, i.e. that its out-of-control ARL is small enough. Return to the ARL tables or software with a value of $2k$ for the shift δ, and check that the OC ARL(δ) corresponding to the envisaged h is adequately small. In general, it is possible to find values of k and h so that the OC ARL does not exceed 3 years.

As pointed out earlier the basic challenge of using and tuning SPC schemes is to find an acceptable compromise between the risk of false alarm and the power to detect shifts that matter in the state of the system, and it is often necessary to iterate through these 4 steps to arrive at that compromise. The notions of "meaningful impact", "acceptable risk" and "desired performance" are very much policy issues and have to be decided in partnership with managers and stakeholders.

B) Assumptions and effects of violations

The main assumptions underlying the statistical properties of Cusum charts are (i) that the monitored variable has a distribution from the exponential family; in particular, the run length characteristics commonly tabled in textbooks or computed with the R software coded for this project are only valid for normally distributed data; (ii) that the in-control process parameters are known rather than estimated; and (iii) that the time series of residual variation has no correlation in time. Violations of these assumptions all go in the same direction: the in-control ARL(0) experienced in practice is shorter than the value computed for the perfect case, i.e. the chances of false alarms are larger than expected (e.g. Section 3.7 in Hawkins & Olwell; Jones et al., 2004; Lu & Reynolds, 1999; Reynolds & Stoumbos, 2004). Smaller values of k (also large h) enhance the robustness to non-normality, but increase the impact of estimating the reference mean and sd from the data. An encouraging note: even though a Cusum tuned with a given k is optimal for detecting shifts of $2*k$ standard deviations, its performance remains high for actual shifts that are 'not too far' (Hawkins & Olwell, p. 54).
Time series of survey data population indicators are often short (< 20 years) with marked deviations and sometimes show correlation or trend. It is advised to check the distribution of the indicator variable as well as its correlation in time. It may be necessary in some cases to transform the variable into a Gaussian or to detrend the time series. The reference period is even shorter, and we use noisy data to estimate the IC process parameters. Since all departures from the assumptions will result in effective RLs being very different (in general shorter) than values publicised for the "clean" case, an ad hoc remedy is to take relatively large h values. Conservative advice is to use (k,h) parameters giving large IC RLs: ARL > 20 years and 25th percentile of RL distribution > 10 years. When some deviations from μ outside the reference period are large in comparison to σ, it may be telling that the variance has changed or that the indicator variable is skewed. In that case, starting Phase II with a large value of k is advisable. When the value of h is small in comparison to an increasing (decreasing) Cusum deviation, it may be telling that there is correlation in time in the indicator series.

C) Strengths and weaknesses

Control charts have been in operation in many branches of industry since the 1930's and their statistical bases have been thoroughly investigated in a huge body of literature (the references below are just a tiny sample).They are still a recurrent topic of specialised journals such as the Journal of Quality Technology or Technometrics. Applications have been extended to environmental surveillance, biomedicine, clinical tests, and public health. The strengths in these domains are that the in-control state is well defined, the monitoring involves numerous samples taken at high frequency through rigorous sampling designs, and measurement errors are often small.

In contrast, this defines the weaknesses for fisheries applications. Perhaps the main limitation is our poor ability to characterise the reference state of fisheries (or of ecosystems) with survey data that just span the recent decade(s) in a background of large variability compounded by substantial sampling variance (i.e. we do not have a proper Phase I). Keep in mind, however, that the reference state does not imply perfect stability; the goal of control charts is to spot those events where the state of the system jumps beyond the domain of its inherent variability.

A virtue of the Cusum approach is that it does not presume the nature of the change (linear, trend or otherwise) and treats positive and negative deviations equally. Cusum charts are best suited to detecting small, persistent changes. Anomalies in the system can take the form of shifts in the mean and/or changes in the variance of the distribution. Specific control charts can deal with both situations. Actually, it is common to combine location and scale charts to enhance the detection performance for both small and large shifts (Reynolds & Stoumbos, 2004).

It has been demonstrated that, among the procedures that have similar in-control ARL(0), the Cusum has the smallest expected time until a change is detected when it occurs. This is the basis of the rationale for tuning the chart, with priority given to achieving large ARL(0). The emphasis on low risk of false alarm has some practical advantage in our application to fisheries management and its overly controversial atmosphere; we have learnt to know that casting assessment noise straight into fisheries regulations has damaged our credibility and our relations with the industry, and a method that explicitly aims to avoid this should help.
Lastly, the biggest advantage of the Cusum is that it is so simple to implement. Yet, it provides a formal framework to establish diagnostics in an objective and replicable way.

References

The oft-quoted textbooks

Derman, C. & S.M. Ross. 1997. *Statistical aspects of quality control*. Academic Press, London.

Hawkins, D.M. & D.H. Olwell. 1998. *Cumulative sum charts and charting for quality improvement*. Springer-Verlag, New York.

Manly, B.F.J. 2001. *Statistics for environmental science and management*. Chapman & Hall/CRC, Boca Raton.

Montgomery, D.C. 1991. *Introduction to statistical quality control*. Wiley, New York. (5th Edition 2005).

Whetherill, G.B. & D.W. Brown. 1991. *Statistical process control: theory and practice*. Chapman & Hall, London.

Articles on SPC methods

Gan, F.F. 1993. The run-length distribution of a cumulative sum control chart. J. Quality Technology 25: 205-215.

Jones, L.A., C.W. Champ & S.E. Rigdon. 2004. The run length distribution of the CUSUM with estimated parameters. J. Quality Technology 36: 95-108.

Lu, C.-W. & M.R. Reynolds Jr. 1999. Control charts for monitoring the mean and variance of autocorrelated processes. J. Quality Technology 31: 259-274.

Luceño, A. & J. Puig-Rey. 2002. Computing the run-length probability distribution for CUSUM charts. J. Quality Technology 34: 209-215.

Mesnil, B. & Petitgas, P. (WD). Statistical Process Control (SPC) schemes for indicator-based assessments. A literature review and manual for the Fisboat project, 2006. 30 pp.

Page, E.S. 1961. Cumulative sum control charts. Technometrics 3: 1-9.

Reynolds, M.R. & Z.G. Stoumbos. 2004. Control charts and the efficient allocation of sampling resources. Technometrics 47: 409-424.

Scandol, J.P. 2003. Use of cumulative sum (CUSUM) control charts of landed catch in the management of fisheries. Fish. Res. 64: 19-36.

Woodall, W.H. & B.M. Adams. 1993. The statistical design of Cusum charts. Quality Engineering 5: 559-570.

5.3.4 Nonparametric statistical methods for assessing trends.

John Cotter
CEFAS, Lowestoft

Introduction

This note describes a selection of nonparametric statistical methods thought to be useful for assessing trends in fishery statistics or indicators, e.g. abundance-at-age, mean length, geostatistical indices, or just about any continuous variable. The trends referred to here are assumed to relate to time but they could also relate to a transect over a spatial dimension. The literature on trends is extensive so this note can only provide a modest introduction to it. Loftis et al. (1991b) point out that formal statistical methods do not usually reveal trends that are not apparent from inspection of the data but they are useful for allowing different data analysts to reach similar conclusions from the same data and assumptions. Several of the references cited come from the literature on monitoring of water pollution where the sporadic and chaotic nature of variation combined with frequent gaps in the time-series has stimulated development of nonparametric methods because of their minimal assumptions. Fisheries scientists typically prefer modelling, i.e. parametric methods for assessing trends in fish stocks and have exploited nonparametric methods relatively lightly. All the same, trends in fish stocks could be established with less reliance on assumptions about the data and models if nonparametric methods were used. Furthermore, interest nowadays is shifting from estimation of quantities of fish in a single commercial stock to assessment of whole ecosystems, a task for which well established, structural models are not always available.

There has been little discussion in the literature of statistical inference in relation to trends. This note therefore begins by proposing some points thought to be important. A variety of nonparametric statistical tests tailored for assessing trends is then introduced, some of which are easy to calculate with a spreadsheet but limited to providing only the most general statements, e.g. the binomial test with the median, and others which are more elaborate and specific, e.g. Mann-Kendall's Tau which finds monotonic trends. Multivariate tests presented include Cochran's *Q* and the the Dietz-Killeen test. A spreadsheet accompanying the paper illustrates application of each method to a single test set of data, namely a set of abundance-at-age figures for cod from the North Sea IBTS quarter 1 survey. There were few problems in Excel for the univariate tests once the relevant functions had been discovered (e.g. RANK(), MEDIAN(), and BINOMDIST()) although some methods were quite labour intensive. The spreadsheet can be downloaded from http://www.ifremer.fr/drvecohal/fisboat/index.htm. Alternatively, the methods could easily be implemented in R (and several already are). R code is available from the same site for the Dietz and Killeen multivariate trend test.

Inferring about trends

A distinction is acknowledged here between the true, unknown trend, called the signal, and the measures of it made with error, called observations. Most analyses of trends have to be based on the following assumptions: at all times, t,

- E(measurement error) = 0
- E[(measurement error at t) * (measurement error at $t+\Delta$)] = 0 where Δ is any lag interval, and
- E(measurement error * signal) = 0.

$E()$ is the operator for statistical expectation. In words, measurement errors should average to zero, have no serial correlation, and be independent of the level of the signal. Failure of any of these assumptions could lead to spurious trends unrelated to the signal.

"Trend" is hard to define more specifically than our intuitive understanding of a general movement up or down of an observed variable. A trend can occur in the observations, in the signal, or in both. Usually we think of a trend as monotonic upwards (or downwards), i.e. with every observed value at time t equal to, or higher (or lower) than that at t-1. However, turning points, either real or error-based, are likely to occur too. Their occurrence in an observed series is not necessarily, by itself, an accurate indicator of the position of the turn in the signal, or of its magnitude at that point. For an explanatory analogy of this, consider flying over a mountain range and dropping weights at fixed intervals without looking where; some might fall on high ground but few, if any, will fall exactly on the turning points of height, e.g. the mountain peaks. Step changes can also occur in time series and may look like smooth trends when obscured by observation errors. Nonparametric methods for inferring the location and magnitude of a step are discussed by Pettitt (1979). The Mann-Whitney nonparametric test is another option when the location of the step is known (Lettenmaier 1976). The binomial test, see below, would be even simpler.

Statistical tests for trend are affected by the statistical approach adopted. There are two accepted ways of thinking about time-series:

(i) *design-based*: the signal is assumed to be unique and fixed over any defined interval of time; the results of a survey depend on its design.
(ii) *model-based*: the signal is assumed to be one of many possible realisations over that interval; results of a survey depend on the model fitted to the data.

Under design-based thinking, the null hypothesis of no trend, meaning exactly equal values of the signal at all observation points, would seldom be plausible for fisheries data unless the locations of observation were extremely close, or all possible causes of variation temporarily ceased to exist. The analogy of the rocky mountain range is again applicable – no two observation points along a transect are likely to be at exactly the same height. Provided that there are enough observations, and measurements are made accurately enough, statistically significant differences in value will be discovered even though these may not be significant in practical terms (Loftis *et al.* 1991b). Model-based thinking comes from the other conceptual direction by assuming that a signal should be assumed to be horizontal until evidence indicates otherwise. The analogy here is of a randomised experiment in which subjects from one defined population are assigned randomly to treatments so that, if the treatments have no effect, the null hypothesis of equal means in each experimental group is readily plausible. This brings in the concept of the statistical power of a test for trend (Lettenmaier 1976; Nicholson and Fryer 1992).

The implications of serial correlation are also affected by whether the approach is design- or model-based. In the first case, the distinction between serial correlation and trend is undefined. Serially correlated values can look like a trend when observed through a narrowed time window, and, vice versa, a trend can look like serial correlation when observed through a widened window. Either situation could cause rejection of the "no trend" hypothesis. Under model-based inference, serial correlation invalidates nonparametric tests that are based on the null assumption that all permutations of values around a horizontal signal are equally likely. Serial correlation can be decreased by increasing the time intervals between observations, or by modelling the serial correlation and subtracting the estimates from the series (Lettenmaier 1976). Seasonal or other cyclical trends add to the complications since the seasonal trends may themselves move independently over years (Van Bell and Hughes 1984), and measurement errors could be serially correlated from one season to the next (Zetterqvist 1991; El-Shaarawi and Niculescu 1992). Several nonparametric methods for seasonal trends are available (Hirsch *et al.* 1982; El-Shaarawi 1993; Esterby 1993; Yu *et al.* 1993) but they are not considered in detail here since most fish survey data are annual.

Monotonic trends might appear *linear* or *curvilinear*. These are easily modelled, of course, but, with the design-based approach, any structural model of the pattern could only be postulated as a rough approximation to the signal from a natural system.

The design-based approach is preferred here for its plausibility, because avoidance of modelling is consistent with the simplicity underlying nonparametric statistics, and because it relieves the analyst of many assumptions associated with model identification and fitting, thereby offering a genuine alternative to modelling. To be consistent with the design-based approach, I suggest replacing the term "hypothesis" with the word "notion" when describing the signal as having no trend or a specific type of trend so as to be clearer about the informality of a test in these circumstances. Trends can be estimated together with confidence limits without testing the usually untenable notion of 'no trend'. Alternatively, the more reasonable null hypothesis (H) : "trend ≤ 0" can be tested with nonparametric methods against the alternative (A) : "trend > 0". This one-sided H encompasses a region of probability, not a point. [Technically, it is a 'composite hypothesis' (Brownlee 1965).] If true, it would not be rejected by a sample, however large or precise, except by chance with probability α, as expected for a statistical test. Usually, this null hypothesis would be the most sensible choice for a test because a one-sided test is consistent with a prior concern that the trend is in one direction. Applying a two-sided test for either a positive or negative trend could suggest that the data are being mined unscientifically for any detectable feature.

A special problem with assessing trends is that they are often noticed in graphical plots before they are tested statistically or confidence limits are fitted. Bearing in mind that trends are often visible in series of random numbers (Kendall 1976), the application of statistical methods *a posteriori* could be misleading. Equally risky is when the terminal points of a trend are decided by inspection. Questions of the type "Is this variable going up or down?", e.g. for the purposes of controlling environmental quality, should be completed with "since when" before assessing statistically because the probability of a trend is likely to depend on the chosen starting point, as well as the end point if not the final observation. If the interest lies in cause and effect, the recommended plan is to decide by prior reasoning when a trend might occur and in which direction, then to apply statistical methods to test whether the trend is present. If it is, linking it with a putative cause in a matching time-frame might be reasonable as a cautious, on-going hypothesis. Loftis et al (1991b) point out that trend analysis cannot establish cause and effect relationships.

Nonparametric statistical methods for trends

1. Example data

Table 5.13.1 shows abundance (N per hour) indices for North Sea cod as found by the ICES International Bottom Trawl quarter 1 surveys from 1976 to 2004 at ages 1 to 6, except that results for ages 3 to 6 were missing in the earlier years. These values were taken from a report of the ICES working group on fish stocks of the North Sea and Skagerrak. They will be used to illustrate application of various nonparametric statistical methods. Some use the full time-series; others have to use only 1983 to 2004 when all age classes were determined.

Table 5.13.1. International bottom trawl survey (IBTS) quarter 1: time series of abundance indices (numbers caught per hour) for cod in the North Sea in 6 age classes. -1 = missing value.

Year	Age1	Age2	Age3	Age4	Age5	Age6
1976	7.9	19.9	-1	-1	-1	-1
1977	36.7	3.2	-1	-1	-1	-1
1978	12.9	29.3	-1	-1	-1	-1
1979	9.9	9.3	-1	-1	-1	-1
1980	16.9	14.8	-1	-1	-1	-1
1981	2.9	25.5	-1	-1	-1	-1
1982	9.2	6.7	-1	-1	-1	-1
1983	3.9	16.6	2.7	1.8	0.8	1.5
1984	15.2	8	3.9	0.9	1	0.9
1985	0.9	17.6	3.5	1.7	0.5	1
1986	17	3.6	6.8	2.3	1.3	1.1
1987	8.8	28.8	1.4	1.7	0.6	0.9
1988	3.6	6.1	5.8	0.6	0.9	1.1
1989	13.1	6.3	5	2.3	0.4	1
1990	3.4	15.2	2	1	1	0.8
1991	2.4	4.1	3.4	0.8	0.4	0.8
1992	13	4.5	1.2	1	0.3	0.5
1993	12.7	19.9	2	0.7	0.6	0.4
1994	14.8	4.4	3	0.8	0.5	0.5
1995	9.7	22.1	2.8	1.1	0.3	0.3
1996	3.5	8	6	0.7	0.6	0.4
1997	40	6.9	2.3	1.1	0.4	0.4
1998	2.7	26.4	2	0.9	0.5	0.4
1999	2.1	1.6	8.1	0.8	0.5	0.5
2000	6.6	3.8	0.7	2	0.4	0.5
2001	2.8	8.7	1.7	0.2	0.4	0.3
2002	7.8	3.4	4.3	0.5	0.1	0.2
2003	0.6	3	1	1.4	0.4	0.3
2004	7.5	1.3	1.2	0.30	0.4	0.01
Median	7.9	8.0	2.75	0.95	0.5	0.5

2. Quantiles and binomial methods

A time-series may be characterised most basically by its median value +/- binomial confidence limits. The latter are found by firstly ranking the observed values, then finding the ranks, conventionally shown in brackets as (a) and (b), with cumulative binomial probabilities nearest to the required confidence limits, e.g. 2.5% and 97.5% for the case of limits of approximately 95%. Limits exactly at some preset, rounded percentage are seldom possible with the binomial distribution. Binomial confidence limits for the median, $\tilde{\mu}$, of a variable X are obtained with

$$\Pr\left\{X_{(a)} \le \tilde{\mu} < X_{(b)}\right\} = \sum_{i=0}^{b-1} B\{N,0.5,i\} - \sum_{j=0}^{a-1} B\{N,0.5,j\}$$

where $B\{\ \}$ is the binomial probability function for sample size N and probability of 'success'=0.5 (for the median). See Connover (1971) for more details. The binomial probabilities are based on an assumption that the observations fall independently and randomly to either side of the median for the tested time period, a questionable assumption if a trend is present. [They can be computed in an MicroSoft Excel spreadsheet with the BINOMDIST function.] Applying these formulae to an example subset of the abundance indices shown in table 5.13.1, the median index for age 1 cod from 1976 to 2004 was 7.9 with 93.86% confidence limits of 3.5 and 12.7 if there was no trend. These correspond to ranks a = 9 and b = 20 with cumulative binomial probabilities of 0.0307 and 0.969 respectively. Binomial confidence limits can also be estimated for more than one percentile simultaneously, e.g. the 10, 50, and 90 percentiles (Cotter 1985).

The binomial distribution can be used to test H : "trend ≥ 0" against A : "trend < 0" by assuming only that the estimated median is close to the true median for the whole tested period. This is a very simple test to carry out but it would often miss trends that would be detected by more elaborate methods. Four quadrants are formed by intersection of the estimated median observed value with the median of the observation times, the latter being the vertical line half way through the observed series. The test could also be applied to look for a step change; the vertical line would then be located at the time when the step change is expected. Each observed value and its associated time of observation is then classified by quadrant. The null hypothesis, H, implies that observations will fall equally into each quadrant or that there will be more in the lower left and top right quadrants. Suppose that x out of N observations fall in either the top left or bottom right quadrants, implying A, a downward trend. The probability that H is true is

$$\Pr\{H = true\} \leq 1 - \sum_{i=0}^{x} B\{N, 0.5, i\}.$$

[The '≤' would be '=' if H only represented independence of the four quadrants.] Calculations for all ages are shown in table 5.13.2. With the probability of rejecting H given that it is true being α=0.05, we would accept downward trends in abundance for cod of ages 2, 5, and 6. Note however that this is a set of univariate tests, so use of a lower value of α might be preferred to allow for the increased possibilities of type 1 errors in multiple tests. A simple, if conservative, way to achieve this is with the Bonferroni inequality (Prins 2006); when conducting $m = 1,...,k$ tests, set $\alpha_m = \alpha/k$. In this case, $\alpha_m = 0.0083$ implying that downward trends should only be accepted for the 5 and 6 year-olds. Binomial tests can also be applied to assess compliance with an ecological quality objective set as a quantile other than the median though larger sample sizes tend to be necessary to find statistical significance (Cotter 1985). Compliance testing with multiple objectives set as quantiles is further discussed by Cotter (1994).

Table 5.13.2. Binomial test for a trend relative to the median value for the abundance indices shown in table 5.13.1. x = observed value, t = observation time; 0 indicates index ≤ median ;1 indicates index > median.

Year	**Age1**	**Age2**	**Age3**	**Age4**	**Age5**	**Age6**
1976	0	1				
1977	1	0				
1978	1	1				
1979	1	1				
1980	1	1				
1981	0	1				
1982	1	0				
1983	0	1	0	1	1	1
1984	1	0	1	0	1	1
1985	0	1	1	1	0	1
1986	1	0	1	1	1	1
1987	1	1	0	1	1	1
1988	0	0	1	0	1	1
1989	1	0	1	1	0	1
1990 = median (1976-2004)	0	1	0	1	1	1
1991	0	0	1	0	0	1
1992	1	0	0	1	0	0
1993 = median (1983-2004)	1	1	0	0	1	0
1994	1	0	1	0	0	0
1995	1	1	1	1	0	0
1996	0	0	1	0	1	0
1997	1	0	0	1	0	0
1998	0	1	0	0	0	0
1999	0	0	1	0	0	0
2000	0	0	0	1	0	0
2001	0	1	0	0	0	0
2002	0	0	1	0	0	0
2003	0	0	0	1	0	0
2004	0	0	0	0	0	0
Number of x > median and t <= med(year)	9	9	6	7	7	9
Number of x < median and t > med(year)	9	10	6	7	10	11
Number of x	29	29	22	22	22	22
Binomial probability (H: trend >=0)	0.068	0.031	0.262	0.067	0.002	<0.001

3. Cochran's Q test

Cochran's Q (Cochran 1950) tests the notion that the probabilities of response are the same (*H*) in different groups or, alternatively (*A*) detectably different. The test can be adapted to look at multiple trends with the aim of avoiding the problem of multiple univariate tests mentioned in connection with binomial tests, above. This is illustrated in table 5.13.3 using the abundance indices for 1 to 6 year-olds from 1983 to 2004. Age classes of cod are treated as groups, and each observed value is marked as a response, i.e. with a 1, if above the median value and located on or before the median time, or if below the median value and after the median time. Otherwise it is marked as a non-response, i.e. with a 0. In other words, each value gets 1 if it is consistent with a downward trend, and 0 otherwise. The markings are shown in table 5.13.3. Let T_j be the column sums in the *j*'th age class, $j = 1,\ldots,C$, and $\overline{T}$ the mean of them. Let u_i be the *i*'th row sum, $i = 1,\ldots,R$. Then Cochran's statistic is defined as

$$Q = \frac{C(C-1)\sum_j \left(T_j - \overline{T}\right)^2}{C\sum_i u_i - \sum_i u_i^2}.$$

Q is distributed as χ^2 with $(C-1)$ degrees of freedom (Brownlee 1965, section 7.10) under *H*. Note that Q is not sensitive to the total number of responses (since many of the u_i may be zero) hence, for our purposes, it does not by itself establish whether or not an overall trend is present. For the example, $Q = 3.56$ which is much less than $\chi^2(5) = 11.07$ implying that the different age classes are not showing detectably different trends, given that a general downward trend exists. Cochran's Q appears to have similarities with Friedman's rank test for blocked data (Brownlee 1965), and with van Belle and Hughes' test for homogeneity of seasonal trend (Van Bell and Hughes 1984). El-Shaarawi (1993) suggests ways of extending the latter method to testing the notions that linear or quadratic patterns exist in the seasonal trends.

Table 5.13.3. Cochran's Q test applied to age classes 1 to 6 from 1983 to 2004 for the abundance indices shown in table 5.13.1 illustrating how observed abundance indices are marked. 0 means index ≤ median and 1 means index > median if year ≤ 1993, and vice versa if year > 1993.

Year	Age1	Age2	Age3	Age4	Age5	Age6
1983	0	1	0	1	1	1
1984	1	0	1	0	1	1
1985	0	1	1	1	0	1
1986	1	0	1	1	1	1
1987	1	1	0	1	1	1
1988	0	0	1	0	1	1
1989	1	0	1	1	0	1
1990	0	1	0	1	1	1
1991	0	0	1	0	0	1
1992	1	0	0	1	0	0
1993	1	1	0	0	1	0
1994	0	1	0	1	1	1
1995	0	0	0	0	1	1
1996	1	1	0	1	0	1
1997	0	1	1	0	1	1
1998	1	0	1	1	1	1
1999	1	1	0	1	1	1
2000	1	1	1	0	1	1
2001	1	0	1	1	1	1
2002	1	1	0	1	1	1
2003	1	1	1	0	1	1
2004	1	1	1	1	1	1

4. Runs test

A ‘run’ is defined as any sequence of 1 or more like elements from two classes. In the present context, this could mean above or below a level line, or a notional trend line. The runs test examines the notion of randomness in a series by looking at the number of runs of observed values above and below the median and comparing with the expected number which, along with variance, can be computed from theory. Non-randomness is usually represented by positive serial correlation of the observations, i.e. fewer than the expected number of runs, hence the test is usually one-sided. Serial correlation may be of interest in itself, e.g. as an interfering factor in a model-based test of trend (Loftis *et al.* 1991b), but could also arise from non-monotonic trends in the underlying signal.

Sources on the runs test are texts by Brownlee (1965, section 6.3) and Conover (1971, p. 349). Let the two classes of elements be a or b for ‘above’ or ‘below’ the sample median. Values equal to the median are ignored. Let the number of a’s be m . Then the number of b’s turns out also to be m, assuming no tied values. The expected number, u, of runs is

$$E(u) = 1 + m$$

and the variance is

$$V(u)=\frac{m(m-1)}{2m-1}.$$

Exact probabilities of runs are available (Swed and Eisenhart 1943) but for series of reasonable length (?), it is easier, and justifiable under the Central Limit Theorem, to assume that the statistic

$$\frac{u+0.5-(1+m)}{\sqrt{m(m-1)/(2m-1)}}$$

is approximately a standard normal variate. The 0.5 in the numerator is a correction for continuity. For a one-sided test, compare the test statistic with the standard normal deviate having cumulative probability of α. If the probability of the observed number of runs is less, serial correlation is detected.

The runs test is illustrated using the abundance indices for 1 to 2 year-olds from 1976 to 2004, and for 3 to 6 year-olds from 1983 to 2004 in table 5.13.4. Scoring of the runs above and below the median is shown in table 5.13.4. It is similar to the markings in table 5.13.2 except that values equal to the median must also be marked and ignored when counting the runs. Some tied values prevented n and m from being equal in each of age classes 2 and 5; m and n were adjusted to the minimum of the pair. The probabilities that the series were random, shown at the bottom of table 5.13.4, indicates that only the 6 year-olds were non-random by this test. The runs test is noticeably less sensitive to pure trend than the binomial test in relation to the median (section 2 above) because the continuity of runs above and below the median is frequently broken by variant observations.

Table 5.13.4. Runs test applied to age classes 1 to 6 from 1983 to 2004 for the abundance indices shown in table 5.13.1 illustrating how runs are obtained. 0 means index ≤ median and 1 means index > median; M means = median (observation ignored).

Year	**Age1**	**Age2**	**Age3**	**Age4**	**Age5**	**Age6**
1976	M	1				
1977	1	0				
1978	1	1				
1979	1	1				
1980	1	1				
1981	0	1				
1982	1	0				
1983	0	1	0	1	1	1
1984	1	M	1	0	1	1
1985	0	1	1	1	M	1
1986	1	0	1	1	1	1
1987	1	1	0	1	1	1
1988	0	0	1	0	1	1
1989	1	0	1	1	0	1
1990 = Median (1976-2004)	0	1	0	1	1	1
1991	0	0	1	0	0	1
1992	1	0	0	1	0	M
1993 = Median (1983-2004)	1	1	0	0	1	0
1994	1	0	1	0	M	M
1995	1	1	1	1	0	0
1996	0	M	1	0	1	0
1997	1	0	0	1	0	0
1998	0	1	0	0	M	0
1999	0	0	1	0	M	M
2000	0	0	0	1	0	M
2001	0	1	0	0	0	0
2002	0	0	1	0	0	0
2003	0	0	0	1	0	0
2004	0	0	0	0	0	0
N runs=>	14	18	13	16	8	2
N>median: *m*	14	13	11	11	8	9
N<=median: *n*	14	14	11	11	10	9
Adjusted *m*,*n*	14	13	11	11	8	9
E(runs)=	15	14	12	12	9	10
V(runs)=	6.74	6.24	5.24	5.24	3.73	4.24
Normal d.f.	-0.19	1.80	0.66	1.97	-0.26	-3.64
Probability that series is random	0.42	0.96	0.74	0.98	0.40	<0.01

5. Mann-Kendall's K

Kendall's tau is used to measure concordance or correlation (Hollander and Wolfe 1973). Slightly adapted (Mann 1945), Kendall's method is considered to be very sensitive to monotonic trends (Esterby 1993). Every observed value is paired with every value observed after it and the pair scored 1 if the first is greater than the second, 0 if the same, and –1 if the first is less than the second. The Mann-Kendall test statistic, *K*, is the sum of these values[2]. The null hypothesis is *H* : "trend ≥ 0" against *A*: "trend is monotonic, negative", or vice versa. Note that, if a turning point is definitely expected at a certain time, the observations could be re-ordered in time to conform with monotonicity under *A*.

K takes on large positive or negative values when a monotonic trend is present. One-tail probabilities of observing *K* under *H* are tabulated by Hollander and Wolfe (1973, Appendix A.21) but, for large samples of size *n*,

$$K^* = K\Big/\sqrt{\left[n(n-1)(2n+5)/18\right]}$$

is distributed as a standard normal variate if no data are tied. If there are ties, the square root denominator representing the standard error of *K* has to be inflated (Hollander and Wolfe 1973, p. 187). According to a citation in Yu et al. (1993), 10 observations are adequate for "large sample".

Scoring of observations for Kendall's *K* is illustrated in table 5.13.5 for age 1 abundance indices for cod from 1976 to 2004. The full table has 29 columns, so only 3 years of scores are shown. The sum of all the scores from 1976 to 2004 (*K*) was -106, and *n*=29, giving a large-sample standard normal approximation of -1.988. The corresponding probability of no trend is 0.02, implying here that a monotonic, downward trend was present in the signal. Note that the test is more sensitive to trend than the median test for trend where the probability of no trend was found to be 0.068 (table 5.13.2). On the other hand, the Mann-Kendall test is much more work to carry out on a spreadsheet. Kendall's test is available in R [cor.test(. . . method="kendall". . .). The observed values are correlated with times of observation, or their ranks, to achieve the Mann-Kendall test.

---

[2] Kendall's $\tau = 2K/n(n-1)$ is that used in correlation studies.

Table 5.13.5. Kendall's Tau test applied to age class 1 from 1976 to 2004 for the abundance indices shown in table 5.13.1 illustrating how the sequence of observed indices is scored for 1976 with subsequent years (column 1), and similarly for 1977 and 1978 (columns 2 and 3). Other years to 2004 not shown. *x* denotes observed value, *k* and *i* are years.

Year	**sign(x(k) - x(i)), i<k**		
	1976	1977	1978
1977	1		
1978	1	-1	
1979	1	-1	-1
1980	1	-1	1
1981	-1	-1	-1
1982	1	-1	-1
1983	-1	-1	-1
1984	1	-1	1
1985	-1	-1	-1
1986	1	-1	1
1987	1	-1	-1
1988	-1	-1	-1
1989	1	-1	1
1990	-1	-1	-1
1991	-1	-1	-1
1992	1	-1	1
1993	1	-1	-1
1994	1	-1	1
1995	1	-1	-1
1996	-1	-1	-1
1997	1	1	1
1998	-1	-1	-1
1999	-1	-1	-1
2000	-1	-1	-1
2001	-1	-1	-1
2002	-1	-1	-1
2003	-1	-1	-1
2004	-1	-1	-1

6. Thiel's or Sen's slope estimator

Thiel's slope estimator is used for a notional linear trend:

$$Y_i = \alpha + \beta x_i + e_i, \qquad\qquad i = 1,\ldots,n$$

The *e*'s must be mutually independent and from the same continuous population (Hollander and Wolfe 1973). Theil's estimator for β is similar in construction to that for Mann-Kendall's *K*. Every observed value is paired with every value observed after it, and the slope, $S_{ij} = (Y_j - Y_i)/(x_j - x_i)$, $i < j$, calculated. Theil's estimator is the median of these values. Sen's estimator, as described by Yu et al. (1993), appears to be exactly the same. Small sample confidence limits are available using Hollander and Wolfe (1973, chapter 9 and table A.21). For large samples, use the rounded integer value of

$$C_\alpha \approx z_{\alpha/2}\left\{\frac{n(n-1)(2n+5)}{18}\right\}^{0.5}.$$

The $1-\alpha$ confidence interval is obtained from the ranked slope values. Use $\{S_{(L)}, S_{(U)}\}$ where rank $L = (N - C_\alpha)/2$ and rank $U = (N + C_\alpha)/2$.

As an illustration, Thiel's slope estimator and 95% confidence limits were calculated for the age 1 abundance indices for cod from 1976 to 2004. The median slope, -0.23, is drawn through the intersection of the median value of Age 1 indices, 7.9 fish per hour, and the median observation time, 1990, in fig. 5.13.1 below. I am not aware of a method for estimating confidence limits for *Y* that takes into account the covariance of estimated α and β.

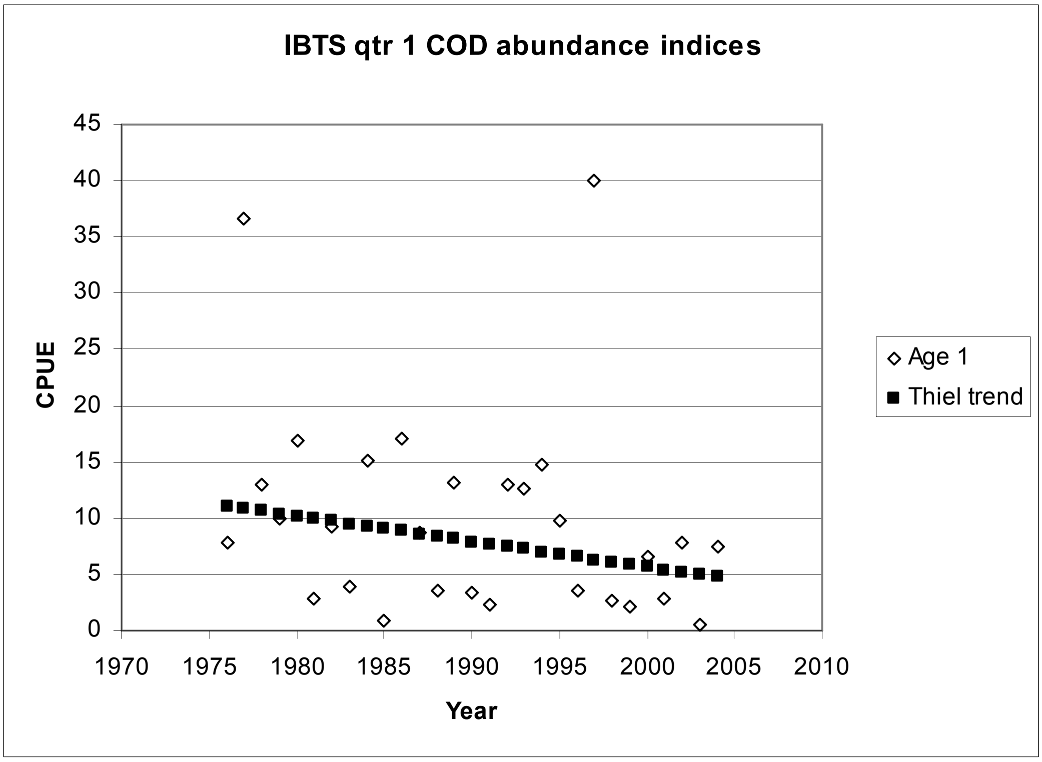


Figure 5.13.1 Thiel's slope estimator for the age 1 abundance indices, 1976 to 2004, shown in table 5.13.1.

7. Spearman's rho

Spearman's rho is the product-moment correlation between the ranks of paired data, the ranking being carried out separately for each variable of the pair. To test for trend, one member of the pair is the time of observation, the other is the observed variable. In practice, the arithmetic needed to calculate rho can be avoided by simply using

$$T = \sum_{i=1}^{n} \left(R(X_i) - i\right)^2$$

where i indexes the observation times and $R(X_i)$ is the rank of the corresponding observation (Lettenmaier 1976). This is also known as the Hotelling-Pabst test. T is small when $R(X_i)$ and i are positively correlated, and large when negatively correlated. Connover (1971, p389) gives quantiles for T for series up to 30 observations. Alternatively, use

$$w_p \approx \frac{1}{6}\left(n\left(n^2-1\right) + \frac{x_p\left(n\left(n^2-1\right)\right)}{6\sqrt{n-1}} \right)$$

where x_p is the p'th quantile of a standard normal deviate. T should be less than w_p (and x_p on the negative side of the normal distribution) for an upward trend, i.e. positive correlation with time, and T should be greater than w_p (and x_p on the positive side of the normal distribution) for a downward trend, i.e. negative correlation (Conover 1971).

Calculation of Spearman's Rho is illustrated in table 5.13.6 for age 1 abundance indices for cod from 1976 to 2004. $T = 5586$ which is greater than $w_p = 5564$ with $x_{0.975} = 1.96$, implying that the downward trend for 1-year olds is significant at $\alpha = 0.025$. This is not as significant as was found with Mann-Kendall's K which gave $\alpha = 0.018$ but it is more significant than was found with the median test, $\alpha = 0.07$. The sequence of probabilities is roughly matched inversely by the work required to carry out the tests on a spreadsheet. Connover states that the normal approximation is better for Kendall's tau than for Spearman's rho with small sample sizes.

Table 5.13.6. Spearman's Rho test applied to age class 1 from 1976 to 2004 for the abundance indices shown in table 5.13.1 illustrating the ranking (*R*) of years (*y*) and observed indices (*a*), and the computation of *T*.

Year	**Age 1 index**	**Year rank**	**Data rank**	**(R(y)-R(a))^2**
1976	7.9	1	15	196
1977	36.7	2	28	676
1978	12.9	3	21	324
1979	9.9	4	19	225
1980	16.9	5	26	441
1981	2.9	6	7	1
1982	9.2	7	17	100
1983	3.9	8	11	9
1984	15.2	9	25	256
1985	0.9	10	2	64
1986	17	11	27	256
1987	8.8	12	16	16
1988	3.6	13	10	9
1989	13.1	14	23	81
1990	3.4	15	8	49
1991	2.4	16	4	144
1992	13	17	22	25
1993	12.7	18	20	4
1994	14.8	19	24	25
1995	9.7	20	18	4
1996	3.5	21	9	144
1997	40	22	29	49
1998	2.7	23	5	324
1999	2.1	24	3	441
2000	6.6	25	12	169
2001	2.8	26	6	400
2002	7.8	27	14	169
2003	0.6	28	1	729
2004	7.537	29	13	256
			Sum, *T*	5586

8. Jonckheere's test

Jonckheere's test is a nonparametric version of a one-way analysis of variance with unequal sample sizes, except that it tests *H*: 'no treatment effect' versus the special alternative, *A*: 'the treatments are ordered in effect' (Hollander and Wolfe 1973, p. 120). This can be applied to trends by equating observation times to treatments and then arranging them in the order implied by the possible trend so as to make a one-sided test. Hypothesis *A* is then equivalent to a monotonic trend, as for Mann-Kendall's *K*. For a time-series of un-replicated observations, Mann-Kendall's *K* is the conventional test to use. However, if observations are independently replicated at some time points, the Mann-Kendall test is not clearly applicable and Jonckheere's test can be applied instead.

Let the number of observation times be *k*. It is necessary first to find $k(k-1)/2$ Mann-Whitney counts U_{uv} where *u* and *v* are times of observation and $1 \le u < v \le k$:

$$U_{uv} = \sum_{i=1}^{n_u} \sum_{j=1}^{n_v} \phi\left(X_{iu}, X_{jv}\right).$$

n_u, n_v are the numbers of observations at times *u* and *v* respectively, while $\phi(a,b) = 1$ if $a < b$, and 0 otherwise. Be warned, this is an extensive job for large *k* using a spreadsheet. The test statistic, $J = \sum_{u<v}^{k} U_{uv}$, can be compared with table A.8 in Hollander and Wolfe (1973), or, when the minimum number of replicate observations at any point is large (?), can be transformed to J* which is approximately normally distributed:

$$J^* = \frac{J - \left\{\left(N^2 - \sum_{j=1}^{k} n_j^2\right)/4\right\}}{\sqrt{\left(N^2(2N+3) - \sum_{j=1}^{k} n_j^2(2n_j+3)\right)/72}}.$$

$N = \sum_{j=1}^{k} n_j$. As for Kendall's test, this would be one-tailed. *H* is rejected if J* exceeds the standard normal deviate at the required level of significance.

Table 5.13.7 presents a fabricated set of data to illustrate computations for Jonckheere's test. Abundance indices for 1-year olds were arbitrarily re-assigned to a shortened series of 10 years so as to create a set of replicate observations. *J* was found to be 209 which together with the totals in the right 3 columns of table 5.13.7a gave *J** = 0.8318 which is less than 1.645, the standard normal deviate corresponding to 95% of the area under the normal curve. Not surprisingly, the arbitrarily re-arranged data did not show a significant monotonic trend.

Table 5.13.7. Jonckheere's test applied to abundance indices for cod of age class 1 from 1976 to 2004 taken from table 5.13.1 and arbitrarily assigned to a shortened series of 10 years with variable replication annually for illustrative purposes. a) resulting test data and computation of terms for the test statistic; b) illustrating scoring for u = 1976 and v = 1977, 1978.

a)

Year	**Dummy replicate observations**					**nj**	**nj^2**	**nj^2*(2nj +3)**
1976	7.9	17	3.5			3	9	81
1977	36.7	8.8				2	4	28
1978	12.9	3.6	2.7			3	9	81
1979	9.9	13.1	2.1	40	7.8	5	25	325
1980	16.9	3.4	6.6			3	9	81
1981	2.9	2.4				2	4	28
1982	9.2	13				2	4	28
1983	3.9	12.7	0.6	2.8		4	16	176
1984	15.2	14.8	7.537			3	9	81
1985	0.9	9.7				2	4	28
					Sum=>	29	93	937

b)

***u* = 1976**	Replicate 1	Replicate 2	Replicate 3
***v* = 1977**			
Replicate 1	1	1	1
Replicate 2	1	0	1
***v* = 1978**			
Replicate 1	1	0	1
Replicate 2	0	0	1
Replicate 3	0	0	0

9. Permutation (randomisation) and bootstrapping methods

Permutation tests are described in connection with benthic studies by Bell et al. (1981). They are alleged to be more sensitive to trend than rank tests. Let Y be the observed value, and t the time of observation, $t = 1,...,T$. Calculate

$$h = \sum_t tY_t$$

and compare with the $T!$ values of h^* computed with the t values permuted. Bell et al. state that if there is no change in Y over time, h is likely to fall near the middle of the range of h^*, otherwise near one of the extremes. Probabilities of h or values more extreme can be found because each permutation is equally likely. Computation of all permutations may be an onerous and unnecessary task. The `sample()` function in the R programming language readily produces random permutations that may suffice for building up a reference set for assessing probabilities. The case for a randomisation test is likely to be strong when the number of observations available is very small, e.g. less than 10. Many other nonparametric methods then need access to special tables to estimate significance because the Central Limit Theorem can not be invoked satisfactorily to argue that test statistics are approximately normally distributed, as for Spearman's rho and Mann-Kendall statistics for example. Ties can further complicate application of competing nonparametric methods.

Edgington (1995, p. 15) points out that "Parametric tests of all kinds, including relatively complex tests, . . . become distribution-free when the significance is determined by a randomization test procedure." Consider for example a least squares estimator of slope. Rather than assume that the observations at each time point were a random sample from some population and looking up the value of *t* or *F* for the slope based on that assumption, the observations and times are permuted to create a reference set of estimated *t* or *F* statistics, then the observed value is compared with the reference set and its statistical significance judged from the proportion of the reference set having a more extreme value. Edgington describes several randomisation tests applicable to trends over treatment levels in randomised experiments but states (p. 217) ". . .randomization trend tests do *not* test hypotheses about trends; they simply utilize test statistics sensitive to trend which test the null hypothesis of no differential treatment effect." His view is possibly related to the focus on experiments in his book. On the other hand, finding that an observed trend over time is unlikely by a randomisation test, even outside a randomised experiment, seems to be no less useful than finding it by means of other nonparametric tests.

Bootstrapping observations could be another way of assessing the significance of observed trends, particularly for fisheries survey data for which abundance indices can be bootstrapped to estimate sampling and measurement errors (Beare *et al.* 2002) even though analytical formulae for variances are hard to derive or non-existent. The trend would be fitted to each bootstrapped series and the distribution determined. To be nonparametric, the estimates should be based on sampling theory and no model assumed. The bootstrap is no less vulnerable than most other statistical methods to small numbers of observation times.

There are subtle differences between bootstrapping and randomisation tests. Bootstrapping uses a sample as a surrogate for the population and (re-)samples the sample *with* replacement. A randomisation test of trend permutes observations to time points *without* replacement and without reference to the sampling process generating the observations. Bootstrapping empirically estimates the distribution of statistics assuming that the observed sample looks like the true population. On the other hand, randomisation tests provide a form of statistical inference when the sample itself is assumed to be the total population of interest. When testing for trend in an environmental context, the quality of the sample appears to be just as important as for the bootstrap.

10. Dietz and Killeen test for multivariate monotone trend (turquoise box)

Dietz and Killeen (1981) derived a formula for the covariance matrix of Mann-Kendall statistics estimated from a multivariate monitoring programme, and proposed a test statistic based on it that is asymptotically distributed as χ^2. Use of multivariate methods "controls the overall significance level" when multiple univariate tests are made with covarying observations, and there were very few cases in a study of lake water quality "where the univariate methods perform better" (Loftis *et al.* 1991b). Multivariate methods appear to be especially applicable when monitoring groups of indicators expected to respond similarly to environmental changes of concern. The Dietz and Killeen test has the same strengths as the Mann-Kendall test for discovering trends that are specifically monotonic. An alternative but related test when the aggregate trend of several variables is of interest, e.g. when correlated measures of one variable are made seasonally, was given by Hirsch and Slack (Hirsch and Slack 1984). Comparisons of nonparametric methods with linear models in a multivariate context were made using simulation by Loftis et al. (Loftis *et al.* 1991a).

Code in R (Mvar.trend.r) for the Dietz and Killeen test is available from http://www.ifremer.fr/drvecohal/fisboat/index.htm; it was trialled with data given by Dietz and Killeen and found to give the same results. Fig. 5.13.2 shows selected output from Mvar.trend.r applied to the cod abundance indices, ages 1 to 6, 1983 to 2004. (About 2 minutes were required for running; times increase roughly in proportion to the product of $n^3 p$.) The first matrix shows the Mann-Kendall statistics. Those along the diagonal are computed as for the univariate statistics; see section 5 above. Those off the diagonal are

$$K_{XY} = \sum_{i<j} sign\left[\left(X_j - X_i\right)\left(Y_j - Y_i\right)\right]$$

where $sign(x) = 1, 0, -1$ for positive, zero, and negative values respectively. K_{XY} is high and positive when both the X and Y variables are showing monotonic trends in the same direction, and high and negative when in opposite directions. The second matrix, **S**, in box 10 shows the covariances of these statistics calculated using formulae given in the appendix of Dietz and Killeen. The third matrix shows the Spearman rank correlations of the observed values. The highest correlations occur among the older age groups. Note that although values along the diagonal of a correlation matrix are normally 1, some here are less than 1 due to tied values within the time-series. The test statistic, $\mathbf{K}^{\mathbf{T}}\mathbf{S}^{-1}\mathbf{K} = 25.09$ is compared with χ^2 with 6 d.f. (all age classes contribute to d.f. since no parameters are estimated) and found to be highly significant, $P < 0.005$; $\chi^2 = 18.548$. The standardised K for each age class and their standard errors are shown below. This is the same information that would come from testing each age class separately. Examination of them when the multivariate null hypothesis is rejected should indicate which of the age classes was responsible (Loftis *et al.* 1991b). For 1-sided tests, standardised K greater in magnitude than 1.64 are likely to be contributing to the significance of the multivariate result. The results suggest that all except 1-year olds are contributing.

Figure 5.13.2. Output from Mvar.trend.r applied to cod abundance indices for ages 1 to 6 from 1983 to 2004 given in table 5.13.1. $Kxy is Matrix of Kendall stats (K), age1 to 6; $S is covariance matrix of K; $Spearman.corr is the rank correlation matrix of the abundance indices; $Standardised.K is K from diagonal of S (above), divided by $St.error.K (last row). Some rows of output are omitted to save space.

```
>Mvar.trend(IBTS.indices)
$Kxy
         [,1]     [,2]     [,3]     [,4]     [,5]     [,6]     [,7]
[1,]            0        0        0        0        0        0        0
[2,]            0      -45       -2       12       32       11       22
[3,]            0       -2      -72      -31       27       34       27
[4,]            0       12      -31      -60      -13       52       66
[5,]            0       32       27      -13      -74       16       91
[6,]            0       11       34       52       16      -97       95
[7,]            0       22       27       66       91       95     -166

$S
         [,1]     [,2]     [,3]     [,4]     [,5]     [,6]     [,7]
[1,]            0        0        0        0        0        0        0
[2,]            0 1257.667 29.33333 108.6667 249.3333 61.66667      140
[3,]            0 29.33333 1256.667     -155 206.3333      309 232.3333
[4,]            0 108.6667     -155     1254 -90.3333      394 475.3333
[5,]            0 249.3333 206.3333 -90.3333     1248 106.6667      662
[6,]            0 61.66667      309      394 106.6667     1215 724.6667
[7,]            0      140 232.3333 475.3333      662 724.6667 1232.667

$Spearman.corr
         [,1]     [,2]     [,3]     [,4]     [,5]     [,6]     [,7]
[1,]            0        0        0        0        0        0        0
[2,]            0        1 0.025409  0.08865 0.202146 0.049125 0.112366
[3,]            0 0.025409 0.999435 -0.12253 0.167137 0.252117 0.189159
[4,]            0  0.08865 -0.12253 0.997741 -0.07284 0.319029 0.383964
[5,]            0 0.202146 0.167137 -0.07284 0.994353 0.085827 0.535009
[6,]            0 0.049125 0.252117 0.319029 0.085827 0.971203 0.586957
[7,]            0 0.112366 0.189159 0.383964 0.535008 0.586957  0.98419

$Standardised.K
[1]       -1.26891 -2.03106 -1.69435 -2.09471 -2.78281 -4.72809

$St.error.K
[1]        35.4636 35.44949 35.41186 35.32704 34.85685 35.10935
```

References

Beare, D., Castro, J., Cotter, J., van Keeken, O., Kell, L., Laurec, A., *et al.* (2002) Evaluation of research surveys in relation to management advice (EVARES). Final report. *DGXIV Fisheries, European Commission, Brussels.* **FISH/2001/02 - Lot 1**. Available from a.j.cotter@cefas.co.uk

Bell, C.B., Conquest, L.L., Pyke, R. and Smith, E.P. (1981) Some nonparametric statistics for monitoring water quality using benthic species counts. *Environmetrics, selected papers, SIAM, Philadelphia* **8**, 100-120.

Brownlee, K.A. (1965) *Statistical theory and methodology in science and engineering.* 2nd. edn., John Wiley & Sons, New York

Cochran, W.G. (1950) The comparison of percentages in matched samples. *Biometrika* **37**, 256-266.

Conover, W.J. (1971) *Practical nonparametric statistics.*, John Wiley & Sons, New York

Cotter, A.J.R. (1985) Water quality surveys: a statistical method based on determinism, quantiles and the binomial distribution. *Water Research* **19**, 1179-1189.

Cotter, A.J.R. (1994) Compliance testing of two or more water quality determinands using quantiles. *Environmetrics* **5**, 29-45.

Dietz, E.J. and Killeen, T.J. (1981) A nonparametric multivariate test for monotone trend with pharmaceutical applications. *Journal of the American Statistical Association* **76**, 169-174.

Edgington, E.S. (1995) *Randomization tests.* 3rd. edn., Marcel Dekker, Inc., New York

El-Shaarawi, A.H. (1993) Environmental monitoring, assessment and prediction of change. *Environmetrics* **4**, 381-398.

El-Shaarawi, A.H. and Niculescu, S.P. (1992) On Kendall's Tau as a test of trend in time series data. *Environmetrics* **3**, 385-411.

Esterby, S.R. (1993) Trend analysis methods for environmental data. *Environmetrics* **4**, 459-481.

Hirsch, R.M. and Slack, J.R. (1984) A nonparametric test for seasonal data with serial dependence. *Water Resources Research* **20**, 727-732.

Hirsch, R.M., Slack, J.R. and Smith, R.A. (1982) Nonparametric tests for trend in water quality. *Water Resources Research* **18**, 107-121.

Hollander, M. and Wolfe, D.A. (1973) *Nonparametric statistical methods.*, John Wiley & Sons, New York

Kendall, M.A. (1976) *Time-series*, Charles Griffin and Co. Ltd., London

Lettenmaier, D.P. (1976) Detection of trends in water quality data from records with dependent observations. *Water Resources Research* **12**, 1037-1046.

Loftis, J.C., Taylor, C.H. and Chapman, P.L. (1991a) Multivariate tests for trend in water quality. *Water Resources Research* **27**, 1419-1429.

Loftis, J.C., Taylor, C.H., Newell, A.D. and Chapman, P.L. (1991b) Multivariate trend testing of lake water quality. *Water Resources Bulletin, American Water Resources Association* **27**, 461-473.

Mann, H.B. (1945) Nonparametric tests against trend. *Econometrica* **13**, 245-259.

Nicholson, M.D. and Fryer, R.J. (1992) The statistical power of monitoring programmes. *Marine Pollution Bulletin* **24**, 146-149.

Pettitt, A.N. (1979) A non-parametric approach to the change-point problem. *Applied Statistics* **28**, 126-135.

Prins, J. (2006). "NIST/SEMATECH e-Handbook of Statistical Methods." from http://www.itl.nist.gov/div898/handbook/prc/section4/prc473.htm.

Swed, F.S. and Eisenhart, C. (1943) Tables for testing randomness of grouping in a sequence of alternatives. *Annals of mathematical statistics* **14**, 66-87.

Van Bell, G. and Hughes, J.P. (1984) Nonparametric tests for trend in water quality. *Water Resources Research* **20**, 127-136.

Yu, Y.-S., Zou, S. and Whittemore, D. (1993) Non-parametric trend analysis of water quality data of rivers in Kansas. *Journal of Hydrology* **150**, 61-80.

Zetterqvist, L. (1991) Statistical estimation and interpretation of trends in water quality time series. *Water Resources Research* **27**, 1637-1648.

5.4 Construction of multivariate indicators.

5.4.1 Principal Components Analysis (PCA) and biological indicators

Pierre Petitgas,
IFREMER, Nantes,
France

Introduction

For each survey, a variety of indices of stock attributes are usually estimated. A multivariate monitoring procedure is then potentially more efficient than procedures based on the analysis of a collection of univariate monitoring charts as the multivariate analysis will make coherent use of the relationships between the many indices of stock attributes. Here we suggest application of PCA to the Fisboat biological indicators (abundance indices, length indices and mortality index) and represent the evolution of the stock by a multivariate distance to a reference gravity centre. Clearly, abundance and length indices are potentially related and more consistency can be obtained by explicitly using these correlations. But the method will not take into account any correlations in time between indicators, in particular lagged effects of one indicator on another.

Method

Consider an array where for each line (observation) we have a vector of measurements for a variety of parameters (variables). Measured values of variables are in columns and each line is one observation (here in time). The variables can be correlated between each others. PCA constructs linear combinations of variables (factors) that are non correlated between each other and that best account for the variability in the data array. Mathematically, this is done by diagonalising the correlation matrix of the variables. Eigen vectors geometrically support principal components. These are ranked by their decreasing importance of data variance explained. The geometrical properties of the method enables representation of the correlation structure among the variables as well as the position of each observation in the space of the principal components (factorial space). The correlation between two variables (variable - variable or variable - principal component) is represented by the angle between vectors. The similarity between observations is represented by their Euclidean distance in the factorial space. It is usual to analyse correlation between variables and similarity between observations in a factorial sub-space made by a reduced number of principal components. These are the first principal components that account for a large percentage of the data variance (e.g., 80%). Such reduction corresponds to filtering variability in the data assumed to be noise. The few retained non correlated factors then summarise the multivariate structure of the data. PCA (e.g., Lebart et al., 1995) is a widely used technique in many fields including marine ecology that was first developed more than fifty years ago.

Software

The R code available from the FISBOAT web site, http://www.ifremer.fr/drvecohal/fisboat/index.htm, is pcachart.R. It is commented. It uses the R library ade4 (Chessel et al., 2006) for performing the PCA. The PCA is applied to the Fisboat Table 2 of biological (non spatial) indicators. A set of reference years is defined, in which the population is considered in acceptable health status. The PCA is performed giving high weight to the set of reference year observations (99.9%). In that way, the factorial space is not affected by years outside the reference period: years outside the reference period play the role of passive variables that are projected in a reference factorial space. First the decrease in the eigen values is illustrated. Then the correlation structure in the indicators is illustrated using the first three principal axes and the loadings of the indicators on these are provided in a table. This allows to interpret the principal axes. Then the similarity between years is illustrated by positioning the years in the plane made by the first two principal axes. On the figure, the reference period years are marked by a symbol while the non reference years to be monitored are labelled by their number. A multivariate distance in each year is then calculated, which quantifies the deviation of that year to the gravity centre of the reference years, and is saved with the name mdbio. The multivariate distance is the euclidean distance between the position of any year *p(y)* and the gravity centre of the reference years c_{ref}: $d_y = \sqrt{d^2(p_y; c_{ref})}$.

Inputs are : Table 2 (from the FISBOAT web-site) of biological non spatial indices, years to consider as reference period, number of principal PCA axes to consider for computing the multivariate distance mdbio. Outputs are: a figure representing the decrease in eigen values, figures illustrating the correlation structure between variables, correlation table of the variables on the PCA axes (loadings), the time series of the mdbio multivariate distance as well as its histogram.

Example

North Sea cod

PCA can be used to set up a multivariate monitoring approach of stock status using the many indicators of biological stock attributes available for North Sea cod, see tables 1 and 2 at http://www.ifremer.fr/drvecohal/fisboat/. Because the indicators are correlated to each other, PCA is useful to summarise the correlation structure between the parameters and reduce the dimensionality of the monitoring scheme using a small number of non correlated factors. The monitoring approach will then take place in the factorial space composed of the first two (or more) principal axes. A reference domain in that factorial space (in-control domain) can be defined based on the position of reference years in that factorial space. The definition of reference years is analogous to monitoring a process when it is in-control: the reference year period is the set of years where the stock could be considered in acceptable health. The multivariate monitoring approach compares the current year vector of stock indicators to that of the reference period. It is therefore suggested to estimate the gravity centre of the reference years and, for each year, to calculate the distance to that reference gravity centre. The time series of the multivariate distance then summarises the deviation of the population biological characteristics from its reference status.

Fisboat table 2 of biological non spatial indicators comprises the year observations as lines and the columns as variables. It is a typical set for input to a PCA. The PCA will display the correlation structure between indicators (abundance, length, mortality) and will allow quantification of which years depart from the others, not just because of one indicator but as a whole in their multivariate characteristics.

Reference years are 1985-1994. Figure 5.10.1 illustrates the correlation structure in the biological indicators. Length50 at maturity heavily determines the first principal axis. The second is determined by total abundance and the opposition between total abundance and length at the third quartile (L75). The third axis is determined by mortality (Z) which is hardly correlated with any other indicator. Overall, length indices and abundance indices are correlated. Figure 5.10.2 illustrates the multivariate monitoring approach. Years 1997, 1999, 2003-2005 are well outside the domain defined by the reference years, meaning that they depart largely from the reference status. The direction of departure is on the first and second axes meaning that departure is primarily guided by changes in abundance and length at maturity. This is quantified by the multivariate distance on which a statistical monitoring scheme can then be applied.

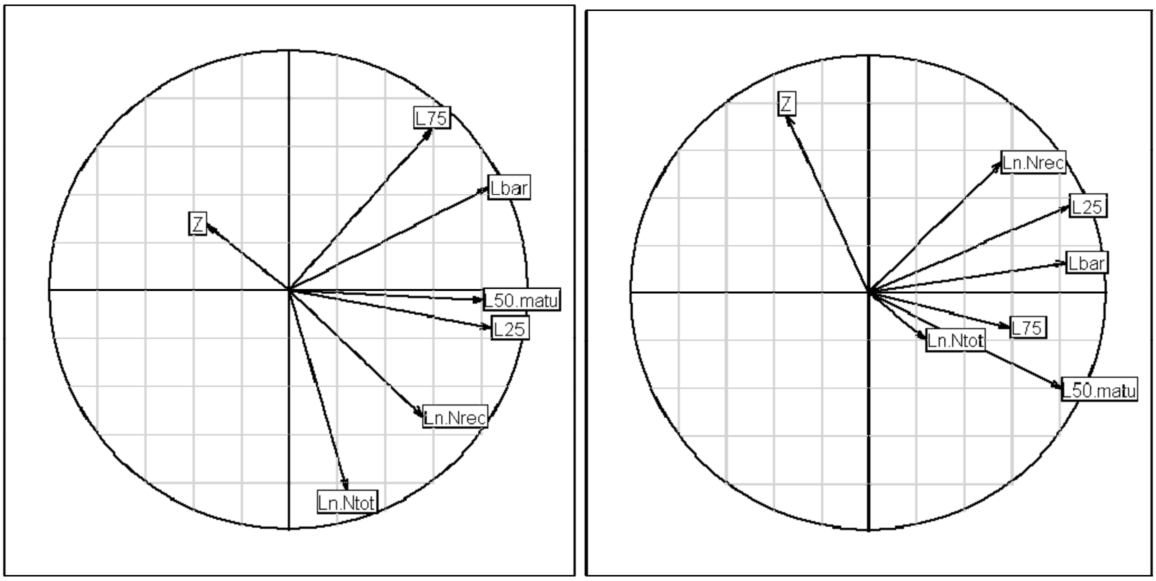


Figure 5.10.1. Correlation structure between the biological non spatial indices (Fisboat Table 2) for North Sea cod. Left: principal axes 1 and 2; right: principal axes 1 and 3.

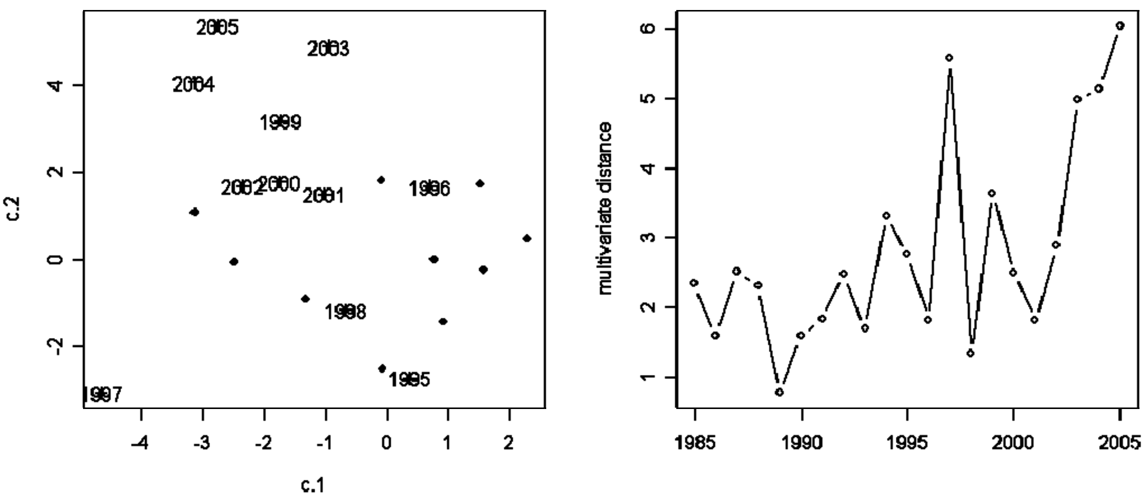


Figure 5.10.2. Monitoring North Sea cod in the factorial sub-space of the two first principal axes using the biological non spatial indicators (Fisboat Table 2). Left: representation of years in the factorial sub-space (the black diamonds are the reference years); right: the time series of the multivariate distance representing the deviation of the stock from its reference status.

References

Chessel, D., Dufour, A. and Dray, S. 2006. The ade4 package. R documentation. R-project.

Lebart, L., Morineau, A. and Piron, M. 1995. Statistique exploratoire et multidimensionelle. Dunod, Paris.

5.4.2 Multi-factorial analysis (MFA) and spatial indicators

Pierre Petitgas, Jean Charles Poulard.
IFREMER, Nantes, France.

Introduction

MFA is a multi-table analysis method (Escoffier and Pagès, 1994 ; Dazy and Le Barzic, 1996) extending PCA methodology (Principal Component Analysis) to the analysis of 3D structured data. In particular, MFA is designed for cases where the same variables (data matrix columns) are measured for the same individuals (e.g., stations : data matrix rows) at various times (third dimension of the data structure). The method allows the construction of a unique factorial space in which to represent each data matrix for each time, each variable and each individual. This unique factorial space is a compromise space that best matches that of each data matrix at each time. It allows extensive tables of data to be represented pictorially in such a way that groupings among variables in space and time are readily identifiable.

The method has been applied in fisheries science to characterise seasonal and inter-annual variation in fish community structure (Gaertner et al., 1998), fishing activity (Poulard and Léauté, 2002), as well as common structure between trophic levels (Petitgas et al., 2006). It has been used in FISBOAT for summarising the average life cycle spatial organisation (Woillez et al., in press). Here we suggest a measure of inter-annual variation in that pattern and an R code for doing so.

Method

The method proceeds in two steps. First a PCA is performed on each data matrix. Then each variable at each time is weighted by the inverse of the first eigen value of that matrix. Then a general matrix is constructed that contains all the weighted variables in columns and the individuals as rows. The PCA of that general matrix constructs the MFA compromise factorial space. Its principal axes are interpretable using the correlation of the variables to them. The interest in the method is the construction of a compromise factorial space in which to represent the 3D structure of the data : each individual is represented by n points (n repetitions in time) as well as by its gravity centre (average position in the compromise factorial space).

Software

The R code available from the FISBOAT web site, http://www.ifremer.fr/drvecohal/fisboat/index.htm, is dmul_mfa.R. It is commented. It uses the R library ade4 (Chessel et al., 2006) for performing the MFA. Using the Fisboat Table 1 of spatial indicators [http://www.ifremer.fr/drvecohal/fisboat/] as input, the code will produce a figure representing the life cycle spatial pattern in the two first principal axes of the MFA space. On the figure, each point represents a particular age in a particular year. The average position for each age is labelled. A multivariate distance is built quantifying in each year the deviation from the average. This distance, named dmul, is the sum over the ages of the distance between the yearly position of each age p(a,y) and its gravity centre c(a): $d_y = \sqrt{\sum_a d^2(p_{a,y};c_a)}$. The code also provides as output a table of the number of times the correlation between each spatial indicator and the MFA axes is greater than 0.5, allowing interpretation of the axes. The output result is the times series of the multivariate dmul distance, which characterises the inter-annual variation around the average spatial pattern of the life cyle.

The position of each age in each year p(a,y) is computed by applying MFA to all years. The age-specific gravity centres relative to which deviations are referenced are calculated considering reference years only, that must be provided as input. The gravity centre for each age is the average position of each age specific point in the MFA space for the years of the reference period.

Inputs are : Table 1 of spatial indices, years to consider as the reference period, and the number of principal MFA axes to consider for computing the distance dmul. Outputs are a figure representing the life cycle pattern in the MFA space, correlation tables of the variables with the MFA axes, the time series of the dmul distance as well as its histogram.

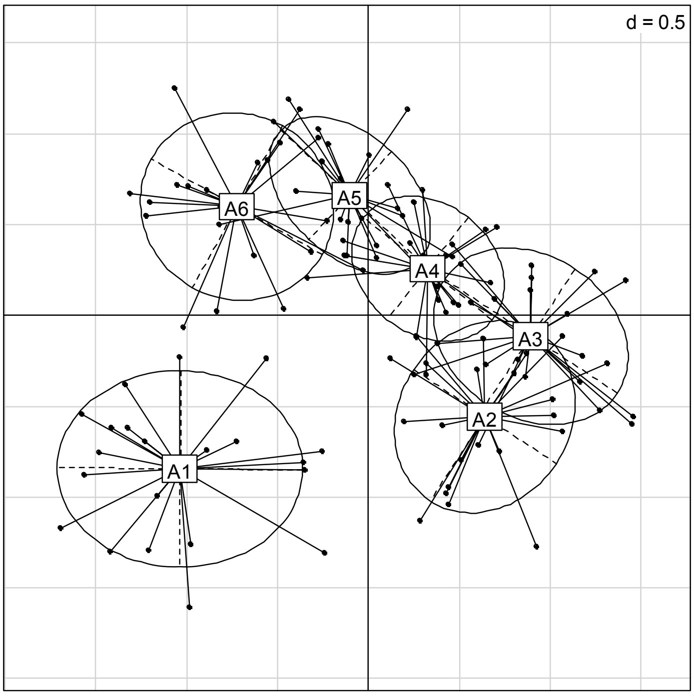
Example

North Sea cod

The life cycle of any fish population is organised in space because the fish will occupy different habitats at different life stages. Therefore, the young and the old can be expected to show different characteristics in their spatial distribution. The spatial indicators developed in Fisboat have been applied to survey numbers-at-age for a series of years. They thus characterise the fish distributions at age in different years. In each year, the matrix made of the spatial indicators in columns and the ages in rows describes the life cycle pattern of the population in that year. In the MFA compromise space, the cloud of points of the compromise individuals represents the average life cycle pattern. In general, structure in the data is strong and the first two principal axes are explained at least by location and aggregation indicators, meaning that there is a change with age in the location and aggregation of the fish. MFA is then used to describe these patterns and their inter-annual variation. In particular, the departure of individual points in each year from the average cloud represents inter-annual variation in the life cycle organisation. An R code has been developed to calculate such variation.

Reference years are 1985-1994. Fig. 5.11.1 illustrates the structure in the spatial organisation and fig. 5.11.2 illustrates the monitoring of the deviations from the mean reference structure. Each point on fig. 5.11.1 (left) represents the position of a given age and year in the MFA factorial sub-space of the first two principal components. The age labels identify the mean position for each age for the reference years, materialising the reference life cycle spatial organisation of North Sea cod. In effect, marked and progressive differences exist across ages. The table on fig. 5.11.1 (right) summarises the correlation structure in the spatial indicators. The first component is determined by having a larger area occupied, a centre of gravity more to the north and west, higher inertia (dispersion), and a lower nugget effect. The second component is determined by having a centre of gravity more to the north, higher anisotropy and a smaller positive area. Marked and progressive differences in the spatial distribution of ages are shown on fig. 5.11.1 (left), which characterise the life cycle spatial pattern of North Sea cod. Young (A1) and old ages (A5-6) differ on the first component from intermediate ages (A2-4). Spatial distributions of young and old ages are more to the east, are less dispersed, occupy a smaller area, and have a higher nugget effect than that of intermediate ages, which are more to the west, more dispersed, occupying larger areas and with smoother correlation. Age 1 and Ages 5-6 differ on the second principal component by the location of their centre of gravity and anisotropy. The spatial distribution of old ages is more to the north, is more anisotropic, and occupies a smaller area than that of the age 1 fish.

In order to monitor the departure of all the ages from the reference mean spatial organisation, the multivariate distance dmul was calculated (fig. 5.11.2): in each year the distance between the age reference point (labelled) and the current year's point (black dot) was calculated and summed over the ages. Ages 4 and 5 show some elongation in the direction of departure from their reference, meaning that these ages tend to show a systematic change, namely a reduction in area occupied and a more northerly centre of gravity. This is quantified by the multivariate distance on which a statistical monitoring scheme could be applied.



	PC1	PC2
PositiveArea	**12+|0-**	**0+|16-**
Inertia	**12+|2-**	1+|2-
Anisotropy	0+|2-	**15+|0-**
xcg	**0+|17-**	4+|4-
ycg	**13+|0-**	**15+|0-**
MicrostructureI	**1+|11-**	3+|3-
EquivalentArea	**12+|0-**	1+|2-
SpreadingArea	**18+|0-**	**0+|7-**

Figure 5.11.1. Left: MFA representation of each age in each year (points) relatively to the age gravity centre. Right: Spatial indicators that are the most correlated to the first two principal axes

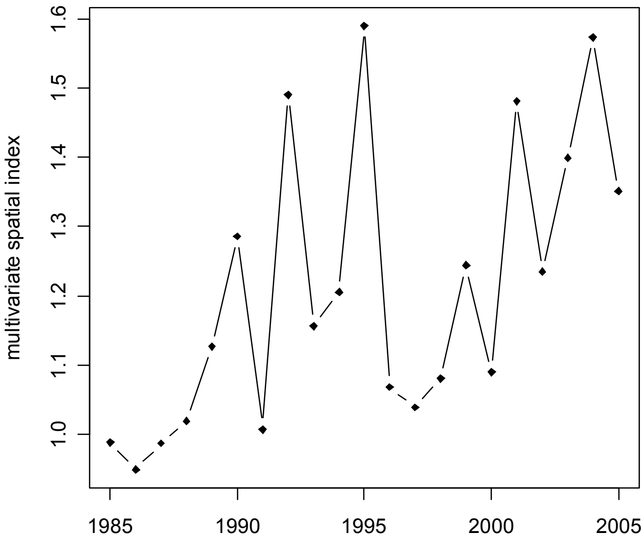


Figure 5.11.2. Monitoring North Sea cod in the sub-space of the two first principal MFA axes using the spatial indicators (Fisboat Table 1) : time series of the multivariate distance representing the deviation of the stock from its reference spatial distribution all ages considered.

References

Chessel, D., Dufour, A. and Dray, S. 2006. The ade4 package. R documentation. R-project.

Dazy, F. and Le Barzic, J.-F. 1996. L'analyse des données évolutives : méthodes et applications. Editions Technip, Paris, 227p.

Escoffier, B. and Pagès, J. 1994. Multiple factor analysis (AFMULT package). Computational Statistics and Data Analysis, 18 : 121-140.

Gaertner, J.-C., Chessel, D. and Bertrand, J. 1998. Stability of spatial structures of demersal assemblages. Aquatic Living Resources, 11 : 75-85.

Petitgas, P., Massé, J., Bourriau, P., Beillois, P. et al. 2006. Hydro-plankton characteristics and their relationship with sardine and anchovy distributions on the French shelf of the Bay of Biscay. Scientia Marina, 70(S1): 161-172.

Poulard, J.-C. and Léauté, J.-P. 2002. Interaction between marine populations and fishing activities: temporal patterns of landings of La Rochelle trawlers in the Bay of Biscay. Aquatic Living Resources, 15: 197-210.

Woillez, M., Poulard, J-C., Rivoirard, J., Petitgas, P. and Bez, N. (accepted February 2007). Indices for capturing spatial patterns and their evolution in time, with application to European hake (Merluccius merluccius) in the Bay of Biscay. ICES Journal of Marine Science.

5.4.3 Min/Max autocorrelation factors (MAFs) and time continuity

Woillez M., Rivoirard J.
Centre de Géostatistique, Armines/Ecole des Mines de Paris,
Fontainebleau, France.

What are MAFs?

Min/max autocorrelation factors (MAFs) are a multivariate statistical method, having similarities with the classical Principal Components of PCA when analyzing repeated values taken by a set of variables. When applied to a time series, the MAFs allow the set of initial variables to be decomposed into factors, the autocorrelation of which decreases from the first factors to the last ones (or more generally, the variogram – half variance of increments – of which increases from the first factors to the last ones). Hence the very first factors extract the part of the variables which is the most continuous in time. A recent application of MAFs in a fishery context was reported by Erzini et al. (2005).

Some details are now provided, beginning with PCs. The PCs are linear combinations of the original variables, each of them explaining a decreasing part of the variability present in the values (these values can be seen as a cloud of points in the space having the initial variables as coordinates; the cloud is centered on the means of the variables, and we are interested in the variability of the cloud around its center, which corresponds to the centered variables). The PCs are uncorrelated with each other. The 1st component explains the highest part of the variability (it corresponds to the direction of maximal variability of the cloud). The 2nd component explains the second highest part of the variability, while being uncorrelated to the 1st one (it corresponds to the direction of maximal variability of the cloud, while being orthogonal to the first direction), and so on. Then the set of variables can be decomposed and represented by these uncorrelated PCs, and they can be summarized by selecting the often few PCs that explain most of the variability of the cloud.

Note that the PCs depend on the magnitude of the values taken by the different variables, and then on the unit used for each of them. Because the different variables may be of different nature, with different and conventional units, PCA is very often performed on the normed variables (i.e. having the variance of each variable set to 1).

PCA is well suited to the case where the repeated values of the variables are independent. In case of repetition in time (time series) or in space, the PCs are uncorrelated with each other at the same time (or location), but may be correlated between different times (or locations), making the representation of the variables by the factors (the PCs) less appropriate.

The MAFs, which are also linear combinations of the original variables, and have a variance of 1, offer a better representation of variables distributed in time or space: in addition to being uncorrelated with each other at the same time (or location), they are uncorrelated with each other for a given time (or space) lag (taken equal to the sampling lag in practice). Moreover they are computed with the aim of: (1) presenting the highest autocorrelation (or smallest variogram) at this lag for the 1st MAF; (2) then presenting the second highest autocorrelation at this lag, while being uncorrelated with the 1st MAF, for the 2nd MAF. Etc. Hence, in a time series, the MAFs offer a way to build the combinations of variables which present the maximal continuity in time (as measured at the lag) for the first MAFs, and the minimal continuity for the last MAFs.

Note that the MAFs depend on the chosen computation lag and may be correlated with each other for other lags. Note also that the MAFs do not depend on a possible normalisation of the initial variables.

From a technical point of view, the MAFs are the solution of a generalized eigenvalues problem. This can be simplified into a simple eigenvalues system in the case where the variables are uncorrelated with each other at the same time (or location). Hence MAFs can be obtained by solving two simple eigenvalue systems: one to transform the initial variables into PCs, the other to obtain the MAFs from the increments of PCs at the computational lag (by maximizing/minimizing the variance of increments, i.e. minimizing/maximizing the autocorrelation).

The MAFs in Fisboat

In Fisboat (work package WP2A), a set of spatial indices were selected to represent a target spatial population over its time series. The estimated indices presented notable variations in time. These may have been due to actual variations but also to various errors. MAFs can be used to extract from the set of indices the very first factors, that present the maximal continuity in time, and that can be thought to be used for a follow-up of the population in time.

Note: Here the continuity in time is measured at the lag of the time series: for instance a lag of 1 year if there is a survey every year, or the varying lag between successive surveys if there are gaps.

The interpretation of the MAFs

The very first MAFs (typically MAF1 and 2) allow us to extract trends in the multivariate time series of a set of indices. A jump upwards or downwards in the trends can be interpreted as a change in the spatial pattern of the considered age or functional group. Loadings informs us about the contribution of each index in the observed trends. Such an index can be used in an indicator approach to qualify a component of stock status (e.g. the spatial component). Moreover, correlation and delayed correlations allow us to put in relation MAFs and the abundance.

The limits of the MAFs

The number of MAFs cannot exceed the number of variables, nor the number of year increments (number of sampled years - 1). If the number of variables tends to be larger than the number of sampled years, the MAF n° i (i = 1, 2...) tends to have a period (number of years - 1)*2/i. In particular there will be evidence of a high continuity with period (number of years -1)*2 for MAF1, (number of years -1) for MAF2, etc, whatever the data, which may not be significant outside this series e.g. for additional years.

To prevent such an overfitting to the very detailed values of the variables, and so to increase the significance of the MAFs, these are computed while adding a repeated random noise to the variables. The noise, which is function of the number of variables and years, vanishes when the time series is longer.

Software

An R script has been written, MAF_noise_script_WP5.R, in which the solution of MAFs is obtained from a double call to the PCA standard routine prcomp(). In the first call, PCs are determined from the initial variables. In the second call, MAFs are computed from the increments of the PCs at the time lag.

MAFs are computed for each of 100 realizations where a random noise is added to the centered and normed indices. The random noise follows a Gaussian distribution with mean 0 and variance 0.1 * (nb.indices / (nb.years - 1)). For each resulting MAF, the 100 values of the loading of each index are available. Their distribution is symmetric. The median of the loading values (which is more robust than the mean) is used to determine the median MAFs profile.

The MAFs that present the lowest variogram (half variance of increments) at the lag are retained finally. By default the two first MAFs are retained. (Note: since each MAF has a variance of 1, it can be shown that a small value for the variogram at short time lags is compensated by some large values at larger time lags. This explains the large variogram values at large time lags that can be found for the first MAFs).

Output per age or functional group for a target spatial population:

For each retained MAF, coming from the 100 realizations:

- the median MAF time series;
- the median MAF variogram;
- the contribution (median of the loading coefficients) of each initial index into the median MAF.

For each retained MAF:

- the median MAF time series;
- the times series of the log of the abundance;
- the regression between the median MAF and the log of the abundance;
- the delayed correlation between the median MAF and the log of the abundance.

Important remark on the sign of MAFs:

Like a PC, a MAF is equivalent to its opposite (= the MAF with changed sign, that would be obtained by changing the sign of each coefficient of its linear combination), since the unit variance and the variogram at the computation lag would be unchanged. Then a MAF that would be monotonic over a time series can indifferently appear to be increasing as well as decreasing. Similarly a MAF with an extremum in the middle of a time series could indifferently present a maximum or a minimum.

References

Switzer, P. and Green, A. A. 1984. Min/max autocorrelation factors for multivariate spatial imaging. Technical report no 6, Department of Statistics, Stanford University, 14 p.

Erzini, K., Inejih, C.A.O. and Stobberup, K.A. (2005) An application of two techniques for the analysis of short, multivariate non-stationary time-series of Mauritanian trawl survey data. *ICES Journal of Marine Science* **62**, 353-359.

5.5 Diagnosing stock status from indicator series.

5.5.1 Combining trend signals using a cause-effects table

Verena M. Trenkel, Marie-Joëlle Rochet and Benoît Mesnil
Ifremer, Nantes, France

Introduction

This method provides guidance for an interpretation of time trends in indicators of the biological status of a stock derived from survey data (Gangl and Pereira 2003; Rochet and Trenkel 2003; Ault *et al.* 2005), and indicators of fishing pressure derived from knowledge of the fishery. Trends are interpreted relative to a past reference period. The aim is to estimate the current state of a stock of interest with respect to management objectives. A fuller account of the method together with a practical example of its application is given by Trenkel et al. (2007).

Method

At the preliminary stage, managers define operational objectives, for example, to obtain landings at a certain level, or individuals of a certain average size. Meanwhile, scientists and stakeholders decide upon a suite of indicators suitable for monitoring the stock and the fishery, possibly using the multi-stage framework proposed by Rice and Rochet (2005). Secondly, scientists examine estimated values of the selected indicators together with reports of any other studies that pertain to the stock at some time in the past, called the reference state, most probably the starting year of a survey time series. They then categorise the stock as having been either 'satisfactory' or 'unsatisfactory' for each indicator at that time. Thirdly, changes in indicator values after this reference time are estimated, interpreted, and combined into a diagnostic that highlights possible causes of the changes observed. Finally, this diagnostic is considered with the managment objectives, indicators of fishing pressure, and past experience of managing the fishery in order to decide appropriate managerial actions.

In more detail, having selected a suite of suitable indicators, there are several steps:

1. Select the reference time and calculate time series of indicators.
2. Determine time trends and status for each indicator in the current year.
3. Evaluate any other relevant information and combine the results of different indicators to provide an interpretation of the changes observed.
4. Agree a final diagnostic, including possible causes.
5. Determine trends in fishing pressure and propose appropriate management actions in the light of the diagnosis and stated management objectives.

For step 2, time trends in population indices from the reference year to the current year allow the current population dynamics to be assessed with respect to the reference situation. Uncertainty and natural variability in the survey data are accommodated through a hypothesis testing framework. A hypothesis test involves two risks of error, the type-I error of detecting a trend when none occurs, and the type-II error of not detecting a real trend. Whereas the α-level of the type-I error can be selected, the probability of type-II errors is generally not known, but it increases as the α-level decreases. Clearly there is a trade-off here between the type of error that is to be most avoided, so the selection of α-levels in hypothesis tests for time trends is a task for managers of the fishery. Time trends over the whole series provide estimates of longer term changes while trends over a more restricted number of years can inform on recent changes. Generally at least five years of data are required to be able to detect linear trends, in many cases, however, much longer series (>20 years) are necessary due to the high uncertainty and large random inter-annual variations in population indices (Nicholson and Jennings, 2004). The choice of the management time scale is clearly a decision for managers. Even with a multi-annual approach it is desirable to detect drastic changes in order to take rapid measures if necessary. For this, a choice of methods is available, as described elsewhere in this FISBOAT manual, e.g. CUSUM charts (Mesnil and Petitgas), and the method of second derivatives (Trenkel). For certain indicators reference points may exist, for example Z* for the total mortality rate (Die and Caddy, 1997) which can be used to evaluate directly the status of an index in the current year. The final aim of step 2 is to determine the direction of the most recent changes for each indicator, i.e. decreasing, stable or increasing.

Step 3 involves combining the results of several population indices. One well known method is the traffic light approach put forward by Caddy and others (Halliday *et al.* 2001; Caddy 2002). Depending on how many indicators are red, i.e. in an undesirable state, the overall assessment is set. For this approach the different indices take equal weights but they could just as well be weighted based on some *a priori* criteria. Rochet et al. (2005)proposed an alternative approach based on combining population and community indicators according to their biological meaning. Step 3 should also involve consideration of biological information additional to that provided by the indicator series (e.g. recruitment estimates, mean weight-at-age) in order to clarify the likely causes of the changes observed. Investigation of time-trends in indicators for fishing pressure (Piet *et al.* 2007) such as days-at-sea, fishing mortality, and quantities landed and discarded will allow corroboration of whether changes in fishing pressure could have been the major cause, before stating the final diagnosis (step 4).

The last step is to propose possible managerial measures linked to the diagnoses. The proposals should depend on whether the reference state was considered satisfactory or not, and whether fishing pressure increased since the reference year. The advice provides the direction of appropriate measures rather than prescribing them in quantitative terms, leaving the final decision to managers who should be guided by past experience.

Example

Here we apply this approach to an imaginary case of five population indicators: total mortality (Z), log-transformed abundance ln N, mean length $\overline{L}$, and length quartiles $L_{25\%}$ and $L_{75\%}$. Table 5.4.1 shows the increasing or decreasing effects (shown with arrows) that the most relevant factors are expected to have on each indicator. We assume that its long term time trend from the reference year to the current year has been categorised as either increasing, stable or decreasing. Using the method proposed by Trenkel (2006), the sign of the recent time trend for ln N is also obtained. Table 5.4.2 is then used to determine which time trends, short or long term, are used in table 5.4.1. Depending on agreement or not between long and short term trends in ln N, the estimated category of the short or long term trends determines which trend is used. The latter case occurs if a recent increase in ln N can be explained by recent management measures that reduce fishing while the long term trend is decreasing. Note that table 5.4.1 only considers single causes and not the expected effects of combined causes such as reduced recruitment and increased fishing. In certain cases additional information such as recruitment estimates, should be sought in order to clarify the causes of the observed changes. Thus, in the proposed approach, biological knowledge is used to point at the possible causes behind the observed changes, at least as far as being able to say whether increased fishing pressure could have contributed.

The last step consists of proposing possible managerial measures. The last column of table 5.4.1 makes suggestions for mitigating measures based on whether the reference state is considered to be impacted by fishing (1), or satisfactory (2). As the diagnosis is qualitative, so are the managerial measures. The direction rather than a prescribed amount of treatment is proposed. The quantitative decision is left to the managers. The adaptive management approach advocated in the 1980's by Walters (1986) seems a natural choice for implementing qualitative management advice, with a careful monitoring of response to the chosen policy being fed into the process for the next time step. The listed measures are by no means exhaustive and are only intended to give a flavour of the kind of advice that could be proposed.

Table 5.4.1: Expected effects of different causes on selected indicators and possible mitigation measures for counterbalancing changes. Δ stands for change, — for no effect, ↗ for increasing and ↘ for decreasing. Population indicators: Z = total mortality, ln-N = log-transformed total abundance, Lbar = mean length, $L_{25\%}$ and $L_{75\%}$ = length distribution quartiles.

Cause	***Z***	**ln *N***	**L-bar**	$L_{25\%}$	$L_{75\%}$	**Mitigation measure** 1 = impacted reference state 2 = satisfactory reference state
↗ fishing mortality	↗	↘	↘	—	↘	1 & 2 : −ΔF: reduction in overall fishing mortality
↘ fishing mortality	↘	↗	↗	—	↗	1: status quo 2: +ΔF allowed
↗ recruitment	—	↗	↘	↘	—	1: status quo 2: +ΔTAC: increase in TAC
↘ recruitment	—	↘	↗	↗	—	1 & 2: −ΔTAC: reduction in TAC
Faster growth	—	—	↗	—	↗	1: status quo 2 : Δselection pattern: decrease selectivity to smaller sizes
Slower growth	—	—	↘	—	↘	1 & 2: Δselection pattern: increase selectivity to larger sizes
Larger fish caught (Δ fishing area, stock distribution or gear)	↘	↗	↗	—	↗	1: status quo 2: Δselection pattern: decrease selectivity to smaller sizes
Smaller fish caught (Δ fishing area, stock distribution or gear)	↗	↘	↘	↘	—	1 & 2: Δselection pattern: increase selectivity to larger sizes

Table 5.4.2. Table for determining whether 'long term' or 'recent term' trends are used in table 5.4.1. If recent time trends are used, the reference state at the beginning of the recent period needs to be used!

<table>
<tr><th rowspan="2">Long term trend in ln N</th><th colspan="3">Recent trend in ln N</th></tr>
<tr><th>↗</th><th>—</th><th>↘</th></tr>
<tr><td>↗</td><td>long term</td><td>long term</td><td>recent term</td></tr>
<tr><td>—</td><td rowspan="2">if recent management measures intended to decrease fishing: continue else: as for long term</td><td>long term</td><td>recent term</td></tr>
<tr><td>↘</td><td>long term</td><td>long term</td></tr>
</table>

References

Ault, J. S., Smith, S. G., and Bohnsack, J. A. 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. ICES Journal of marine Science, 62: 417-423.

Caddy, J. F. 2002. Limit reference points, traffic lights, and holistic approaches to fisheries management with minimal stock assessment input. Fisheries Research, 56: 133-137.

Die, D. J., and Caddy, J. F. 1997. Sustainable yield indicators from biomass: are there appropriate reference points for use in tropical fisheries? Fisheries Research, 32: 69-79.

Gangl, R. S., and Pereira, D. L. 2003. Biological performance indicators for evaluating exploitation of Minnesota's large-lake walleye fisheries. North American Journal of Fisheries Management, 23: 1303-1311.

Halliday, R. G., Fanning, L. P., and Mohn, R. K. 2001. Use of the Traffic Light method in fishery management planning. CSAS Research Document, 2001/18, 41 pp.

Rochet, M. J. 2000. Does the concept of spawning per recruit make sense? ICES Journal of marine Science, 57: 1160-1174.

Rochet, M.-J., Trenkel, V., Bellail, R., Coppin, F., Le Pape, O., Mahé, J.-C., Morin , J., Poulard, J. C., Schlaich, I., Souplet, A., Vérin, Y., and Bertrand, J. A. 2005. Combining indicator trends to assess ongoing changes in exploited fish communities: diagnostic of communities off the coasts of France. ICES Journal of marine Science, 62: 1647-1664.

Rochet, M. J., and Trenkel, V. M. 2003. Which community indicators can measure the impact of fishing? a review and proposals. Canadian Journal of Fisheries and Aquatic Sciences, 60: 86-99.

Trenkel, V. M. 2006. Nonparametric method for determining recent trends in indicator time series. Working document for Fisboat project. June 2006, 7 pp.

Walters, C. 1986. Adaptive management of renewable resources. Biological resource management, Macmillan Publishing Company, New York. 374 pp.

5.5.2 A 'traffic light' procedure based on Cusum out-of-control tables.

Pierre Petitgas,
IFREMER, Nantes,
France

Introduction

Usually, more than one indicator is used to assess the status of a fish stock. The important question then arises of how to combine the various results into a single assessment. One simple, illustrative way of doing so is to use a 'traffic light approach' (Caddy et al., 2005). An alternative approach is that of multivariate statistical process control. Both are complementary as the CUSUM diagnostic table with all indicators may assist interpretation of multivariate alarm signals. Part II (section 5.9) of this 2-part section illustrates the multivariate statistical approach.

Example

To illustrate, a CUSUM procedure (see Mesnil and Petitgas elsewhere in this manual) was applied separately to time-series of each of several attributes (indicators) for the North Sea cod stock studied under FISBOAT and available at http://www.ifremer.fr/drvecohal/fisboat/. The CUSUM procedure comprises three steps (Hawkins and Olwell, 1997; Montgomery, 2005). First, a reference period is defined as a period when the health of the stock was considered acceptable (in-control). This serves to estimate the in-control reference mean and variance for each indicator. The same reference period was applied to all indicators with the results shown in table 5.8.1. Second, the CUSUM was tuned for each indicator to signal important deviations from the reference mean in years outside the reference period. The tuning of the CUSUM results in statistically defining the false alarm rate, and the no alarm rate associated with the in-control limits that are set to enclose acceptable deviations from the reference mean. The application of the CUSUM to each indicator for the cod stock resulted in an array of deviations from the reference mean vector expressed in standard deviation units. This is the CUSUM diagnostic table shown in table 5.8.2. Each column of the array corresponds to each indicator time series of deviations. Because the deviations being expressed in units of standard deviation, comparisons between indicators are immediate. Setting the non-alerting deviations to zero, the diagnostic CUSUM table provides the quantitative values of the deviations from the reference means with a + or – sign which trigger alarm signals. Clearly, in some years only a small number of indicators signal alarms, some perhaps with high deviations, whereas, in other years, many will signal. Here, we used expert judgement in assigning each year as 'in-control' or 'out-of-control' based on which indicator signalled alarm and how many of them there were (table 5.8.2). Cells in the table may be coloured red, orange, or green, as for traffic lights, to show at a glance the perceived seriousness of the state indicated.

Table 5.8.1: Parameters of the CUSUM control scheme for North Sea cod for the biological indicators. Reference period is 1985-1994. Parameters are: mu, sd: mean and standard deviation in reference period; k: allowance in sd units; h: decision interval in sd units; ic.arl (years): in-control ARL (average run length to a false alarm) also noted ARL(0); ic.rl.25 (years): RL value at the first quartile of the RL distribution; oc.arl (years): out-of-control ARL (average run length to signal real change) also noted ARL(2k). Indicators are: Survey index: abundance index for ages 1 to 6; Recruit index: abundance at age 2; Lbar: average length in the population; L25: length value of the first quartile; L75: length value of the third quartile; L50 maturity: length value at which 50% of the population is mature; Z: apparent total mortality.

	Survey index	**Recruit index**	**Lbar**	**L25**	**L75**	**L50 maturity**	**Z**
mu	19.12	18.00	34.77	20.69	41.70	65.44	1.12
sd	0.26	0.77	4.80	5.16	6.45	5.24	0.44
k	1.3	0.9	1.2	0.9	0.8	1.1	1.0
h	1.0	1.0	1.0	1.0	1.2	1.1	1.0
ic.arl	79.3	27.5	60.0	27.5	30.0	56.2	35.3
ic.rl.25	23.0	8.0	17.0	8.0	9.0	16.0	10.0
oc.arl	1.5	1.9	1.6	1.9	2.3	1.8	1.8

Table 5.8.2: CUSUM diagnostics table for North Sea cod using biological population indicators. Values are the deviation from the reference mean for each indicator in standard deviation units. Reference period is 1985-1994. The procedure signals an alarm from 1999. Indicators are: Survey index: abundance index for ages 1 to 6; Recruit index: abundance at age 2; Lbar: average length in the population; L25: length value of the first quartile; L75: length value of the third quartile; L50 maturity: length value at which 50% of the population is mature; Z: apparent total mortality.

Year	Survey Index	Recruit index	Lbar	L25	L75	L50 mat-urity	Z	diag-nostics
1985	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1986	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1987	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1988	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1989	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1990	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1991	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1992	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1993	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1994	0.00	0.00	0.00	0.00	0.00	0.00	0.00	ref
1995	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
1996	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
1997	0.00	0.00	**-1.84**	**-1.03**	**-2.21**	**-1.85**	0.00	
1998	0.00	0.00	0.00	0.00	**-2.00**	**-3.44**	0.00	
1999	**-1.27**	**-1.30**	0.00	0.00	0.00	**-7.36**	0.00	alarm
2000	**-1.65**	0.00	0.00	0.00	0.00	**-9.34**	0.00	alarm
2001	**-3.04**	0.00	0.00	0.00	0.00	**-9.84**	0.00	alarm
2002	**-3.96**	0.00	0.00	0.00	0.00	**-12.78**	0.00	alarm
2003	**-7.48**	0.00	0.00	0.00	0.00	**-15.95**	0.00	alarm
2004	**-10.50**	**-1.18**	0.00	0.00	0.00	**-19.33**	0.00	alarm
2005	**-14.97**	**-2.02**	0.00	0.00	**1.23**	**-23.10**		alarm

References

Caddy, J., Wade, E., Surette, T., Hebert, M. and Moriyasu, M. 2005. Using an empirical traffic light procedure for monitoring and forecasting in the Gulf of St. Lawrence fishery for the snow crab, Chionoecetes opilio. Fisheries Research, 76: 123-145.

Hawkins, D. and Olwell, D. 1997. Cumulative Sum Charts and Charting for Quality Improvement. Springer Verlag, New York.

Montgomery, D. 2005. Introduction to statistical quality control. Wiley, New York. 5th Edition.

5.5.3 A multivariate statistical procedure.

Pierre Petitgas,
IFREMER, Nantes,
France

Introduction

Part I of this presentation (see section 5.8) illustrated the 'traffic light approach' for combining results from time-series of several indicators into a single assessment of a fish stock. A complimentary approach is to use multivariate statistical methods. An illustrative example is given here. More detail on the PCA and MFA approaches discussed can be found in sections 5.10 and 5.11 of this manual, respectively.

Multivariate statistical process control methods use the relationships existing among control variables to prevent having to deal with false alarms in the many individual control charts. Thus multivariate process control methods are potentially more efficient than control methods based on the analysis of a collection of univariate charts (e.g., Hawkins and Olwell, 1997). Various multivariate control methods are available. Hotelling's T^2 statistic (Hotelling, 1947) is the analogue in the multivariate case of the Shewart chart in the univariate case (e.g., Hawkins and Olwell, 1997). It is best suited to detect large shifts in the mean as it uses the current sample only at each time step. To detect rapidly small shifts in the mean, consecutive samples need to be considered for which multivariate CUSUM methods (Crozier, 1988) and multivariate exponentially weighted moving average (EWMA) methods (Lowry et al., 1992) have been developed. Scranton et al. (1996) used multivariate EWMA on a reduced number of principal components with increased shift detection capability. In environmental monitoring, Manly and MacKenzie (2000) made use of Principal Component Analysis (PCA) to project multivariate observations in a factorial space and assess whether they were inside or outside the area within which the process could be considered in control.

Example

To illustrate a multivariate approach, the same CUSUM results for North Sea cod discussed previously in relation to the Traffic Light approach (section 5.8) will be used with the addition of parallel time series for some spatial indicators. The approach considered is inspired by the PCA procedure of Manly and MacKenzie (2000). First, the centre of gravity in factorial space was estimated for the in-control reference years. Next, the distance in factorial space of the observation in each year to that centre of gravity was computed and a new time series constructed from them. Then a CUSUM procedure was applied to that distance. PCA-based distances were applied to the biological (non-spatial) indices, while, for the spatial indices, MFA-based distances (see section 5.11) were used. Evolution of the stock can then be summarised with two distances, one for the spatial and one for the non-spatial indicators. A traffic light type table with two columns (one for each distance) of CUSUM deviations from the in-control domain represents simply the multivariate process control scheme for all the attributes of the stock (tables 5.9.1, 5.9.2).

Table 5.9.1: Parameters of the CUSUM control scheme for North Sea cod for the multivariate indicators. Reference period is 1985-1994. Parameters are: mu, sd: mean and standard deviation in reference period; k: allowance in sd units; h: decision interval in sd units; ic.arl (years): in-control ARL (average run length to a false alarm) also noted ARL(0); ic.rl.25 (years): RL value at the first quartile of the RL distribution; oc.arl (years): out-of-control ARL (average run length to signal real change) also noted ARL(2k). Indicators are: Spatial.mul.mfa: MFA-based multivariate distance for the spatial indices; Biol.mul.pca: PCA-based multivariate distance for the biological indices.

Parameter	Spatial.mul.mfa	Biol.mul.pca
mu	1.35	1.67
sd	0.23	0.76
k	0.9	1.5
h	1.2	1.0
ic.arl	39.2	142.2
ic.rl.25	11.0	41.0
oc.arl	2.1	1.4

Table 5.9.2: CUSUM diagnostics table for North Sea cod using multivariate indicators, one for all the spatial indices and the other for all the biological indices. Values are the deviation from the reference mean for each indicator in standard deviation units. Reference period is 1985-1994. The procedure signals an alarm from year 2000. Indicators are: Spatial.mul.mfa: MFA-based multivariate distance for the spatial indices; Biol.mul.pca: PCA-based multivariate distance for the biological indices.

Year	Spatial.mul.mfa	Biol.mul.pca	diagnostics
1985	0.00	0.00	ref
1986	0.00	0.00	ref
1987	0.00	0.00	ref
1988	0.00	0.00	ref
1989	0.00	0.00	ref
1990	0.00	0.00	ref
1991	0.00	0.00	ref
1992	0.00	0.00	ref
1993	0.00	0.00	ref
1994	0.00	0.00	ref
1995	**1.40**	0.00	
1996	0.00	0.00	
1997	0.00	**1.46**	
1998	0.00	0.00	
1999	0.00	**3.46**	
2000	**1.77**	**3.80**	alarm
2001	**2.66**	**4.06**	alarm
2002	**2.00**	**4.63**	alarm
2003	**1.69**	**8.36**	alarm
2004	**2.42**	**11.96**	alarm
2005	**2.85**	**17.13**	alarm

References

Crosier, R. 1988. Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. Technometrics, 30: 291-303.

Hawkins, D. and Olwell, D. 1997. Cumulative Sum Charts and Charting for Quality Improvement. Springer Verlag, New York.

Hotelling, H. 1947. Multivariate Quality Control. *In*: Techniques of Statistical Analysis, Eds. C. Eisenhart, M. Hastay and W. Wallis, New York, McGraw-Hill, pp.111-184

Lowry, C., Woodall, W., Champ, C., Rigdon, S. 1992. A Multivariate Exponentially Weighted Moving Average Control Chart. Technometrics, 34: 46-53.

Manly, B. and MacKenzie, D. 2000. A cumulative sum type of methods for environmental monitoring. Environmetrics, 11: 151-166.

Scranton, R., Runger, G., Keats, B. and Montegomery, D. 1996. Efficient shift detection using multivariate exponentially weighted moving average control charts and principal components. Quality and Reliability Engineering International, 12: 165-171.

Appendix 1

Format for documenting indicators

The format adopted here for presenting indicators is derived from
Halliday, R.G. and Mohn, R. (2001). Proceedings of the Fisheries Management Studies WG, 8-11 January 2001. Canadian Science Advisory Secretariat, Proceedings Series 2001/08 (Appendix 5, pp. 45-48).

Description of indicators/indices

INDEX : descriptive name + acronym
<u>Description</u> : short description of what it is (and for case studies in second stage, survey in which it is measured, e.g. "mean length of NEA cod caught during winter Barents Sea surveys")

<u>Stock attribute</u> :
- attribute(s) that the indicator is deemed to reflect (e.g. abundance, productivity, recruitment, mortality, ecosystem, …)

<u>Derivation</u> :
- document briefly how the index is derived from raw data for each station, and then integrated for the whole survey year (+ ref. published manuals for details).
- and how variance of total index is obtained

<u>Reference points</u> :
- bases for setting the RPs (as target, limit or trigger)
- alternative choices and their rationale

<u>Interpretability</u> :
- how does the indicator reflect stock status or the identified attribute?
- what caveats exist regarding interpretation? (e.g., whole stock vs. population in survey area, if they differ)
- processes involved in changes (is indicator's response specific?)

<u>Measurability</u> : (*keyword here is*: confidence in estimates of the indicator)
- statistical properties of estimator (variability, bias, skewness, …)
- transformations required before use
- alternative formulations for the same estimators
- alternative estimators of same indicators ; pros & cons

<u>Sensitivity</u> :
- how rapidly & accurately does indicator respond to changes in stock status?
- does natural variability likely masks real changes?

<u>Review of performance :</u>
- performance of the indicator in hindsight, to infer stock status
- document the adequacy of the indicator, its estimator and its RPs (or problems encountered for the specific case study)

<u>References :</u>
(e.g., for case studies, references to published manuals of procedures, articles in which the indicator has been applied in advisory context, …)

Document 2: Indicator-Based Assessment – Application of methods

ICES CM 2007/O:16

Comprehensive indicator-based diagnostics of fish stocks using fishery-independent survey data: the FISBOAT report on case studies

Pierre Petitgas [(1)], J.-C. Poulard [(1)], K. Radtke [(2)], M.-T. Spedicato [(3)], L. Ibaibarriaga [(4)], C.-Y. Politou [(5)], K. Korsbrekke [(6)], C. Deerenberg [(7)] and P. Fernandes [(8)]

[(1)]: IFREMER, France; [(2)]: SFI, Poland; [(3)]: COISPA, Italy; [(4)]: AZTI, Spain; [(5)]: HCMR, Greece; [(6)]: IMR, Norway; [(7)]: IMARES, Holland; [(8)]: FRS, Great Britain

Research fisheries surveys are now implemented as monitoring programs of fish stocks and provide a large set of measurements on the evolution of their state. Here we show how fishery-independent diagnostics of fish stocks can be achieved using a comprehensive set of indices and analysis procedures inspired from environmental monitoring.

We present fish stock indices, analysis procedures and diagnostics results for nine stocks in European waters. The set of indices considered comprises two population abundance indices, four indices for population vital traits and nine indices for spatial organisation by age. The indices are combined and selected using multivariate techniques that maximise correlation between variables and also continuity in time. Trend detection and quality control techniques are then applied on the time series of the combined and selected indices. Based on these analyses diagnostic tables are filled, leading to comprehensive indicator-based diagnostics of fish stocks.

Similar analysis procedures are applied to all case studies and results are reported using standardised templates. The application to a wide range of fish stocks in different health conditions with different behaviours and past histories demonstrates the potential of the tools and indices for delivering diagnostics in operational mode. The paper is intended to be a manual summarising the results of the EU-project Fisboat (Fishery-independent survey-based operational assessment tools) for general use outside the project.

Key words: Fishery-independent assessment, indicators, quality control, spatial statistics, vital traits, anchovy, hake, cod, herring, red mullet.

Contact author: Pierre Petitgas: IFREMER, BP 21105, 44311 cdx3 Nantes, France [tel: +33 240 374 163, fax: +33 240 374 075, e-mail: pierre.petitgas@ifremer.fr].

1. Introduction

Can survey data by their own be used to make diagnostics on fish stock health? If so, with which methods? The UE project FISBOAT (Fishery Independent Survey-based Operational Assessment Tools, 2004-2007 ; http://www.ifremer.fr/drvecohal/fisboat/) was set up to answer these questions. The project investigated two approaches. One was the simulation evaluation of traditional stock assessment analytical models (Kell et al., 2007) using abundance survey indices at age. The question was: given that survey abundance indices are non absolute and given the uncertainty in the estimates, what harvest control rule can be set to manage stock abundance levels? The other approach, which is our interest in this paper, was the development of a comprehensive indicator-based monitoring methodology making full use of all the biological information (not just the abundance at age indices) available in survey data. Metrics characterising fish stock attributes (here after termed indicators) were developed and estimated from survey data, resulting in the construction of time series of a variety of indicators of stock attributes. Such indicators were used as control variables with which the state of the stock was monitored. Methods for analysing the time series of indicators were developed as well as methods for making diagnostics based on the analysis of the indicator time series. Indicators and their methods are documented in a companion paper (Cotter et al., 2007). The project methodological developments resulted in the set up of statistical monitoring procedures of fish stock status using a comprehensive list of indicators of stock attributes. In this paper we summarise the results obtained by applying the methods to the indicators on the project case studies.

The project case studies scanned nine different stocks across European waters in the demersal and pelagic domains with different vital traits and stock histories and survey methodologies. The case studies were: cod in the Barents sea the Baltic and the North Sea, hake in the Bay of Biscay, the Ionnian sea and the Aegean sea, herring in the North Sea, anchovy in the Bay of Biscay and red mullet in the Thyrhenean sea. Case study individual reports followed the same template and these have been the basis for the present compilation. Individual reports are available on Fisboat website at http://www.ifremer.fr/drvecohal/fisboat/. Three major steps were followed. First, indicators of population attributes were calculated and time series of indicators were produced. Then, the time series of indicators were statistically analysed to detect changes. Last, results of the previous step for the variety of indicators were combined in diagnostic tables to formulate a diagnostic.

The statistical identification of changes in the time series relied on the definition of a reference period to which compare the indicator values for years outside that period. The reference period was defined as that in which the stock status was thought to be acceptable, based on historical knowledge. Such strategy is similar to that in statistical process control, where two phases are distinguished (e.g., Montgomery, 2005). Phase I is where the process is sampled to acurately define the ‘in-control’ state of the indicators. Phase II is the monitoring phase where statistical procedures are applied to detect any departure from the ‘in-control’ state. Phase I was here replaced by the definition of a reference period. We shall not discuss the definition of the reference period for each stock. We shall be concerned only by the monitoring of the stock relatively to the reference defined. The diagnostic is then relative to a reference and not absolute.

The population indicators were raw indicators estimated from the survey data (e.g., mean length in the population or gravity center in the spatial distribution) as well as multivariate combined indicators derived from the raw indicators (e.g., principal components or departure from a reference domain in factorial space). The time series were analysed to detect trends and changes in trends between different sets of years (e.g., trend over all years as compared to trend in the recent years only). Also the Cusum statistical control scheme was used to detect changes in the mean along the series. The Cusum procedure led to the construction of a traffic light type diagnostic table were departures from reference values triggered alarm signals with set risks of false alarm and non alarm. The trend analysis procedure led to the construction of a cause-effects diagnostic table were trends of different indicators were combined and interpreted using background biological knowledge. In all, changes were detected and were assigned causes when possible. The comprehensive indicator-based monitoring system developed produced coherent results which should complement the traditional assessment and thus increase the reliability of diagnostics on fish stocks.

2. Methods

In this section we summarise the procedures applied to the case study fish populations (Table 1) in order to explicit the monitoring system of population status that was set up based on fisheries research surveys only. For each of the indices and methods used below, documented computer code in R is available on Fisboat website at http://www.ifremer.fr/drvecohal/fisboat/.

How were populations described ?

The evolution of the state of populations was characterised by time series of a variety of indicators.

Raw indices. These were estimated directly from the survey data. Two groups of indices were considered: biological (non spatial) and spatial indices (Tables 2 and 3). Biological indices were estimates at population level to characterise abundance, recruitment, length structure, maturity and mortality. Spatial indices characterised the different aspects of a map: location, dispersion, patchiness, occupation, correlation, aggregation. Spatial indices were estimated by age to characterise the spatial distribution in each age and thus characterise the spatial organisation of the life cycle. The biological and spatial indices are fully described in Cotter et al. (2007) and in Woillez et al. (2007a).

Multivariate indices. These were derived from the raw indices and were (composite) multivariate summaries of the many raw indices considered. They were defined as multivariate distances to the gravity center of the reference period. Principal Components Analysis (PCA) was used for constructing a multivariate biological index and Multiple Factorial Analysis (MFA) was used for constructing a multivariate spatial index as the spatial information was 3D (indices, ages and years). PCA and MFA allowed to evidence the linear correlations existing between the indicators. For the biological indices the PCA-based index was the distance in the first factorial plane between the position of the gravity center of the reference period and that of the current year. For the spatial indices the MFA-based index was the sum over all ages of age-specific distances. Each age-specific distance was calculated in the first factorial plane between the position of the age-specific gravity center for the reference period and that of the current year. The multivariate indices are fully described with their methods in Cotter et al. (2007) and Woillez et al. (2007a).

Selection of raw indices. The MAF method (Min/Max Autocorrelation Factors) was used as an automated procedure to select those indices that best summarised the multivariate information with highest continuity in time. The MAF method allowed to construct pincipal components (factors), the autocorrelation of which decreases from the first factors to the last ones. The very first factors (MAFs) extracted the part of the multivariate information which was the most continuous in time. Therefore the loadings of the indices on the two first MAFs were used to select those indices that showed highest continuity in time as well as carrying the most of the multivariate information. The MAF method is fully documented in Cotter et al. (2007) and in Woillez et al. (2007b).

How were changes identified in the indicators time series ?

Change in population status was identfied by analysing the indicator time series. The detection of linear trends and changes in trends were considered. Another anlysis was the detection of shifts in the mean value of the indicator relatively to that in the reference period using the Cusum control charts as in industrial quality control.

Trend plots. Linear trends were estimated and their significance tested using the p-value that measures the risk of type-I (risk of identifying a trend when non exists). The linear trend in the last years of the series were also tested so as to detect change in the slope between the long-term trend and that in the last years of the series. A method based on the value of the second derivative was used to identify change points and detect change in slope for the last years. The derivative's method is fully documented in Cotter et al. (2007).

DI-Cusum plots. Here, we are interested in detecting shifts in indicator mean level relatively to that in the reference period, irrespective of the type of change, whether linear or not. The decision interval form of the Cusum was used. Values outside the interval were considered significantly different than those in the reference period (in-control) and therefore out-of-control. The in-control interval was statistically defined with set risks of false alarm and no-alarm rates. The DI-Cusum procedure is fully documented in Cotter et al. (2007) and in statistcal quality control litterature (e.g., Montgomery, 2005).

How were diagnostics made and interpreted ?

Results of the analyses of the many indicator time series were combined in diagnostic tables to elaborate a diagnostic of the state of the populations. Each method (trend and di-cusum) led to a particular diagnostic table (full documentation in Cotter et al., 2007).

Trend analysis: interpretation using cause-effects tables. A particular cause inducing variation in biological indices can be translated into an expected combination of trends in the indices, e.g., an increase in fishing mortality is expected to translate into an increase in Z, a decrease in Lbar and a decrease in Ln-Ntot. The cause-effects table (Table 4) provided a list of causes with their expected resulting combination of trends in the indicators, thus helped identify potential causes to the observed trends in the indicators (Trenkel et al., 2007).

Di-Cusum analysis: interpretation using traffic light tables of out-of-control signals. The application of the Di-Cusum to each indicator resulted in an array of out-of-control deviations from the reference mean vector. This was the Cusum diagnostic table. Each column of the array corresponded to each indicator time series of deviations. Setting the non-alerting deviations to zero, the diagnostic Cusum table provided the quantitative values of the deviations from the reference means with a + or – sign which triggered alarm signals. Cells in the table may be coloured red, orange, or green, as for traffic lights, to show at a glance the perceived seriousness of the state indicated.

Summary sheet and case study reporting. The results for each case study were reported in a summary sheet with a defined format. The sheet documented the survey time series, the indicators were used (raw, multivariate), the reference period, the methods were used to analyse the indicator time series, the resulting diagnostic on the stock. A template for reporting case study results was defined which comprised the following items: Data, Looking for change, Interpretation, What has been learned, Summary sheet, Comparison with traditional assessment, Formulation of advice. The template is annexed.

3. Results from the case studies

Indices characterising population status

The different indicators (raw and multivariate) that were computed in each case study were compiled in Table 5. PCA applied on the raw biological indices revealed a strong correlation structure between the indices. Across the different case studies, the correlation structure showed some consistency as can be seen from the loadings of the indices in the principle components (Table 6). The first principal component was always made of the length indices, which are much correlated to each other. Depending on the case study, the second component was correlated either to abundance, recruits or Z. For most stocks, only the first 2 components could be interpreted with high enough loadings of particular indices. For 3 stocks only, the third component was well related to one index, either Z or Recruits. The fact that length indices, abundance indices and Z did not always show a correlation structure easily interpretable (e.g., opposition) was perhaps due to the fact that no time lag was considered in the analysis. The multivariate biological index was a measure of inter-annual departures from the correlation structure as observed in the reference period.

MFA applied on the spatial indices also revealed marked and progressive changes in the spatial distributions with age (Tables 8a-b; Figs. 1a-c), which characterised the life cycle spatial pattern in each case study. Across all case studies, the area indices, inertia and gravity center were the spatial indices that were mostly invilved in explaining best the principal components.
For cod in the Barents Sea, the different ages were progressively aligned along the first component. Spatial distributions of young ages were less dispersed, more to the East and occupied less area than for older ages. Spreading area decreased slightly in the mid-ages (A4-7).
For cod in the North Sea, young (A1) and old ages (A5-6) differred on the first component from intermediate ages (A2-4). Spatial distributions of young and old ages were more to the East, less dispersed, occupied smaller area, and were more uneven (higher microstructure or nugget effect) than for intermediate ages. Age 1 and Ages 5-6 differed on the second component by the location of their centre of gravity and anisotropy. The spatial distribution of old ages was more to the north, more anisotropic, and occupied a smaller area than that of the age 1 fish.
For cod in the Baltic, young (A1-2) differred from old ages (A3-5) on the first component. Spatial distributions of young ages were more to the South, less dispersed and less anisotropic than old ages. Spatial distributions of ages A1-2 and A5 differred from that of other ages on the second component by positive area occupied. Ages A1-2 and A5 occupied a smaller less area than ages A3-4.
For herring in the North Sea, young (A1-2) differred on the first component from old ages (A7-9). Spatial distributions of young (immature) ages were more to the East and South, less dispersed and less spread than older ages (A4-9). The acquisition of maturity marked a clear difference in the spatial distribution for ages A2-3 as the distribution of mature A2-3 were more alike than that of older ages A4-7. Spatial distributions of mature ages A2-A9 occupied larger positive areas with age, which was visible on the second component.
For hake in Biscay, young (A0-3) differred on the first component from old ages (A4-5). Spatial distributions of the old A4-5 were more to the West occupying a larger area with more spread. Ages A0-1 differred on the second component from the other ages as their spatial distribution was more anisotropic.
For hake in the Ionian Sea, young (A0-3) differred on the first component from old ages (A4-5). Spatial distribution of the old A4-5 were more to the South and West, occupying a smaller area and were more uneven (larger microstructure index). Age A0 differred on the second component from the other ages as its spatial distribution was more anisotropic.
For hake in the Aegan Sea, ages A0 and A5 differred on the first component from ages A2-3. Their spatial distributions were more to the North and West and were less dispersed with smaller spreading and equivalent areas. The second component distinguished the spatial distribution of ages A0-1 from that of A4-5 as the young ages A0-1 occupied a larger positive area.
For anchovy in Biscay (acoustic surveys), the spatial distribution of ages A1-3 had similar characteristics, though A1 was slightly more dispersed and anisotropic. The spatial distribution of the anchovy eggs shared similarities to that of the adults but was also different (Table 8b). Both adult fish and egg distributions showed the same opposition on the first component between the area indices and the longitude of the gravity centre. The microstructure index (uneveness in the distribution) was characteristic of the egg distributions (component 2) which was less important in characterising the distribution of the adult fish. The anisotropy index characterised the adult fish distribution (component 2) but did less so for the egg distributions as that index corrrelated to component 3 of the egg distributions.
For red mullet in the Thyrrhrenian Sea, the characterisation of the spatial distributions have been separated in two sub region with marked different orientations, GS10a (western coast of mainland Italy) and GS10b (northern coast of Sicily). In GS10a, Ages A1 and A2 differred on the first and second components. Age A1 was more distributed to the SE and more uneven but occupying a larger

positive area. In GS10b, age A1 differred from A3 on the first component. Age A1 was distributed more to the East with larger equivalent and spreading areas and less uneven than A3.
The multivariate spatial index was a measure of inter-annual departures from the average spatial patterns as observed in the reference period.

Identification of changes and formulation of diagnostics

Table 9 compiles what methods were applied on what indicators in each case study. We now summarise the results obtained in each case study.

Cod in the Barents Sea (Fig.2). Time series of raw indices were visually inspected. The time series of Ln.Ntot, Ln.Rec and Z showed clear troughs at the begining of the series (90-94). This particular situation made it difficult for the trend methods to capture the signals due to scale and position of the changes in the time series. In constrast, the Cusum method was able to detect these changes (note that the reference period was at the end of the time series). The multivariate indices with the Cusum analysis allowed to achieve a diagnostic. Both spatial and abundance indices triggered alarms at the beginning of the 90s. The series of survey Z compared well with that of the ICES VPA esitmate. The survey coverage may be hampered by the presence of sea ice in the eastern Barents Sea, limiting the use of the survey indicators.

Cod in the North Sea (Fig.3a-b). In constrast to Barents Sea cod, North Sea cod showed clear trends in many indices, either long term or in the recent years since 2000. Trend and Cusum methods agreed and raw indices and multivariate indices were in agreement as well. The MAF selection of raw indices selected the following indices as carrying the variability in the stock: L50.matu, Ln.Ntot.matures, PA.matures, xcg.matures, ycg.matures, MI.recruits, MI.immatures, ycg.recruits, Anisotropy.recruits, ycg.immatures. Length at maturity has been decreasing all along the survey time series, total abundance and recruits decreased seriously since 2000 and so did the spatial indices of area and location with more northerly distribution of old fish but also recruits. Out-of-control alarm signals were confidently trigerred with the Cusum diagnostic table since 2003 as all indices have been out-of-control since that year. An alarm could be trigerred as early as 2000, if less weigth was given in the analysis to the length indices. Recent trends in were estimated for the last 5 years using the derivatives methods. The cause-effects table and the trend results table suggested that the closest cause to the recent trends identified was an increase in fishing mortality.

Cod in the Baltic Sea (Fig.4). The survey series began in the mid-90s at a time when the stock was already at a low abundance level. Therefore the survey series could not trace the historical evolution of the stock but its recent evolution within a degraded state. The index L50.matu was unreliable because of the seasonal timing of the survey. Visual inspection of the raw indices suggested that abundance at age A5 and Positive area of A5 showed obvious long-term decreasing trends. The other indices contained much variability. Recent trends were estimated for the last 5 years using the derivatives methods. Comparison of the cause-effects table with the trend results table suggested that the closest cause to the recent trends identified was an increase in fishing mortality. Based on age A5 series, the Cusum traffic light diagnostic table suggested to signal alarms since 2000. Results were in agreement with ICES assessment.

Hake in the Bay of Biscay (Fig.5). The time series in the different indices were variable enough to make visual inspection difficult. Trend analysis revealed no long-term trend but the derivatives method identified recent trends in length indices, Z and some spatial indices. It is noteworthy that the derivatives method identified changes where a linear approach did not. The recent increase in the length indices together with an a recent increase in Z were inconsistent with the cause-effects table and therefore difficult to interpret. The Cusum also diagnosed increase in Z and L25 indices in the recent years. Therefore that increase was considered real. The selection of indices using the MAF procedure resulted in selecting the area indices for the older fish: SA.A4, EA.A5, SA.A5, PA.A5, and xcg.A3. The Cusum procedure identified changes for these indices when the trend method did not, supposedly because of the type of variability in the time series. The old ages A4-5 showed decreased abundance, decreasing area indices and the age A3 a shift of its gravity centre to the West. The multivariate spatial index gave similar results with the Cusum as the MAF selected indices. Departure of the multivariate biological index from its reference domain had different causes that can be assigned based on the PCA loadings of the indices and the Cusum diagnostic table of the raw indices: in 98 total abundance A1-5 is low, in 99 and 2003 L25 has increased, in 2004, recruitment (A0) has increased. In all, though some amelioration of recruitment occurred in 2004, deterioration of abundance and spatial indices for old ages justified signaling alarms since 2000.

Hake in the Ionian Sea (Fig. 6). The survey series was short (1994-2003) and all indicator time series had high variability. Trend and Cusum methods did not agree in the fluctuations that could be identified, due to the variability in the series. The derivatives trend method identified declines in the last 5 years for the Length indices while the Cusum detected no out-of-control fluctuation in these indices. The trends identified in the biological indices could not be interpreted using the cause-effects table as the combination of trends was inconsistent with any of the causes suggested in the table. The multivariate biological index was declared out of control by the Cusum analysis for years in which the recruitment index was high (1995, 2003). The multivariate spatial index was declared out of control for years within the reference period. Given the intrinsic variability in the time series, a longer series seemed necessary to formulate any diagnostic.

Hake in the Aegean Sea (Fig.7). As for Ionian hake, the survey series was short (1994-2003) and all indicator time series had high variability. Trend and Cusum methods did not agree in the fluctuations that could be identified, due to the way in which the variability is disposed in the series. Here the trends method identified no trend while the Cusum identified poor abundance until 1997 as well as positive and negative alarms in L25, L50 and L75 until 1997. At the beginning of the series (1994), the abundance is low and is progressively increasing until 1997. The out-of-control alarms on the length indices could have resulted from the poor abundance, in coherence with the cause-effects table. The Cusum analysis triggered out-of-control signals for the multivariate biological index in 1994-95 as a result of low abundance and increase in length indices. Until 1997, the multivariate spatial index is identified to be out-of-control by the cusum analysis. The diagnostic is thus an alarm signal at the beginning of the series in the years 1994-97: poor abundance, decrease in length, departure in the spatial distribution.

Herring in the North Sea (Fig.8a-c). This case study has been analysed using multivariate methods and Cusum analysis only. Similarly to Barents Sea cod, the time series of abundance showed a clear trough in the middle of the series, peaking low in 1994, the increasing in the recent years until 2002. The selection of indices using the MAF procedure resulted in selecting those raw indices that carried the major signals. Amongst these, 6 indices were selected by visual inspection: Ln-N.matures, Ln-imatures, xcg.matures, ycg.matures, I.matures, SA.matures. The mature fish decreased in abundance reached a low in 1994 and increased again in 2000-02. During the low abundance period, the fish was less northerly distributed but came back latitude of the gravity center came back to previous values with increasing abundance. It is noteworthy that some spatial indices did not came back to their previous values of before the abundance low. In particular the spreading area, the equivalent area and the inertia have stayed low even after the abundance rebuild. Also the longitude of the gavity centre stayed to the West and did not came back to previous values. The abundance of imatures has increased in recent years, rebuilding the population. The Cusum analysis of the multivariate spatial index revealed out-of-control values during the years of low abundance. But the multivariate biological index was not so much influenced by the decrease in the old fish in the mid-90s probably because of the small response of other biological indices (length indices). In contrast, the multivariate biological index responded to the increase in young fish after the mid-90s. Its Cusum analysis identified out-of-control values for the recent period that revealed increase in abundance. In the Cusum diagnostic table, the multivariate spatial index revealed the period of alarm while the multivariate biological index that of recovery. The choice of reference period (1989-93) perhaps influenced the sign of the out-of-control signals for the multivariate biological index.

Anchovy in the Bay of Biscay (Fig.9). The anchovy population in the Bay of Biscay is monitored by two independent surveys in spring, an acoustic survey and an egg survey. Here, the acoustic survey provided the biological indices and the spatial indices at age and the egg survey provided the spatial indices for the eggs. In the last last years of the survey series, total abundance and recruitment droped to extremely low values. Length indices increased. The spatial distributions in the egg and the adult fish were more coastal and anisotropic. The Cusum analysis and the trends method provided similar results as they both identified the important changes in the last years of the series. The Cusum diagnostic table allowed to trigger alarm signals in 2004 and 2005. ICES recommended closure of the fishery in 2005 only. Early warnings seemed possible with the present indicator-based monitoring.

Red mullet in the Thyrrhenian Sea (Fig. 10). As for eastern mediterranean hake, the survey time series was short and variability in the indices high. The derivatives trend method identified significant recent (5 last years) changes in some indices while the Cusum did not. The reference period was defined as the second half of the series because of higher and more stable abundance levels. The choice of the reference period may have generated mismatch between the search for recent trends and the identification of departures from a reference with Cusum. Age group 3 disappeared from the experimental catches of the last 2 years (2002-2003) in the sub-unit 10a and a decrease of indices of length structure and recruit abundance, referred to the whole area, was occurring in the last two years. Highest values for Z were reached in the last years (2001-02) of the series whereas the survey index, Lbar and L75 were all, although not significantly, decreasing in 2003. A decline in length indicators was also observed from 1995 to 1998, but in that period the total mortality was lower and, in addition, in 1998 the survey and recruit indices increased, remaining almost stable since then. Spatial indices, especially those regarding location in subunit 10a, displayed a long-term trend and a tendency to change in recent years. But this was not identified using the Cusum analysis which considered that variability was such that no out-of-control value was reached. Di-cusum analysis allowed the triggering of alert signal for the survey index in 1997, when it reached the lowest level. An alert signal was also obtained for the multivariate spatial indices in the years 1995-1996 in the subunit 10b, probably as result of change in location and occupation indices. Older ages were more dispersed westwards and slightly offshore. The retained diagnostic was the following: recent increase in Z, low abundance reached in 1997 and change in spatial distribution in 10b in 1995-96.

4. Discussion

Research survey series have been systematically undertaken since 20 years in the most favourable cases and since 10 years otherwise. These series captured stock fluctuations at a time when stocks were already in a degraded situation. Short series with high variability in the indices resulted in statistical difficulty in detecting change. In the worst case, the derivatives method would detect recent trends and the Cusum no signal.

Assigning one particular cause to a combination of trends using the cause-effects table was not always easy as many causes may occur jointly thus providing signals that are difficult to interpret. Surprisingly, variations in length indices have not always been straightforward to interpret and in some cases have been conflicting with the variation in other indices. The result was that diagnostics have always relied more on abundance and spatial indices and in some cases only on length indices.

Refinements of the procedures in the application of the methods is indeed to welcome. For the Cusum analysis these are anticipated to be the definition of the reference period and the accepted risks of false and no alarm with which to trigger an out-of-control signal. For the trends method, the scale at which to identify short-term trends along the series has been a difficulty in those case studies where the change in slope was not at the end of the series. The interest in the Cusum procedure has the potential advantage of suggesting reference values for the indicators.

In all, the methods have shown potential across the case studies to monitor population status using fishery-independent survey-based indices of population biological and spatial attributes. The system was intended to be a monitoring system of the state of fish stocks. As such, it is hoped that it complements the traditional assessment, providing comprehensive biological and spatial information on the evolution of the stocks. Procedures can now be applied in operational mode to provide results to assessment working groups for any stock that is monitored with research surveys.

Indicator-based diagnostics, because they are based on spatial indices as well as abundance and length indices can justify alternative management strategies to TAC such as the protection of juveniles or closed areas.

References

Cotter, J., Mesnil, B., Trenkel, V., Rochet, M.-J., Petitgas, P., Woillez, M., Rivoirard, J., Uriarte, A., Witthames, P., Spedicato, M.-T. and Lembo, G. 2007. Manual of indicators and methods for assessing fish stocks using only fishery independent survey based information. ICES CM 2007/O:27.

Kell, L., Mosqueira, I., Grosjean, P., Fromentin, J.M. et al. 2007. FLR : an open source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64 (accepted january 2007).

Trenkel, V., Rochet, M.-J. and Mesnil, B. 2007. From model-based prescriptive advice to indicator-based interactive advice. ICES Journal of Marine Science 64, (accepted January 2007).

Montgomery, D. 2005. Introduction to statistical process quality control. Wiley, new York. 5th Edition.

Woillez, M., Poulard, J.-C., Rivoirard, J., Petitgas, P. and Bez N. 2007a. Indices for capturing spatial patterns and their evolution in time with an application on European hake (*Merluccius merluccius*) in the Bay of Biscay. ICES Journal of Marine Science, 64: 537-550.

Woillez, M., Rivoirard, J., Petitgas, P. and Deerenberg, C. 2007b. Selecting and combining survey based indices of fish stocks using their correlation in time to make diagnostics of their status. ICES CM 2007/O:07.

Table 1: Fish stocks case studies of the Fisboat project on which the indicator-based monitoring methodology was applied.

Stock	Behaviour	Life span	Survey Type	Survey time series used	Reference period used	Age range in survey data
Barents Sea Cod	Demersal	Long	Botttom trawl	1989-2004	1996-2004	1-10
North Sea Cod	Demersal	Long	Bottom trawl	1985-2005	1985-1994	1-6
Baltic Sea Cod	Demersal	Long	Bottom trawl	1994-2004	1994-1999 [excluding 97]	1-5
Biscay Hake	Demersal	Long	Bottom trawl	1987-2004 [excluding: 91,93,96]	1987-1997	0-5
Ionian Sea Hake	Demersal	Long	Bottom trawl	1994-2003 [exclusing: 02]	1998-2001	0-5
Aegean Sea Hake	Demersal	Long	Bottom trawl	1994-2003 [exclusing: 02]	1998-2001	0-5
North Sea Herring	Pelagic	Long	Acoustics	1989-2002	1989-1993	0-9
Biscay Anchovy	Pelagic	Short	Acoustics Eggs	1989-2005 [excluding: 91-93,95,96,99] 1989-2005 [excluding: 93,96,99-00]	1990-2001 1990-2001	1-3 -
Thyrrhenian Sea Red mullet (GS10)	Demersal	Short	Bottom trawl	1994-2003	1999-2003	1-3

Table 2 : Raw biological (non spatial) indices used in the study. All Fisboat biological indices are fully described in Cotter et al. (2007).

Population attribute	Index name	Index symbol	Index description
Total abundance	Abundance	Ln-Ntot	Natural logarithm (total surveyed fish numbers all ages pooled +1)
Recruit abundance	Recruit abundance	Ln-Rec	Natural logarithm (fish numbers at recruiting age +1)
Length structure	Mean length	Lbar	Mean length of the fish length histogram
Length structure	First quartile of length	L25	25th percentile of the fish length histogram
Length structure	Last quartile of length	L75	75th percentile of the fish length histogram
Reproductive capability	Length at 50% maturity	L50matu	Length at which 50% of the individuals have reached reproductive maturity
Total mortality	Mortality Z	Z	Mortality rate between years $t-1$ and t of all individuals aged $a_{\min}$ to $a_{\max-1}$

Table 3 : Raw spatial indices used in the study. All Fisboat spatial indices are fully described in Cotter et al. (2007) and in Woillez et al. (2007).

Population attribute	Index name	Index symbol	Index description
Location	Longitude gravity center	Xcg	Weighted average of sample longitudinal positions
Location	Latitude gravity center	Ycg	Weighted average of sample latitudinal positions
Patchiness	Number of Patches	NbPatch	Concentration of abundance in patches with spatially distant local gravity centers
Dispersion	Inertia	I	Weighted variance of sample positions around a gravity centre
Dispersion	Anisotropy	A	Ratio of inertia for directions carrying minimal and maximal inertia
Occupation	PositiveArea	PA	Area of non null values
Correlation	Microstructure	MI	Decrease of correlation at short distance on the relative covariogram
Correlation	EquivalentArea	EA	Integral range of the relative covariogram
Aggregation	SpreadingArea	SA	Concentration of abundance relative to the homogeneous distribution

Table 4 : Cause-effects table linking one cause (first column) to a combination of expected trends in biological indicators. (after Trenkel et al., 2007). 0: no trend; -1: decreasing trend; 1: increasing trend.

Cause	Z	ln-Ntot	Lbar	L25	L75	ln-Rec
F: increase	1	-1	-1	0	-1	0
F: decrease	-1	1	1	0	1	0
Recruit: increase	0	1	-1	-1	0	1
Recruit: decrease	0	-1	1	1	0	-1
Faster growth	0	0	1	0	1	0
Slower growth	0	0	-1	0	-1	0
Larger fish caught (or change in fishing area, stock distribution or gear)	-1	1	1	0	1	0
Smaller fish caught (or change in fishing area, stock distribution or gear)	1	-1	-1	-1	0	0

Table 5: Indices calculated in each case study

	Cod			Hake			Herring	Anchovy		Red mullet	
	Barents sea	Baltic sea	North Sea	Bay of Biscay	Ionian sea	Aegean sea	North Sea	Bay of Biscay	Bay of Biscay	Thyrhenian sea GS10a	Thyrhenian sea GS10b
Survey	BT	BT	BT	BT	BT	BT	AC	AC	EG	BT	BT
Age groups	1-10	1-5	1-6	0-5	0-5	0-5	1-9	1-3	-	1-2	1-3
Biological Indicators											
Ln-Ntot	x	x	x	x	x	x	x	x	x	x	
Ln-Rec	x	x	x	x	x	x	x	x	x	x	
Lbar	x	x	x	x	x	x	x	x		x	
L25	x	x	x	x	x	x	x	x		x	
L75	x	x	x	x	x	x	x	x		x	
L50matu	x	x	x	x	x	x	x	x		x	
Z	x	x	x	x	x	x	x	x	x	x	
PCA-based	x	x	x	x	x	x	x	x		x	
Spatial indicators by age											
PositiveArea	x	x	x	x	x	x	x	x	x	x	x
Inertia	x	x	x	x	x	x	x	x	x	x	x
Anisotropy	x	x	x	x	x	x	x	x	x	x	x
Xcg	x	x	x	x	x	x	x	x	x	x	x
Ycg	x	x	x	x	x	x	x	x	x	x	x
NbPatches	x	x	x	x	x	x	x	x	x	x	x
Microstructure	x	x	x	x	x	x	x	x	x	x	x
EquivalentArea	x	x	x	x	x	x	x	x	x	x	x
SpreadingArea	x	x	x	x	x	x	x	x	x	x	x
MFA-based	x	x	x	x	x	x	x	x	x	x	x
MAF-based									x	x	x

Table 7: Loadings of the biological indices on their Principal Components for each case study. Asbolute values greater than 0.6 are in bold characters.

Case study	Index	Comp1	Comp2	Comp3
Cod	Ln.Ntot	0.376	**-0.761**	-0.248
Barents Sea	Ln.Nrec	**-0.636**	0.050	**-0.616**
	Lbar	**0.912**	0.205	-0.020
	L25	**0.747**	0.354	-0.065
	L75	**0.878**	0.187	-0.015
	L50.matu	-0.023	**-0.757**	0.441
	Z	-0.567	0.557	0.369
Cod	Ln.Ntot	0.241	**-0.834**	-0.197
North Sea	Ln.Nrec	0.558	-0.527	0.550
	Lbar	**0.831**	0.429	0.129
	L25	**0.845**	-0.157	0.365
	L75	0.594	**0.671**	-0.144
	L50.matu	**0.813**	-0.039	-0.406
	Z	-0.345	0.277	**0.751**
Cod	Ln.Ntot	0.104	**-0.842**	-0.183
Baltic Sea	Ln.Nrec	**-0.690**	-0.565	0.036
	Lbar	**-0.823**	0.227	-0.263
	L25	**-0.853**	0.172	-0.165
	L75	**-0.841**	0.123	-0.257
	L50.matu	**-0.620**	-0.527	0.322
	Z	**0.639**	-0.301	-0.516
Herring	Ln.Ntot	**-0.66**	**0.60**	0.00
North Sea	Ln.Nrec	-0.48	0.41	-0.62
	Lbar	**1.00**	0.54	-0.03
	L25	0.39	**0.77**	0.04
	L75	**1.00**	-0.14	-0.11
	L50.matu			
	Z	0.27	-0.38	**-0.60**
Anchovy	Ln.Ntot	0.177	**0.908**	-0.014
Biscay	Ln.Nrec	-0.172	**0.907**	-0.070
	Lbar	**0.874**	-0.179	-0.195
	L25	**0.841**	-0.058	-0.352
	L75	**0.815**	-0.126	0.383
	L50.matu			
	Z	**-0.795**	-0.381	-0.182

Case study	Index	Comp1	Comp2	Comp3
Hake	Ln.Ntot	**0.832**	0.353	0.180
Biscay	Ln.Nrec	**0.876**	0.267	0.115
	Lbar	**-0.749**	0.520	0.152
	L25	-0.452	**0.762**	-0.259
	L75	**-0.869**	0.145	0.273
	L50.matu			
	Z	-0.590	-0.564	0.027
Hake	Ln.Ntot	**0.851**	0.157	-0.040
Ionnian Sea	Ln.Nrec	0.568	**-0.641**	-0.143
	Lbar	**-0.844**	-0.197	-0.007
	L25	**-0.771**	-0.366	0.147
	L75	**-0.853**	0.139	-0.061
	L50.matu			
	Z	**-0.833**	0.120	-0.205
Hake	Ln.Ntot	**0.759**	0.387	0.160
Aegean Sea	Ln.Nrec	**0.857**	0.118	0.025
	Lbar	**-0.828**	0.260	-0.018
	L25	-0.410	**-0.652**	0.409
	L75	**-0.739**	0.448	-0.072
	L50.matu			
	Z	0.146	**-0.800**	-0.298
Red Mullet	Ln.Ntot	0.356	**0.799**	0.184
Thyrhenian	Ln.Nrec	**0.888**	-0.082	0.050
Sea	Lbar	**0.890**	0.053	-0.020
	L25	**0.823**	-0.266	-0.208
	L75	**0.869**	0.150	-0.138
	L50.matu	**0.878**	0.038	-0.165
	Z	**-0.642**	0.317	-0.537

Table 8a: Interpretation of the principal components (PCs) resulting from applying MFA on the spatial indicators at age. The table shows the number of times that each index has shown a correlation greater than +0.5 or lower than –0.5 whith the PCs along the data series. Values in bold character signal a number of times greater than half the number of years.

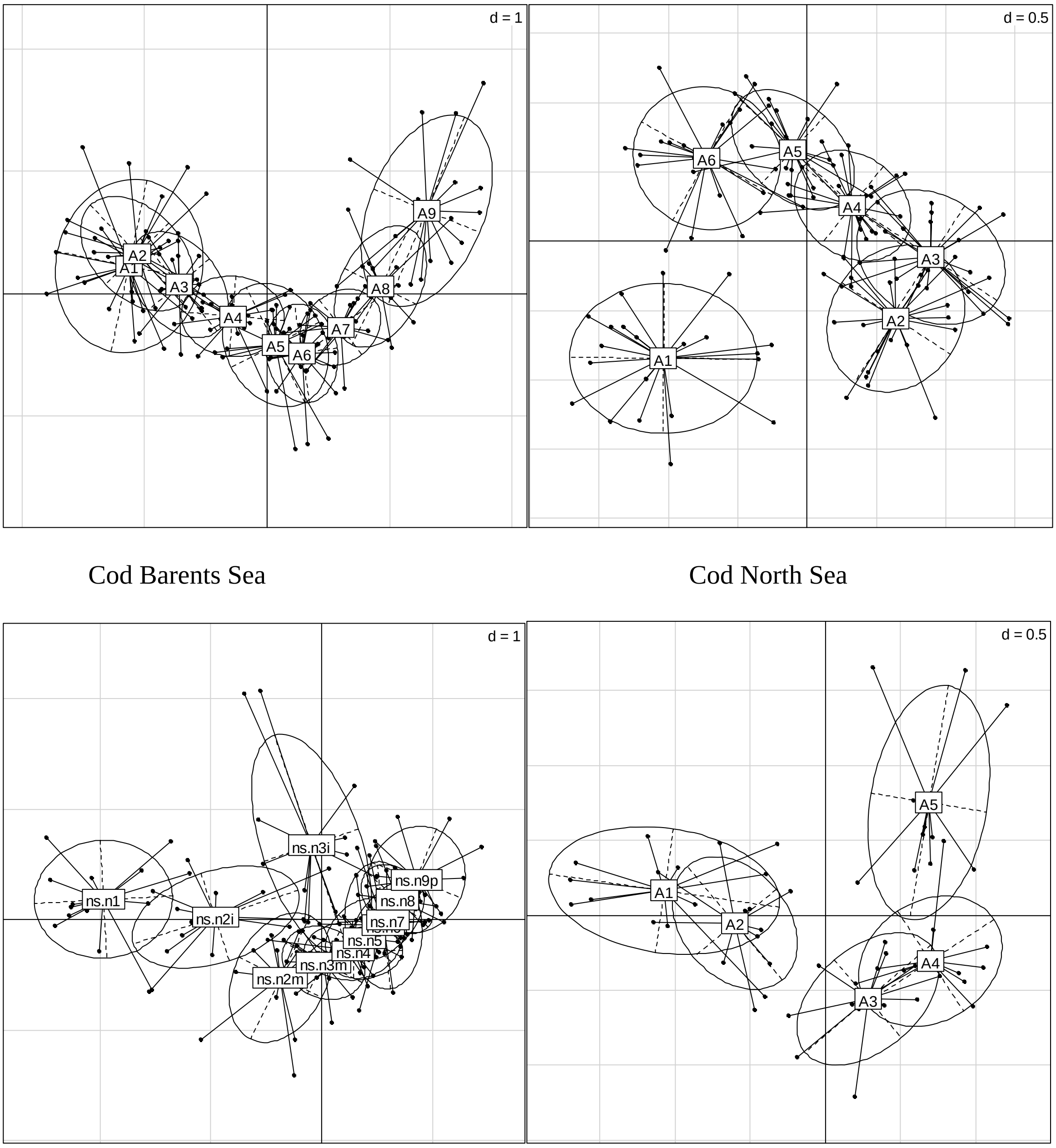
Case study	Index	Comp1	Comp2	Comp3	Case study	Index	Comp1	Comp2	Comp3
Cod	PositiveArea	**0+|16-**	0+|3-	0+|0-	Hake	PositiveArea	**0+|13-**	1+|0-	0+|0-
Barents Sea	Inertia	**1+|12-**	1+|0-	2+|0-	Biscay	Inertia	**9+|0-**	0+|6-	0+|1-
	Anisotropy	**1+|9-**	1+|0-	7+|0-		Anisotropy	7+|0-	**10+|0-**	1+|1-
	xcg	**0+|16-**	0+|0-	1+|0-		xcg	**1+|9-**	2+|3-	1+|1-
	ycg	1+|7-	0+|5-	0+|1-		ycg	7+|2-	4+|1-	2+|0-
	MicrostructureIndex	5+|2-	4+|0-	4+|1-		MicrostructureIndex	4+|0-	2+|4-	2+|2-
	EquivalentArea	1+|3-	0+|6-	2+|7-		EquivalentArea	**0+|12-**	0+|1-	2+|0-
	SpreadingArea	1+|5-	**0+|12-**	1+|3-		SpreadingArea	**0+|11-**	0+|7-	2+|0-
Cod	PositiveArea	**12+|0-**	**0+|16-**	0+|0-	Hake	PositiveArea	**0+|7-**	0+|1-	1+|1-
North Sea	Inertia	**12+|2-**	1+|2-	2+|5-	Ionnian Sea	Inertia	**4+|2-**	2+|0-	2+|1-
	Anisotropy	0+|2-	**15+|0-**	3+|0-		Anisotropy	2+|1-	3+|0-	1+|3-
	xcg	**0+|17-**	4+|4-	2+|0-		xcg	**7+|0-**	2+|0-	1+|0-
	ycg	**13+|0-**	**15+|0-**	0+|0-		ycg	**1+|6-**	2+|2-	0+|1-
	MicrostructureIndex	**1+|11-**	3+|3-	1+|5-		MicrostructureIndex	**5+|0-**	3+|0-	1+|1-
	EquivalentArea	**12+|0-**	1+|2-	2+|4-		EquivalentArea	**2+|5-**	0+|3-	1+|1-
	SpreadingArea	**18+|0-**	0+|7-	1+|0-		SpreadingArea	**1+|5-**	0+|3-	2+|0-
Cod	PositiveArea	3+|0-	**0+|7-**	0+|1-	Hake	PositiveArea	3+|0-	**9+|0-**	0+|0-
Baltic Sea	Inertia	**1+|7-**	0+|3-	2+|2-	Aegean Sea	Inertia	**5+|0-**	**0+|5-**	1+|1-
	Anisotropy	**1+|6-**	3+|1-	5+|0-		Anisotropy	0+|2-	1+|4-	1+|1-
	xcg	2+|3-	3+|0-	1+|2-		xcg	**7+|0-**	0+|4-	0+|0-
	ycg	**9+|0-**	1+|0-	0+|3-		ycg	**2+|6-**	1+|2-	0+|0-
	MicrostructureIndex	1+|3-	1+|1-	2+|1-		MicrostructureIndex	**0+|5-**	1+|1-	0+|1-
	EquivalentArea	**6+|0-**	2+|5-	2+|0-		EquivalentArea	**7+|0-**	0+|1-	3+|0-
	SpreadingArea	**8+|0-**	0+|2-	2+|1-		SpreadingArea	**8+|0-**	0+|0-	2+|0-
Herring	PositiveArea	**0+|12-**	**0+|7-**	0+|0-					
North Sea	Inertia	**0+|10-**	2+|0-	0+|0-					
	Anisotropy	0+|3-	3+|0-	3+|0-					
	xcg	**0+|14-**	0+|0-	0+|1-					
	ycg	**14+|0-**	0+|1-	0+|0-					
	MicrostructureIndex	2+|2-	2+|1-	4+|0-					
	EquivalentArea	**0+|8-**	0+|5-	0+|4-					
	SpreadingArea	**0+|8-**	0+|5-	1+|0-					

Table 8b: Interpretation of the principal components (PCs) resulting from applying PCA to the spatial indicators at age. PCA was applied instead of MFA when the stock has too few age classes. Values are the loadings of the indices on the PCs. Values in bold character signal a correlation greater than 0.6 in absolute value.

Case study	Index	Comp1	Comp2	Comp3	Case study	Index	Comp1	Comp2	Comp3
Anchovy	PositiveArea	**-0.799**	0.244	-0.291	Red Mullet	PositiveArea	0.38	**0.867**	-0.088
Biscay	Inertia	-0.185	**-0.807**	-0.291	GS10a	Inertia	**0.62**	**0.654**	-0.255
AC	Anisotropy	0.276	**-0.722**	0.416		Anisotropy			
	xcg	**0.595**	-0.032	0.135		xcg	**-0.795**	0.382	0.457
	ycg	-0.56	-0.504	-0.382		ycg	**0.816**	-0.355	-0.439
	MicrostructureIndex	0.593	0.328	-0.383		MicrostructureIndex	**-0.84**	0.312	-0.393
	EquivalentArea	**-0.745**	0.119	0.582		EquivalentArea	**0.932**	-0.02	0.338
	SpreadingArea	**-0.952**	0.13	0.036		SpreadingArea	**0.937**	0.148	0.285
Anchovy	PositiveArea	**0.914**	-0.225	0.131	Red Mullet	PositiveArea	0.586	-0.48	0.585
Biscay	Inertia	**0.711**	**0.649**	0.194	GS10b	Inertia	0.374	**-0.657**	-0.543
EG	Anisotropy	-0.598	-0.17	**0.725**		Anisotropy			
	xcg	**-0.617**	**-0.619**	0.183		xcg	**0.892**	0.378	0.087
	ycg	**0.871**	0.128	0.238		ycg	**0.762**	0.506	0.104
	MicrostructureIndex	-0.469	**0.743**	0.196		MicrostructureIndex	**-0.642**	-0.534	0.346
	EquivalentArea	**0.903**	-0.354	-0.008		EquivalentArea	**0.897**	-0.281	-0.083
	SpreadingArea	**0.924**	-0.197	0.196		SpreadingArea	**0.909**	-0.315	0

Table 9: Analysis methods applied to detect changes in the time series of raw and multivariate indices by case study

	Cod			Hake			Herring	Anchovy		Red mullet	
	Barents sea	Baltic sea	North Sea	Bay of Biscay	Ionnian sea	Aegean sea	North Sea	Bay of Biscay	Bay of Biscay	Thyrhenian sea GS10a	Thyrhenian sea GS10b
Survey type	BT	BT	BT	BT	BT	BT	AC	AC	EG	BT	BT
Biological Indices : raw											
Trend	x	x	x	x	x	x		x	x	x	
Di-Cusum		x	x	x	x	x			x	x	
Biological Indices : multivariate PCA-based											
Trend			x		x	x					
Di-Cusum	x	x	x	x	x		x	x		x	
Spatial Indices : raw											
Trend		x	x	x	x	x			x	x	x
Di-Cusum		x		x					x		
Spatial Indices : multivariate MFA-based											
Trend			x		x	x			x		
Di-Cusum	x	x	x	x	x	x	x	x	x	x	x
Selection of raw indices											
MAF selection		x		x			x	x			



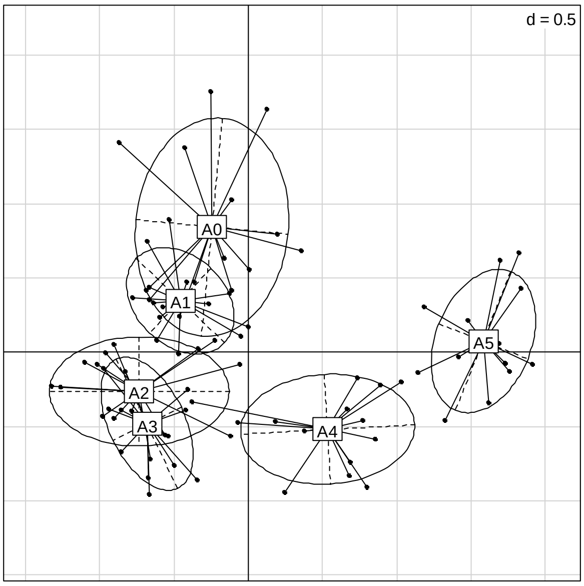
Cod Barents Sea

Cod North Sea

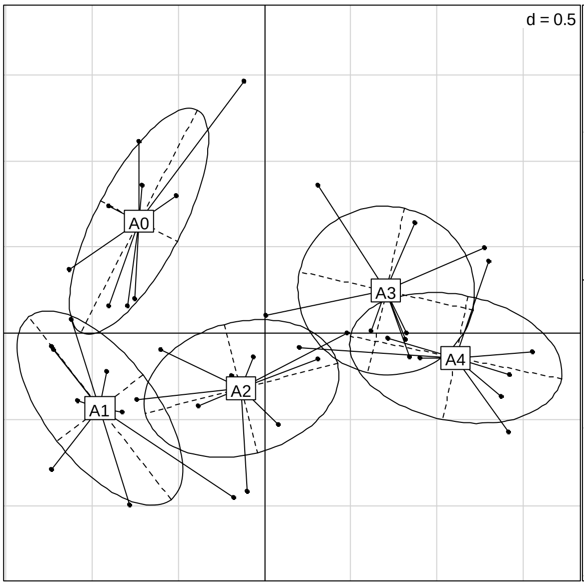
Herring North Sea

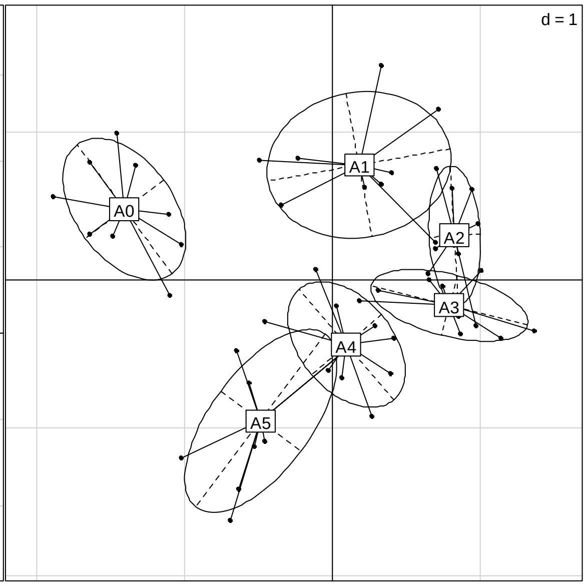
Cod Baltic Sea

Figure 1a: Representation of the life cycle spatial pattern and its inter-annual variations in the first factorial plane of the MFA applied on the spatial indicators at age. Each point represents the position of each age in each year. The gravity center of each age is labelled. Representations for cod and herring.



Hake Biscay

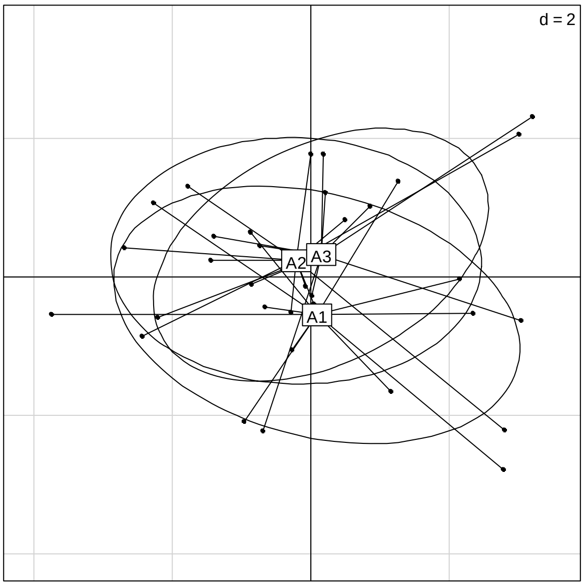




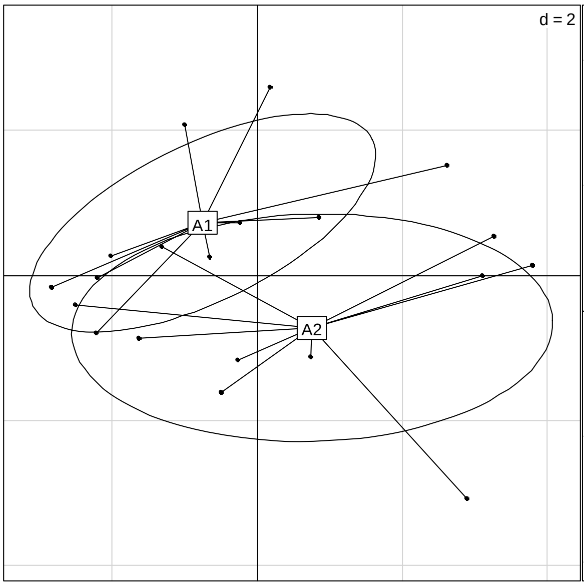
Hake Ionnian Sea

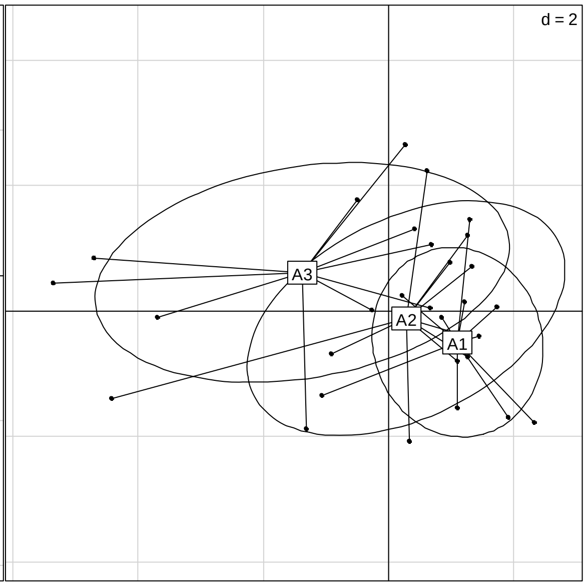
Hake Aegean Sea

Figure 1b: Representation of the life cycle spatial pattern and its inter-annual variations in the first factorial plane of the MFA applied on the spatial indicators at age. Each point represents the position of each age in each year. The gravity center of each age is labelled. Representations for hake.



anchovy Biscay AC





Red Mullet GS10a

Red Mullet GS10b

Figure 1c: Representation of the life cycle spatial pattern and its inter-annual variations in the first factorial plane of the PCA applied on the spatial indicators at age. Each point represents the position of each age in each year. The gravity center of each age is labelled. Representations for Anchovy and Red Mullet.

BS COD		CUSUM diagnostics table							
Years	**MFA_spatial**	**PCA_biological**	**Ln_Ntot**	**Ln_Rec**	**Lbar**	**L25**	**L75**	**Z**	**diagnostic**
1989	0	**3.4**	0	**-0.8**	**1.7**	**4.0**	0		
1990	**1.5**	**3.7**	**-1.6**	**-2.7**	**1.8**	0	**1.5**	0	alarm
1991	**1.1**	**3.4**	**-4.6**	**-4.0**	0	0	0	0	alarm
1992	**2.7**	**3.9**	**-6.9**	**-2.8**	0	0	0	0	alarm
1993	0	**1.0**	**-4.9**	0	0	0	0	**-1.1**	
1994	0	0	0	0	0	0	0	**-1.0**	
1995	0	0	0	0	0	0	0	0	
1996	0	0	0	0	0	0	0	0	ref
1997	0	0	0	0	0	0	**-1.1**	0	ref
1998	0	0	0	0	0	0	0	0	ref
1999	1	0	0	0	0	0	0	0	ref
2000	0	0	0	0	0	0	0	0	ref
2001	0	0	0	0	0	0	0	0	ref
2002	0	0	0	0	0	0	0	0	ref
2003	0	0	0	0	0	0	0	0	ref
2004	0	0	0	0	0	0	0	0	ref

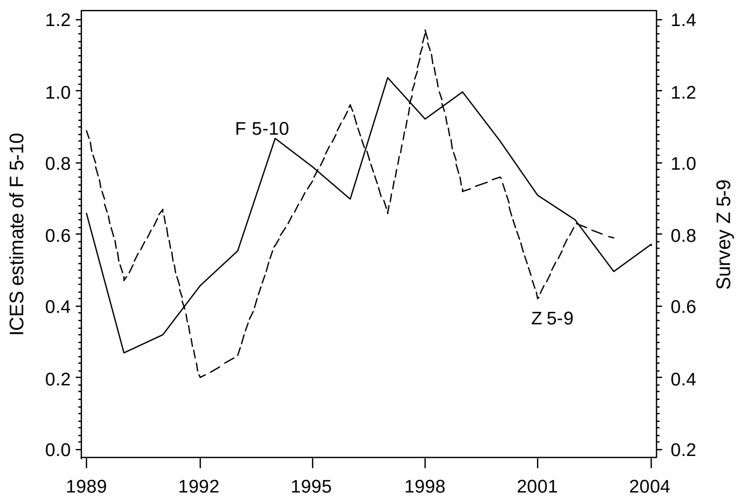


Fig. 2: Barents sea cod. Cusum diagnostic table for multivariate indices and raw biological indices (above). Comparison of survey Z estimate with ICES estimate (below), showing the low in the beginning of the 90s.

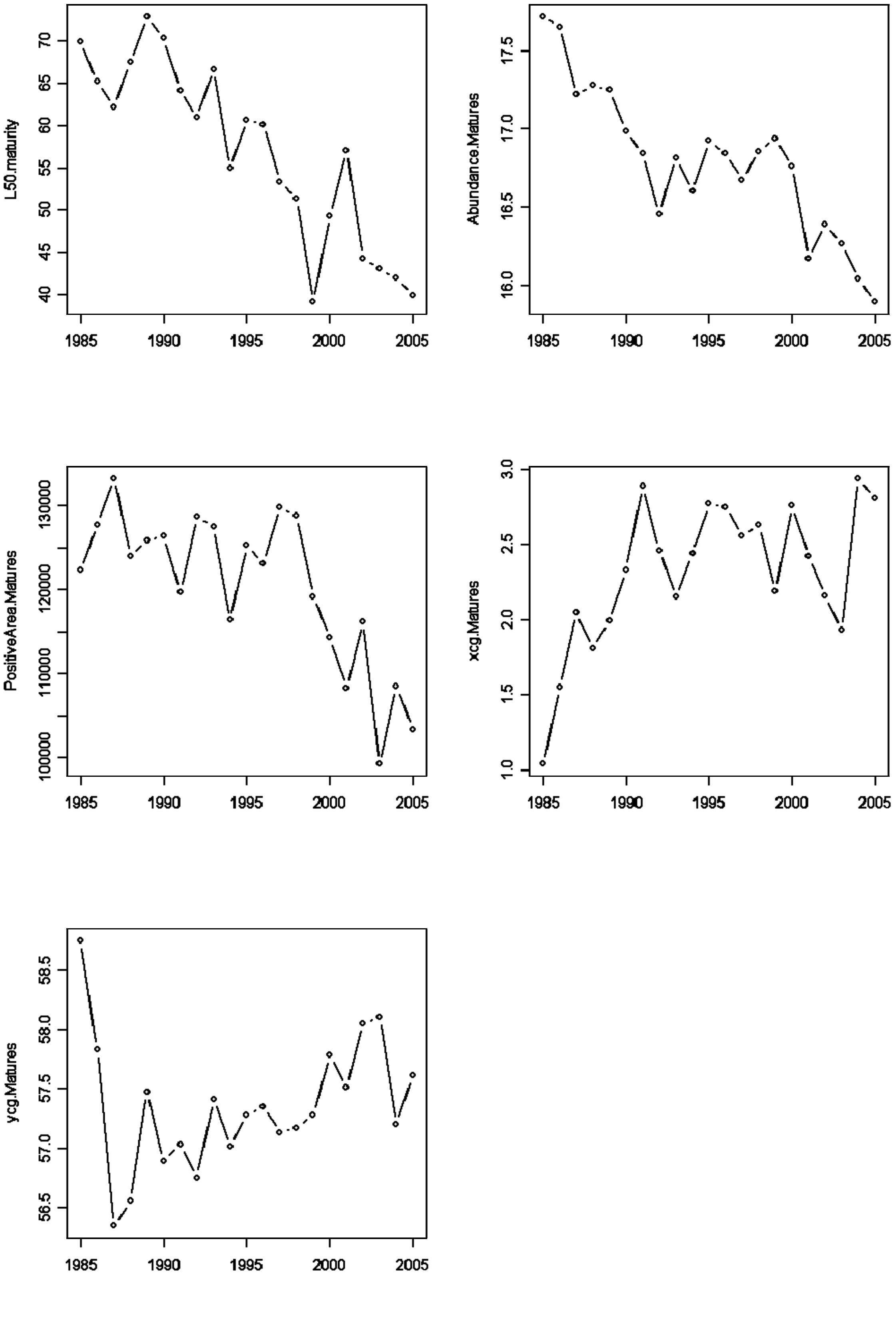


Fig.3a : Example of MAF selected raw indices that express the trend variation in biological and spatial indices

Trend method result table: 1/-1 indicates linear (+/-) trend, 1*/-1* indicates recent change only. recent is from 2001 to 2005 (5 last years).

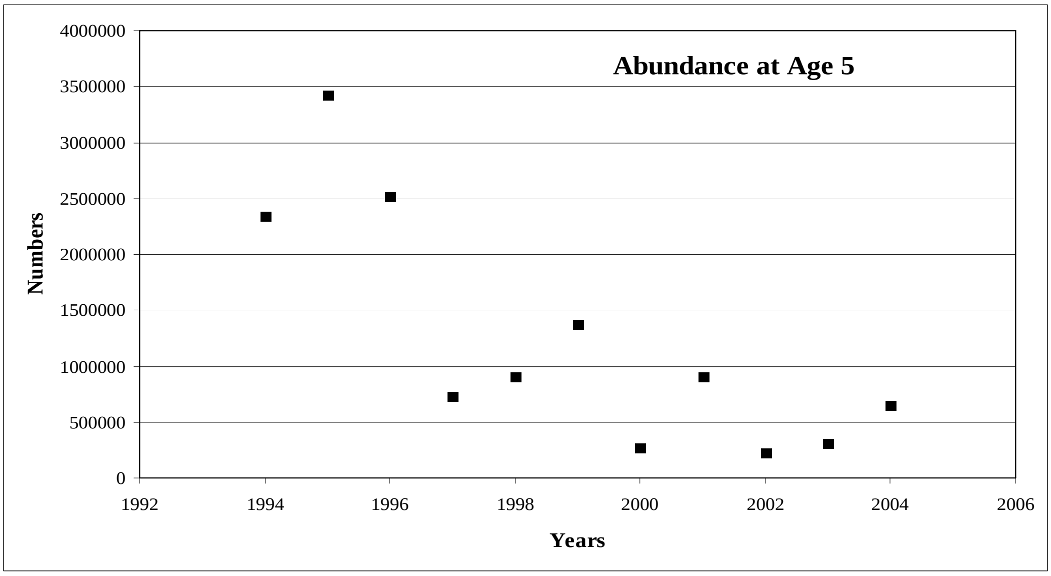
Non-spatial indices	**all**	**recent**
Ln_Survey.index	-1	-1
Ln_Abundance (recruits)	-1	-1
L25	0	1*
Lbar	0	1*
L75	0	1*
L50.maturity	-1	-1
Z	0	-1*
md	1	0

	Age 1		**Age 2**		**Age 3**		**Age 4**		**Age 5**		**Age 6**	
Spatial indices	**all**	**recent**	**all**	**Recent**	**all**	**recent**	**all**	**recent**	**all**	**recent**	**all**	**recent**
xcg	0	0	0	0	1	0	1	0	0	0	0	-1
ycg	1	0	0	0	0	0	0	0	1	1	0	1
Inertia	0	1*	-1	0	0	0	0	0	-1	0	0	0
Anisotropy	1	0	0	-1	-1	0	0	0	0	0	0	1
Positive area	0	-1*	-1	-1*	0	-1*	-1	-1*	-1	-1	-1	0
Equivalent area	0	-1*	0	-1*	1	1	0	0	0	-1	0	0
Spreading area	0	-1*	0	0	0	-1*	0	0	-1	-1	-1	0
Microstructure	0	0	-1	0	0	0	0	0	0	0	0	0
No. of patches	0	1*	0	0	0	0	0	0	0	0	0	-1
dmul (all ages)	1	1										

cod NS **Cusum diagnostic table**

Year	**MFA_spatial**	**PCA_biological**	**Ln_Ntot**	**Ln_Rec**	**Lbar**	**L25**	**L75**	**L50.matu**	**Z**	**diaqnostic**
1985	0	0	0	0	0	0	0	0		ref
1986	0	0	0	0	0	0	0	0	0	ref
1987	0	0	0	0	0	0	0	0	0	ref
1988	0	0	0	0	0	0	0	0	0	ref
1989	0	0	0	0	0	0	0	0	0	ref
1990	0	0	0	0	0	0	0	0	0	ref
1991	0	0	0	0	0	0	0	0	0	ref
1992	0	0	0	0	0	0	0	0	0	ref
1993	0	0	0	0	0	0	0	0	0	ref
1994	0	0	0	0	0	0	0	0	0	ref
1995	**1.4**	0	0	0	0	0	0	0	0	
1996	0	0	0	0	0	0	0	0	0	
1997	0	**3.4**	0	0	**-1.8**	**-1.0**	**-2.2**	**-1.8**	0	
1998	0	0	0	0	0	0	**-2.0**	**-3.4**	0	
1999	0	**1.3**	**-1.3**	**-1.3**	0	0	0	**-7.4**	0	
2000	**1.8**	0	**-1.7**	0	0	0	0	**-9.3**	0	alarm
2001	**2.7**	0	**-3.0**	0	0	0	0	**-9.8**	0	alarm
2002	**2.0**	0	**-4.0**	0	0	0	0	**-12.8**	0	alarm
2003	**1.7**	**2.5**	**-7.5**	0	0	0	0	**-16.0**	0	alarm
2004	**2.4**	**5.3**	**-10.5**	**-1.2**	0	0	0	**-19.3**	0	alarm
2005	**2.9**	**9.3**	**-15.0**	**-2.0**	0	0	**1.2**	**-23.1**	0	alarm

Fig. 3b: Trend result table (above) and Cusum diagnostic table (below) for North Sea cod.



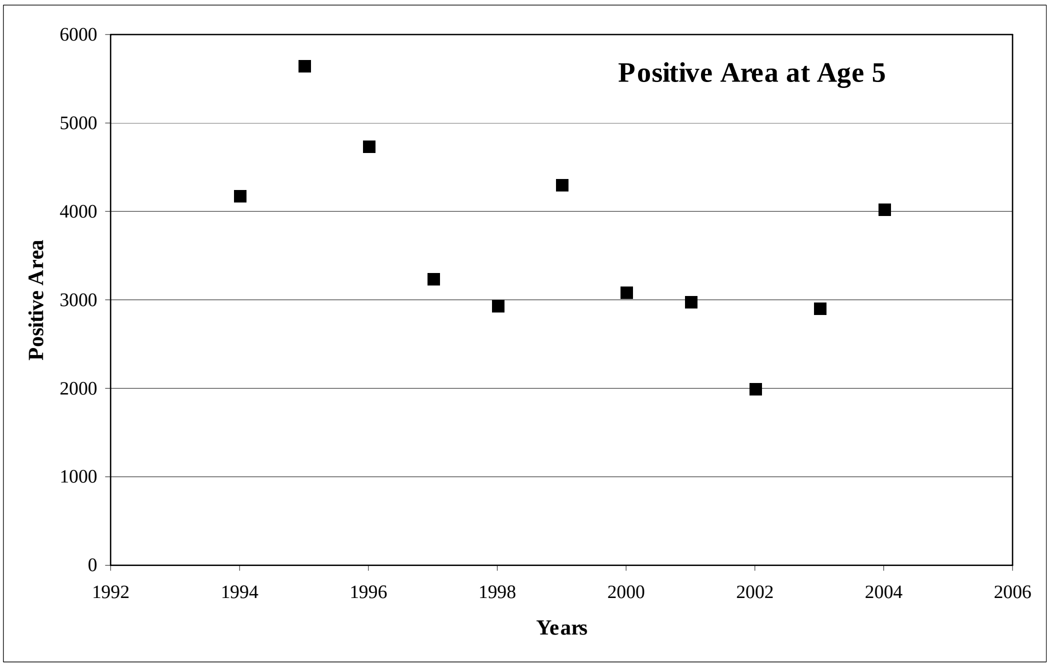


Fig.4a: Time series of the indices that convey the major signal in the evolution of Baltic Sea cod. Indices are Abundance at age 5 and Positive area at age 5.

Results of trend analysis

	all period	recent
Z	1	-1
Ln_Abdnce	0	1
Lbar	-1	0
L25	0	0
L75	-1	0
Ln_Recruit	1	0

cod BA **Cusum diagnostic table**

Year	Ln.Nb.A5	PositiveArea.A5	Ln_Ntot	Ln_Rec	Lbar	L25	L75	Z	diagnostic
1994	0	0	0	0	0	0	0	0	ref
1995	0	0	0	0	0	0	0	0	ref
1996	0	0	0	0	0	0	0	0	ref
1997	0	0	**-6.19**	**-2.2**	0	0	**3.9**	0	
1998	0	0	**-4.44**	0	**2.3**	0	**4**	0	ref
1999	0	0	0	0	0	0	**3**	0	ref
2000	**-2.24**	0	0	0	0	0	0	0	
2001	**-2.23**	0	0	0	0	0	**1.2**	0	alarm
2002	**-4.88**	**-3.4**	0	0	0	0	0	0	alarm
2003	**-6.86**	**-4.1**	0	0	0	0	0	0	alarm
2004	**-7.47**	**-3.4**	0	0	0	0	**-1.4**	0	alarm

Fig. 4b: trend results table (above) and Cusum diagnostic table (below) for Baltic Sea cod.

Nonparametric derivatives method for determining recent trends in indicator time series. For diagnostic recent (7 last years) trends: 1=increase, –1=decrease and 0=no change.

Indicator	LinearSlope	PvalueAll	LinSlopeLastYears	PvalueLast	7 last years diagnostic	
					Linear	Non Linear
L25	0.11	0.06	0.35	0.41	0	**1**
Lbar	0.02	0.83	0.50	0.49	0	**1**
L75	-0.03	0.86	1.09	0.38	0	**1**
ln_recruit_index	0.02	0.56	0.08	0.78	0	0
ln_survey_index_a1a5	-0.01	0.59	0.16	0.21	0	0
Z	0.04	0.22	0.30	0.25	0	**1**
Anisotropy.A0	-0.01	0.75	-0.12	0.42	0	0
Anisotropy.A1	0.03	0.43	-0.15	0.44	0	0
Anisotropy.A2	-0.01	0.65	-0.11	0.15	0	**-1**
Anisotropy.A3	0.03	0.07	-0.06	0.53	0	0
Anisotropy.A4	0.08	**0.03**	-0.01	0.98	0	**1**
Anisotropy.A5	0.05	0.35	-0.33	0.09	0	**-1**
EquivalentArea.A0	3.40	0.97	-432.71	0.37	0	**-1**
EquivalentArea.A1	-219.04	**0.04**	-377.25	0.34	0	0
EquivalentArea.A2	-163.07	0.19	139.89	0.83	0	0
EquivalentArea.A3	33.86	0.74	-7.25	0.99	0	0
EquivalentArea.A4	-280.06	**0.01**	-455.14	0.15	0	0
EquivalentArea.A5	-337.58	**0.00**	119.75	0.55	0	0
Inertia.A0	-72.08	0.36	-317.71	0.45	0	**-1**
Inertia.A1	-101.26	0.10	-243.79	0.35	0	0
Inertia.A2	-61.37	0.49	-189.64	0.66	0	0
Inertia.A3	219.90	**0.03**	116.00	0.83	0	0
Inertia.A4	495.50	**0.01**	561.93	0.56	0	**1**
Inertia.A5	198.69	0.46	739.79	0.58	0	**1**
MicrostructureIndex.A0	-0.01	0.14	-0.01	0.62	0	**-1**
MicrostructureIndex.A1	0.00	0.67	0.01	0.65	0	0
MicrostructureIndex.A2	0.00	0.30	-0.01	0.76	0	0
MicrostructureIndex.A3	0.00	0.96	0.02	0.13	0	**1**
MicrostructureIndex.A4	0.01	0.43	0.01	0.62	0	**1**
MicrostructureIndex.A5	0.01	0.20	-0.01	0.69	0	**-1**
PositiveArea.A0	-77.87	0.60	759.89	0.36	0	**1**
PositiveArea.A1	-59.05	0.63	1117.89	0.04	**1**	**1**
PositiveArea.A2	-190.58	0.23	903.25	0.13	0	0
PositiveArea.A3	-128.10	0.53	274.21	0.65	0	**-1**
PositiveArea.A4	-351.11	0.06	-19.79	0.96	0	0
PositiveArea.A5	-424.21	**0.02**	796.18	0.09	0	0
SpreadingArea.A0	2.96	0.97	-127.36	0.78	0	**1**
SpreadingArea.A1	-193.72	**0.01**	-194.71	0.35	0	0
SpreadingArea.A2	-127.92	0.15	244.21	0.53	0	0
SpreadingArea.A3	12.39	0.89	-129.64	0.73	0	**-1**
SpreadingArea.A4	-286.57	**0.00**	-420.07	0.05	0	**-1**
SpreadingArea.A5	-300.03	**0.01**	315.29	0.06	0	0
xcg.A0	0.02	0.26	0.07	0.34	0	**1**
xcg.A1	0.01	0.30	0.07	0.11	0	0
xcg.A2	0.04	**0.03**	0.03	0.76	0	0
xcg.A3	0.07	**0.00**	-0.08	0.20	0	0
xcg.A4	0.12	**0.01**	0.08	0.73	0	**1**
xcg.A5	0.16	**0.04**	0.25	0.44	0	0
ycg.A0	0.01	0.44	0.08	0.42	0	0
ycg.A1	0.02	0.08	0.08	0.16	0	0
ycg.A2	0.01	0.36	0.04	0.57	0	0
ycg.A3	0.01	0.38	-0.05	0.28	0	0
ycg.A4	0.03	0.20	-0.06	0.60	0	0
ycg.A5	0.05	0.09	0.14	0.31	0	0

Hake Bay of Biscay			**CUSUM diagnostics table**												
Years	**MFA_spatial**	**PCA_biological**	**Ln_N_A0**	**Ln_N_A1-5**	**L25**	**Lbar**	**L75**	**Z**	**EA.A5**	**SA.A5**	**PA.A5**	**EA.A4**	**SA.A4**	**xcg.A3**	**Diagnostic**
1987	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1988	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1989	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1990	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1991															
1992	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**-1.4**	**-1.3**	0.0	ref
1993															
1994	0.0	0.0	0.0	0.0	0.0	**-1.0**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1995	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**-1.2**	0.0	0.0	0.0	0.0	0.0	ref
1996															
1997	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	ref
1998	0.0	**1.5**	**-1.1**	**-4.1**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**-1.8**	
1999	0.0	**1.4**	**-1.1**	**-2.6**	**2.0**	0.0	0.0	0.0	0.0	**-2.0**	**-1.2**	0.0	**-1.1**	**-3.0**	
2000	**1.5**	0.0	0.0	**-4.3**	**1.7**	0.0	0.0	0.0	0.0	**-2.4**	**-1.8**	**-1.6**	**-1.7**	**-3.0**	alarm
2001	**3.3**	0.0	0.0	**-3.2**	**2.6**	0.0	0.0	0.0	**-1.4**	**-2.3**	0.0	**-2.8**	**-2.3**	**-3.1**	alarm
2002	**4.3**	0.0	0.0	**-2.2**	**1.5**	0.0	0.0	1.5	**-1.6**	**-2.7**	0.0	**-4.3**	**-4.1**	**-3.3**	alarm
2003	**2.6**	**1.1**	0.0	**-2.3**	**2.8**	**1.1**	**1.1**	**1.2**	**-1.3**	**-2.7**	0.0	**-5.2**	**-5.3**	**-3.2**	alarm
2004	**3.5**	**1.1**	**1.2**	0.0	**3.9**	0.0	0.0	**1.5**	**-1.3**	**-2.6**	0.0	**-6.0**	**-6.8**	**-3.9**	alarm

Fig. 5: Trend results table (above) and Cusum diagnostic table for Bay of Biscay hake.

Results of trend analysis

	all period	recent
Z	NA	NA
Ln_Abdnce	1	1 (linear)
Lbar	0	-1
L25	0	-1
L75	0	-1
Ln_Recruit	0	0

diagnostic No clear diagnostic can be deduced.

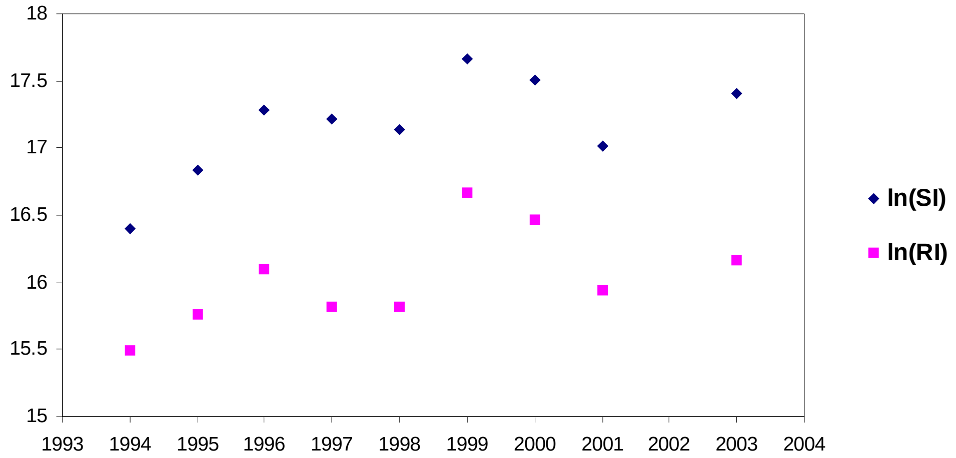
a) The senario of increased recruitment is not supported by the 0 trend of ln_rec and by the recent decreasing trend of L75.

b) The senario of slower growth is not supported by the increasing trend of abundance and the recent decreasing trend of L25.

Hake Ionian **CUSUM traffic light diagnostic table**

Year	MFA_Spatial	PCA_biological	ln_Not	ln_Rec	Lbar	L25	L75	ln_Matures	ln_A2	ln_A3	ln_A4	ln_A5	diagnostic
1994	0	0	**-1.8**	**-1.9**	0	0	0	0	0	0	0	0	
1995	0	**1.6**	0	**1.5**	0	0	0	0	0	0	0	0	
1996	0	0	0	0	0	0	0	0	0	0	0	0	
1997	**4.0**	0	0	0	0	0	0	0	0	0	0	0	
1998	**1.9**	0	0	0	0	0	0	0	0	0	0	0	ref
1999	**2.1**	0	0	0	0	0	0	0	0	0	0	0	ref
2000	0	0	0	0	0	0	0	0	0	0	0	0	ref
2001	0	0	0	0	0	0	0	0	0	0	0	0	ref
2002													
2003	**2.0**	**2.6**	0	**2.0**	0	0	0	0	0	0	0	0	

Fig. 6: Trend results table (above) and Cusum diagnostic table (below) for Ionian hake.



Results of trend analysis

	all period	recent
Z	0	0
Ln_Abdnce	0	0
Lbar	0	0
L25	0	0
L75	0	0
Ln_Recruit	0	0

diagnostic No apparent trends during the studied period.

Hake Aegean Sea **CUSUM diagnostics table**

Year	MFA_spatial	PCA_biological	Ln_Ntot	Ln_Rec	Lbar	L25	L75	Z	Ln_Matures	diagnostics
1994	0	**10.8**	**-2.1**	0	**2.9**	**5.4**	**1.5**	0	0	alarm
1995	**3.2**	**6.8**	**-2.7**	0	0	**3.3**	0	0	**-2.5**	alarm
1996	**3.6**	0	**-1.9**	0	**-2.4**	0	**-1.7**	0	**-2.6**	alarm
1997	**1.2**	0	**-1.3**	0	0	**-2.1**	0	0	0	alarm
1998	0	0	0	0	0	0	0	0	0	ref
1999	0	0	0	0	0	0	0	0	0	ref
2000	0	0	0	0	0	0	0	0	0	ref
2001	0	0	0	0	0	0	0	0	0	ref
2002										
2003	**2.9**	0.0	0	0	0	0	0	0	0	

Fig. 7: Time series of the survey index (SI=Ln_Ntot) and the recruit index (RI=Ln_rec), Trend results table (centre) and Cusum diagnostic table (below) for Aegean hake.

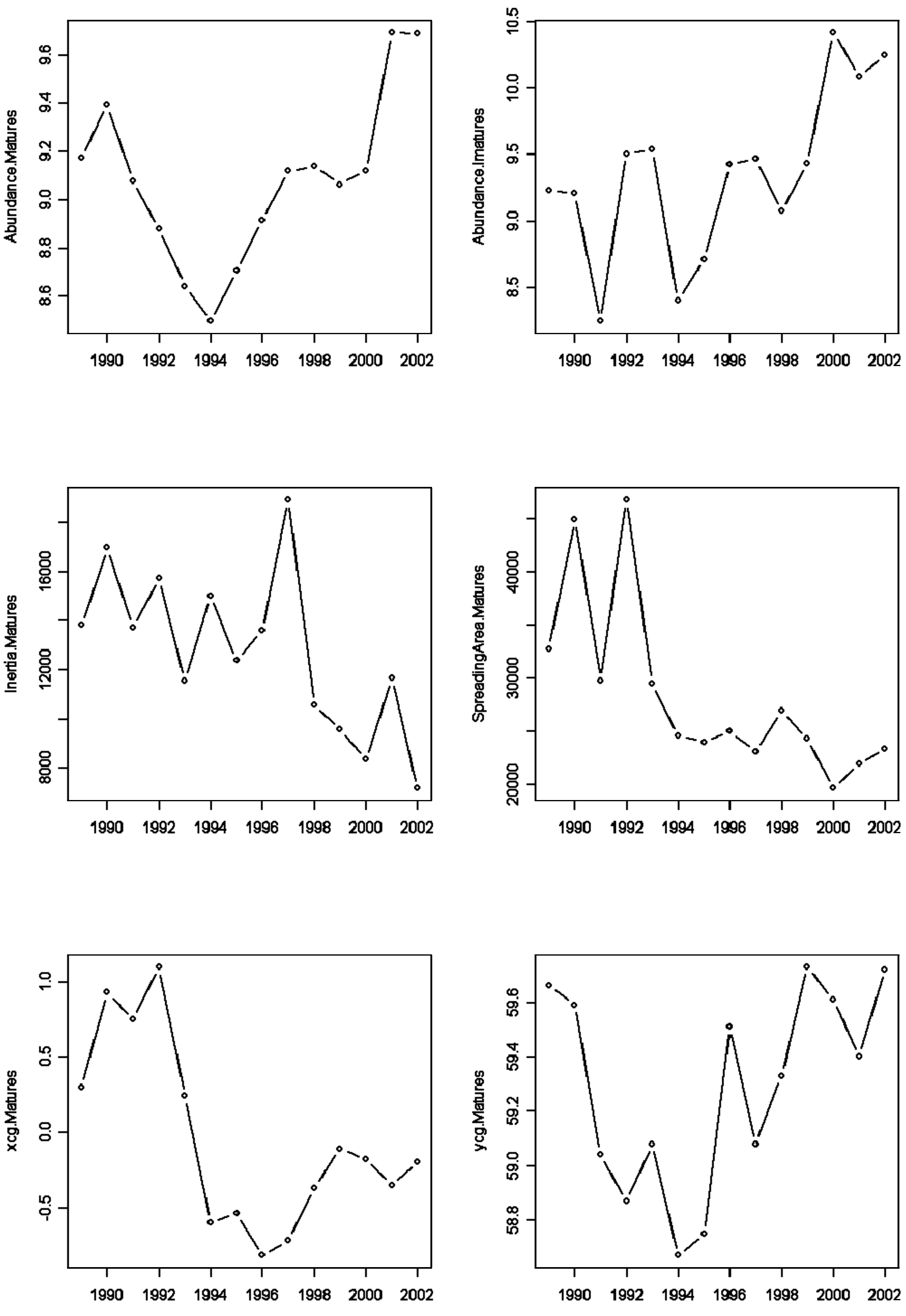


Fig. 8a: North Sea herring raw indices selected using the MAF procedure then visually chosen to evidence the major changes

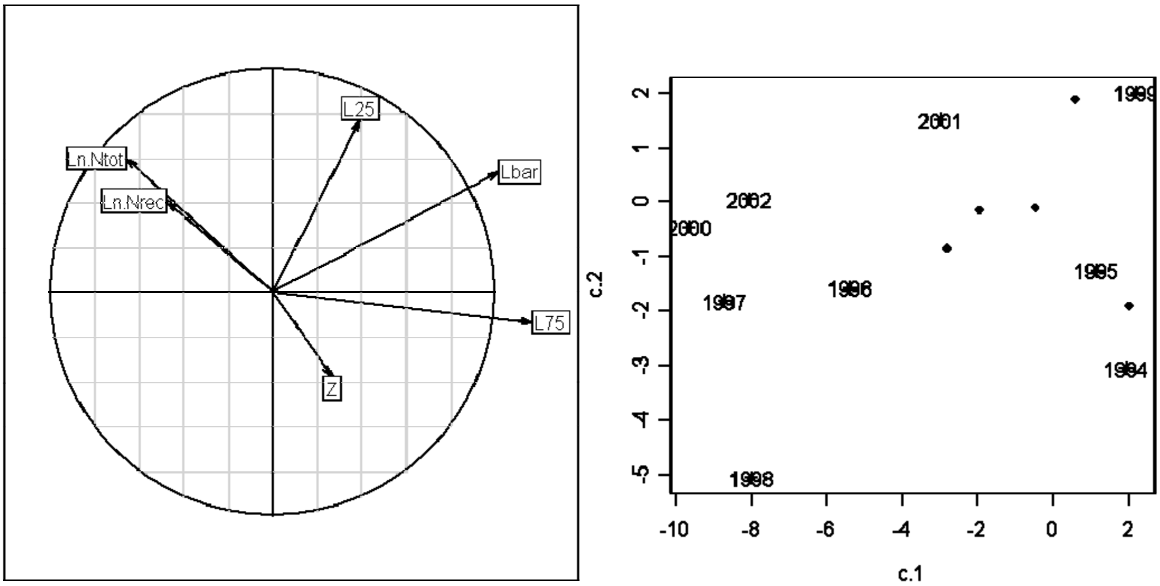


Fig. 8b : North Sea herring multivariate representation of the biological (non spatial) indices (left) and the years (right), in the first plane of the PCA. The reference years of represented by black dots.

North Sea herring **CUSUM traffic light diagnostic table**

Years	MFA_Spatial	PCA_biological	diagnostic
1989	0	0	ref
1990	0	0	ref
1991	0	0	ref
1992	0	0	ref
1993	0	0	ref
1994	**1.2**	0	alarm
1995	**4.6**	0	alarm
1996	**2.2**	**1.1**	alarm
1997	0	**5.1**	
1998	0	**9.6**	
1999	0	**9.4**	
2000	0	**14.1**	
2001	0	**13.5**	
2002	0	**16.8**	

Fig. 8c: North Sea herring Cusum diagnostic table for the multivariate indices

Anchovy Bay of Biscay **CUSUM diagnostics table**

Years	PCA_spatial_AC	PCA_biological_AC	PCA_spatial_EG	diagnostic
1989	0	**5.9**	0	
1990	0	0	0	ref
1994	0	0	0	ref
1997	0	0	0	ref
1998	0	0	0	ref
2000	0	0	0	ref
2001	0	0	0	ref
2002	0	**9.0**	0	
2003	0	**21.3**	0	alert
2004	**3.0**	**16.7**	0	alarm
2005	**6.5**	**48.1**	0	alarm

Fig. 9: Bay of Biscay anchovy

Biological and Spatial indices	all period 10a&b	recent 10a&b	all period 10a	all period 10b	recent 10a	recent 10b
Z	0	1				
Ln_Abdnce	0	0				
Lbar	0	0				
L25	0	0				
L75	0	0				
L50mat	0	0				
Ln_Recruit	0	0				
xcg (age1)			-1	1	0	0
xcg (age2)			-1	1	0	0
xcg (age3)			-1	0	-1	0
ycg (age1)			1	0	0	0
ycg (age2)			1	1	0	0
ycg (age3)			1	0	NA	NA
Inertia (age1)			1	0	0	0
Inertia (age2)			0	0	-1	-1
Inertia (age3)			0	0	1	0
Anisotropy (age1)			0	0	ND	ND
Anisotropy (age2)			1	0	ND	ND
Anisotropy (age3)			0	0	ND	ND
Positive area (age1)			0	0	0	0
Positive area (age 2)			0	0	-1	0
Positive area (age 3)			0	0	0	-1
Equivalent area (age1)			0	0	0	0
Equivalent area (age2)			0	0	0	0
Equivalent area (age3)			0	0	0	0
Spreading area (age1)			1	0	0	0
Spreading area (age2)			0	0	0	0
Spreading area (age3)			0	0	0	0
Microstructure (age1)			0	0	ND	ND
Microstructure (age2)			0	0	ND	ND
Microstructure (age3)			-1	0	ND	ND

ND=not determined

mul TS				**CUSUM diagnostics table**							
Year	**PCA_spatial_10a**	**PCA_spatial_10b**	**PCA_biological**	**Ln_Ntot**	**Ln_Rec**	**Lbar**	**L25**	**L75**	**L50.matu**	**Z**	**alert**
1994	0		0	0	0	0	0	0	0	0	
1995	0	**1.2**	0	0	0	0	0	0	0	0	
1996	0	**1.2**	0	0	0	0	0	0	**1.7**	0	
1997	0	0	0	**-2.0**	0	0	0	0	0	0	alert
1998	0	0	0	0	0	0	0	0	0	0	
1999	0	0	0	0	0	0	0	0	0	0	ref
2000	0	0	0	0	0	0	0	0	0	0	ref
2001	0	0	0	0	0	0	0	0	0	0	ref
2002	0	0	0	0	0	0	0	0	0	0	ref
2003	0	0	0	0	0	0	0	0	0	0	ref

Fig. 10: Trend results table (above) and Cusum diagnostic table (below) for the Red mullet in the Thyrrhenian Sea

Annex 1: Template for reporting case studies indicator-based diagnostics

Case study NAME

Each of the following items with comments (NA if not done)

Data :
- Map of all survey stations overlaid showing polygon used.
- For spatial indices : 2 maps of gravity centres across years for selected ages in immature and mature ages
- Input parameters for spatial indices : function infl() , function NBPatches() , function Microstructure()
- Raw indices : Tables of spatial and non-spatial indices (wp2a tables 1 and 2)
- Combined indices : (retain the 2 first principal axes) fig. of factorial representation, table of indices values

Looking for changes :
- visual inspection : plots of selected indices (raw & combined, expert or MAF-based)
- trend plots of selected indices (provide plots, specify trend method used, fill trend diagnostic table)
- di-cusum plots of selected indices (provide plots, fill cusum diagnostic table)

template for diagnostic tables are in file : indic_diagno_tables_nantes.xls

Interpretation :
comment diagnostics tables results
- trend analysis : interpretation using cause-effects table as guide line
- cusum analysis :
- interpretation using cusum table of selected indices
- interpretation using cause-effects table as guide line

Compare approaches (cusum/trends)

What have you learned ?

Summary sheet
- Survey series (Periods / Seasons / Type)
- Non-spatial indices (a few words : has index been analysed ? what method for change? change detected ?)
  - Abundance index, Recruitment index
  - Lbar, L75, L25
  - L50.maturity
  - Z by year
- Spatial indices (a few words : index analysed ? by age or stage ? what method ? change detected ?)
  - Positive Area, Spreading area, Equivalent area
  - Centre of gravity, Inertia, Anisotropy
  - Microstructure
- Composite (derived) indices (a few words : method ? index used ? components 1 & 2 dominated by which raw indices ? change detected ?)
  - MAF, MFA, PCA
- Reference period (which years ? comments on choice of period)
- Summary of results on the stock (comments on data series, ref period, changes evidenced, which method support summary)

Comparison with traditional assessment of stock status :
traditional assessment = scientific diagnostic by expert groups, not official advice
short text with following topics : have alerts been triggered for similar years ? has an early warning been possible using indicators ? what do we gain with all indicators in comparision to abundance only ?

Formulation of advice (based on all the above, can you formulate an advice ?)

Document 3: Survey data only Assessment models

ICES CM 2007/O:04

A review of Fishery-Independent assessment models, and initial evaluation based on simulated data

John Cotter, Rob Fryer, Benoit Mesnil, Coby Needle, Dankert Skagen, Maria-Teresa Spedicato, Verena Trenkel

Abstract

Large uncertainties in the catch data (official landings and discards) are undermining ICES' ability to provide valid management advice based on the conventional approach of analytical assessments. There is thus an urgent need to consider alternative tools that do not depend on long series of precise catches, with their age composition. This paper presents a few fishery-independent assessment models developed by the EU project FISBOAT (Fishery Independent Survey Based Operational Assessment Tools). It also reports on rudimentary tests based on simulated data, following the same protocol as an evaluation study conducted by the US National Research Council in 1997. It appears that the survey-based assessment models at hand are able to reliably capture the major signal in biomass and recruitment, although they smooth out transient changes. However, they cannot provide absolute abundance estimates, but only relative values on an arbitrary scale. Their operationalisation in ICES would thus require an adaptation of the advisory framework, in terms of nature of the advice and definition of reference points; indeed, this might be needed anyway, if we were more lucid about the myth of VPA estimates being absolute. It is also noted that survey-based approaches have the potential to provide much more rapid updates of the state of stocks than catch-based methods.

Contact author: Benoit Mesnil; Département Ecologie et Modèles pour l'Halieutique; IFREMER, B.P. 21105, Nantes Cedex 03, France. [tel. +33 240 37 40 09; fax +33 240 37 40 75, e-mail Benoit.Mesnil@ifremer.fr]

Introduction

All stock assessment methods (whether they involve surplus-production, delay-difference, stock-reduction, Collie-Sissenwine or analytical dynamic pool models) used by scientific organisations to advise fisheries managers on the state of fish stocks require a knowledge of total catches to estimate the model parameters and other quantities of management interest. Errors in the input catch figures translate directly into similar errors in stock abundance estimates (e.g., Quinn and Deriso, 1999), and if their magnitude varies from year to year the assessments may not even reflect the relative changes in the state of the resources. When catch is also the support of management control, like in TAC systems, there is often a temptation for fishers or states to mis-report for tactical reasons, especially when catch quotas become very restrictive. This has been the case in Europe in recent years and for over a decade ICES has repeatedly stated that the deterioration of the catch data was threatening its ability to provide managers with the type of advice they require to run the current policies.

In the face of this threat, the European Commission has identified the development of operational fishery-independent (i.e. catch-free) assessment tools as one of the priority topics of research in support of the Common Fisheries Policy in the 2003 round of calls for scientific projects. This has been taken up by the FISBOAT (Fishery Independent Survey Based Operational Assessment Tools) project consortium, with participants from 11 research institutes, which completed its work by mid-2007. In this paper we only report on the findings of one work package which was tasked to "*supply methods for analysing fishery independent stock assessment data to provide managers with relevant information about the stock and its exploitation*". Methods are here understood to mean both the mathematical models and the procedures to estimate their parameters.

Section 1 provides a concise overview of the six fishery-independent (F-I) stock assessment methods that were specifically developed, or elaborated upon for use without catch data, during the project. The models are presented with comments on parameter estimation issues, and practical guidelines or caveats regarding their use in assessment and advisory groups are provided.

One final products of the FISBOAT project is a full evaluation of the survey based methods through a simulation-testing evaluation framework (operating model, harvest rule, etc.), but some elements are missing to run that at the moment. However, in order to gain some early understanding of the capabilities of the F-I models, the group had decided to carry out a simpler testing exercise using artificial data with known properties, following the same protocol as an evaluation study conducted in the USA on catch-based (mostly age-structured) assessment models (NRC, 1998). Sections 2 and 3 recount the conditions and results of these preliminary probing tests carried out on four of the six models. Section 4 concludes on the insight gained during the project into the potential performance of the F-I methods for assessment of stock status, and on some implications for the European (ICES) advisory system.

1. Methods considered

The F-I methods developed, or adapted, for the FISBOAT project fall into two categories:
1) methods intended to estimate abundance, or trends thereof; this group includes stage-structured BREM, age-structured SURBA, TSA and YCC, and length-structured LENSUR;
2) simulation methods, to assess the effects of changes in biological or management parameters, represented by ALADYM.
Key features of each method are described in this section, starting with estimation methods. A summary categorisation of the methods with regards to data needs and estimation approach is provided in Table 1. The computer codes and documentation are available on the FISBOAT website.

1.1. Biomass random effects model (BREM)[1]

- Model description
The population dynamics is formulated as the difference model from Hilborn and Walters (1992, p. 336):

$$B_t = R_t + g_t B_{t-1} \qquad (1)$$

where B_t is the total population biomass, R_t the recruitment in biomass in year t-1 and g_t the net biomass growth rate, which is the balance between individual growth and total (natural + fishing) mortality. Recruitment is assumed to follow a logNormal distribution without any stock-recruitment relationship:

$$\log(R_t) \sim N(\mu, \sigma_R^2) \qquad (2)$$

Biomass growth is modelled by a random walk on the log-scale, to reflect the assumption that Z, which is part of g, does not vary wildly from year to year:

$$\log(g_t) = \log(g_{t-1}) + \varepsilon_t^g \text{ with } \varepsilon_t^g \sim N(0, \sigma_g^2). \qquad (3)$$

$$g_t = g_{t-1} \exp(\varepsilon_t^g)$$

Thus both recruitment R_t and biomass growth g_t are treated as random effects with parameters μ and σ_R^2, and g_1 (t=1) and σ_g^2 respectively.

The observation model has two components. The first one is for an index of total biomass at time t (recruits included) and the second for an index of recruits only. Both are assumed to follow logNormal distributions with common variance and catchability coefficient:

$$\log(IB_t) \sim N(\log(q_b B_t), \sigma_{ib}^2) \qquad (4)$$

$$\log(IR_t) \sim N(\log(q_r R_t), \sigma_{ir}^2). \qquad (5)$$

In order to ensure identifiability, the following constraints are imposed: $q_b = 1$ and $\sigma_{ib}^2 = \sigma_{ir}^2$.

- Sensitivity and robustness issues
Convergence of the parameter estimation algorithm depends critically on sensible starting values. The above mentioned constraints allow parameter identifiability, but the effect of setting $q_b = 1$ is that biomass estimates can only be relative not absolute. In addition, the estimates of recruitment and catchability for recruits q_r are confounded to some degree. This appears as strong correlation between estimates.

- Input and Output
BREM only requires two series of survey indices in mass, one for the total population (adults + recruits) and one for the recruits alone; splitting out the recruits can be based on age readings but there are favourable cases where a reasonable cut-off size may be identified by inspection of the length compositions. Note that knowledge of M is not required, and that occasional gaps in survey series are not likely to affect the estimation. An extension handling two series of indices per category (e.g. acoustic and egg surveys) has been developed (Trenkel, 2006, 2007).

---

[1] Contributed by Verena Trenkel, Ifremer, France

Seven parameters are estimated: B_1 (biomass in year 1), g_1 (biomass growth in year 1), $\log(\sigma_g)$ (standard deviation of growth), μ (mean recruitment for base normal), $\log(\sigma_R)$ (standard deviation of recruitment for base normal), q_r (catchability of recruits) and σ_i (standard deviation of observation error for base normal). Plugging converged estimates into Eq. 1 yields estimated time trajectories of relative total biomass and annual recruitment. In addition, standard deviations are available for biomass estimates, but NOT for recruitment estimates as these are random effects, not real parameters.

- Implementation issues
Parameter estimation by maximum likelihood is implemented in AD model builder (Fournier 2005) using the random effects module. Run time for NRC set 1 was about 20 sec. Note that run time does not increase with the number of years. Rather it depends on how good the starting values are.

- Miscellaneous comments
Future recruitment could be predicted using the fitted logNormal distribution, either as expected recruitment or by drawing a random recruitment value from the distribution. The relationship between model predictions and commercial quantities is not obvious.

1.2. LENSUR[2]

- Model description
Lensur is a newly written program for assessing a stock with only length-structured data. *Lensur* has an operating model that generates an artificial population in numbers by length class and time step, as specified by a set of parameters. Model observations are derived from the operating model in an observation model, and parameters are estimated by minimising the deviation of the model observations from real observations, as expressed by an objective function. The objective function so far is a sum of squared log residuals. A minimisation routine searches over the space of parameters and calls the objective function for each parameter set, to find the parameter set that gives the best value, i.e. the best model fit. This is regarded as the estimate of the population according to the data. This places the method within the framework of ‘statistical models’, where in this case the population is constructed so that stock numbers are represented by length.

The method used to obtain length distributions in the population is to follow an ensemble of trajectories representing ‘super-individuals’ over time, each with its own growth characteristics, and time and length at entry (a 'Lagrangian' approach). Hence, internally the population is represented both by length according to a growth model, and age, represented by the time that has passed for each trajectory since it entered the populations. Each trajectory enters the population at a randomly drawn time with a certain number of fish and randomly drawn growth parameters. The abundance and length of each trajectory is calculated for each time step. The whole population is the sum of all trajectories.

- Sensitivity and robustness issues
The implementation in Fisboat further restricts the data to survey data only. This implies strong limitations on what can be inferred from the data in terms of population abundance and exploitation rates, and the method has so far primarily been used to study these limitations (see miscellaneous comments below).

- Input and Output
Operating model parameters are initial numbers at length, numbers (recruits) entering the population each year, growth parameters (k and Linf), selection at length in the fishery, annual

[2] Contributed by Dankert Skagen, IMR, Bergen, Norway

fishing mortality and natural mortality. Parameters in the observation model are survey catchabilities. These are specified as separable, i.e. with a catchability at length and a year factor. The latter normally is assumed to be constant. The program allows the user to specify, for each parameter, whether this parameter shall be estimated by the optimisation routine, or remain fixed at a given value.

Catches at length in each time step are derived from the abundance at length and fishing mortalities at length. These results are not used in the Fisboat framework, but may be useful if catch data (at length or in biomass) are available.

- Implementation issues

Lensur is programmed in Fortran 77, and will be implemented in FLR in the near future. The FLR version is a slightly reduced one, in particular with respect to input-output.
Optimisation is by a searching routine, which is slow, but very robust. This may be an advantage when the method is incorporated in a framework like FLR where there is no user interaction when the program is run.

- Miscellaneous comments

The program can also be used as a data generator, by extracting catches or the output from the survey observations model as artificial data, and numbers from the operating model as the true stock. Noise can be added to the data output, either as random noise, as a random year factor or both. Such data were used in studies on performance. First, it has been confirmed that with no noise in the data, the model fits the data virtually perfectly, and with the right population numbers. Further exploration of the model performance has concentrated on the limitations in what can be inferred with these sparse data. Theoretical considerations and the experience gained using the model on artificial data, leads to the following conclusions about the method:

- Survey indices at length by themselves carry insufficient information for a full stock assessment. Firstly, all such data are relative, and some additional constraint is needed to scale the stock abundance to absolute values. Furthermore, growth and mortality are confounded in the sense that they influence the length distributions in the surveys in a complementary way. Hence, mortality estimates are conditional on assumptions on growth rate. Finally, noise in the data is amplified when translated into mortality estimates. Therefore some constraining assumptions have to be made on the mortalities, and the results are conditional on these assumptions.
- The experience so far is that simple smoothing of the indices is clearly insufficient to avoid undue influence of random noise in the data. Applying a penalty on the year-to-year variation in F takes most of the noise away, and combining that with smoothing of the survey indices removes even more of the noise from the results. When the true fishing mortality is variable, such variations become damped, however.
- Given the theoretical limitations outlined above, it appears that by assessing a stock with only survey indices at length in a statistical method with a length disaggregated model population, it is not possible to provide reliable estimates of variations in exploitation. However, estimates of the stock and the level of, and trends in, the exploitation can be achieved conditional on assumptions about growth rate. The estimates will also be conditional on assumptions about trends in exploitation. If these assumptions are realistic, the estimates obtained will be so as well.

Therefore, in a management context, having length disaggregated survey data as the primary source of information about the stock is only likely to work if there is additional information on trends in the exploitation rate. The assessment will then provide the information about trends in stock abundance on a relative scale, which at least in principle can be translated into harvest rules.

1.3. SURBA[3]

- Model description

The basis of *SURBA* is a simple survey-based separable model of mortality. This model was first applied to European research-vessel survey data by Cook (1997, 2004), but it has a long history in catch-based fisheries stock assessment (Pope and Shepherd 1982, Deriso *et al* 1985, Gudmundsson 1986, Johnson and Quinn 1987, Patterson and Melvin 1996; see Quinn and Deriso 1999 for a summary). The separable model used in *SURBA* assumes that total mortality $Z_{a,y}$ for ages *a* and *y* can expressed as:

$$Z_{a,y} = s_a \times f_y,$$

where s_a and f_y are respectively the age and year effects of mortality. Note that this differs from the usual assumption in that total mortality *Z* is the quantity of interest, rather than fishing mortality *F*. Then, given $Z_{a,y}$, abundance $N_{a,y}$ can be derived as:

$$N_{a,y} = r_{y_0} \exp\left(-\sum_{m=a_0}^{a-1} \sum_{n=y_0}^{y-1} Z_{m,n} \right)$$

where a_0 and $y_0 = y - a - a_0$ are respectively the age and year in which the fish measured as $N_{a,y}$ first recruit to the observed population. Thus the abundance at each age and year of a cohort is given by the recruiting abundance r_{y_0} of the relevant cohort modified by the cumulative effect of mortality during its lifetime. Parameters are estimated by minimising the weighted sum-of-squares of observed and estimated abundance indices. All abundance estimates are relative.

This simple basis has been expanded considerably over recent years, as the model has been road-tested in ICES assessment working groups (and elsewhere) and modified where necessary. The development is summarised in Needle (2002b, 2002d, 2003d, 2004a, 2004b) and Beare *et al* (2005), but in brief:

- Index catchabilities and SSQ weightings can both be defined by the user.
- Biomass indices can be used, as well as multiple age-structured indices.
- The year-effect for the last year is set to the mean of the previous three year effects, as the terminal year-effect cannot be determined directly from the data (although work is progressing on improving this estimate; see below).
- Age-structured indices are all back-shifted to the start of the year, using the current estimate of *Z*. This allows them to be compared directly, and ensures firstly, that abundance indices refer to Jan 1, and secondly, that mortality estimates relate to the calendar year rather than the year between successive cruises of a given survey.
- Biomass indices are shifted forwards to spawning time before inclusion in the parameter estimation process.
- Optionally, a smoothing term can be added to the SSQ to penalise excessive inter-annual variation in estimated year effects. The degree of smoothing is determined by a user-defined variable λ.
- The reference age (that is, the age at which the age-effect *s* is fixed to 1.0) can also be defined by the user.
- Estimated variances (and thereby confidence intervals) of mean *Z* and recruitment are derived from the variance-covariance matrix of the model fit, using the delta method. Variances for abundance and SSB are currently being implemented.
- Retrospective runs are generated automatically, with the last year of data being moved back one year at a time until half of the original time-series remains. This facility can be switched off by the user if required.
- A scan facility has recently been added. With this, the user can automatically run assessments with a range of choices for smoothing, the reference age, and catchability on the first age, and evaluate model sensitivity to these essentially *ad hoc* settings.

[3] Contributed by Coby L. Needle, FRS Marine Laboratory, Aberdeen, UK.

Planned future work includes:

- Improving the terminal year-effect estimate. This may be possible if there are two or more surveys at different times of year, in which case the relative decline in indices during the year may give an idea of mortality during the year.
- Implementation of bootstrap uncertainty estimation, via a multivariate parametric bootstrap.
- Improved scanning procedure.
- Restructured input dialogues.
- Inverse-variance SSQ weighting.
- The results of the scan procedure cannot currently be plotted within *SURBA* itself, although an SPLUS script is provided. The plotting procedure needs to be updated to accommodate this.

- Sensitivity and robustness

The model is most sensitive to assumptions about catchability. In particular, estimates of *Z* can be very different under different assumptions about catchability; SSB estimates are more robust. *Z* estimates can be very uncertain in any case, and it is not uncommon for there to be no significant evidence of any changes in the levels of *Z*. However, this may well be true for most models in which uncertainty is estimated. The ICES North Sea Demersal WG encountered difficulties in fitting *SURBA* to flatfish survey data during its 2005 meeting, and these are still not resolved. Finally, the automated scanning routine sometimes fails – values scanned over need to be interactively defined in future.

- Inputs and output

SURBA uses the Lowestoft VPA input format, and currently expects to see the full set of such files – which means that dummy catch-based data files had to be set up in order to analyse the NRC datasets. The inputs that are actually required for fitting the model are age-structured tuning indices, and (optionally) biomass tuning indices. The user can also define catchability and SSQ weightings for both types of index, along with values for the smoother λ and the reference age.

Both text and graphical outputs are provided by the program. Text outputs include parameter estimates with variances, mortality and relative abundance estimates, estimated variances for mean *Z* and recruitment, log residuals, stock summaries (SSB etc.), results of retrospective and scan runs, and goodness-of-fit statistics. Plots include exploratory raw-data figures (such as catch curves), model fits and stock summaries, residuals, and retrospective summaries.

- Implementation issues

SURBA (currently Version 3.0) is implemented in a Windows user interface, in which diagnostic plots are automatically generated. The run time for NRC set 1 was 6 s (standard), 40 s (standard + 15 retrospective runs), and 7 m 47 s (105-run scan).

- Predictive ability

SURBA does not currently feature a forecasting mode, although this is planned in the near future. It is intended that this will roll forwards the population from different starting points arising from the bootstrap runs mentioned above, leading to stochastic forecasts. This will need assumptions about weights, exploitation, and recruitment.

- Relation to management indicators

Abundance estimates (and therefore biomass measures) are currently generated by *SURBA* on a relative scale only, and are usually plotted as mean-standardised values for ease of comparison. Furthermore, *SURBA* provides estimates of total mortality Z rather than fishing mortality *F* (although, given the tentative nature of most natural mortality estimates, this is true of catch-at-age methods also). Therefore *SURBA* can be used to provide advice on relative trends in abundance and total mortality, but not absolute levels. It is possible to generate pseudo-absolute abundance estimates by using a catch-at-age VPA to estimate survey catchabilities-at-age using data from some period in the past, and then applying these to recent *SURBA*-derived relative population estimates to scale them up to a level commensurate with that indicated by catch data (Needle 2004a). However, this requires assumptions that there was a period when catch data were reliable, and that the relationship between survey and fishery catchability has remained constant ever since, and these can be hard to maintain. It is also possible, of course, to produce F estimates by subtracting fixed M values from the Z estimates produced by *SURBA*.

If *SURBA* (or any other survey-based approach) is to be used as a management tool, there needs to be a clear idea of the management framework in which such a tool would be used. In other words, reference points for mortality and biomass would need to be redefined on the basis of total mortality and relative biomass, respectively.

1.4. Time series analysis (TSA)[4]

- Model description

TSA, or 'Time Series Analysis', is a state space framework for modelling a fishery. The initial implementation, by Gudmundsson (1994), modelled commercial catch-at-age data, with survey indices-at-age used as auxiliary information. Here, the framework is adapted to model the indices-at-age from a single survey. The state equations relate the log numbers-at-age and fishing mortalities-at-age in year y to those in year y–1. Log numbers-at-age in year y are given by:

$$n(a,y) = n(a-1, y-1) - Z(a-1, y-1) \quad a > 1$$

$$n(1,y) = \mu + \mathrm{NID}(0, \sigma^2_{\mathrm{recruit}})$$

where NID stands for Normal Independent Deviate. Fishing mortalities evolve according to the following model:

$$\log F(a,y) = U(a,y) + V(y)$$

$$U(a,y) = U(a, y-1) + \mathrm{NID}(0, \sigma^2_U) \quad \text{with the constraint that } \textstyle\sum_a U(a,y) = 0$$

$$V(y) = Y(y) + \mathrm{NID}(0, \sigma^2_V)$$

$$Y(y) = Y(y-1) + \mathrm{NID}(0, \sigma^2_Y)$$

Thus, log fishing mortality is separated into an age component $U(a,y)$ and a year component $V(y)$, both of which can evolve over time. Finally, the state vector consists of the $n(a,y)$, log $F(a,y)$, $U(a,y)$, $V(y)$ and $Y(y)$.

The observation equations are given by:

$$i(a,y) = q(a) + n(a,y) + \varepsilon(a,y)$$

where $i(a,y)$ are the log indices-at-age, $q(a)$ are the survey log catchabilities, and the $\varepsilon(a,y)$ are assumed to be NID with zero mean and standard deviation $\sigma_{\mathrm{survey}}\ \lambda(a)\ \delta(a,y)$. The $\lambda(a)$ are initially taken to be unity, but can be adjusted later if the errors associated with some ages are larger than for others. The $\delta(a,y)$ are also initially taken to be unity, but can be inflated to decrease the influence of outliers. It is assumed that the survey takes place at the start of the year.

The model is fitted using the Kalman Filter, with the parameters μ, $\sigma_{\mathrm{recruit}}$, σ_{survey}, σ_U, σ_V, σ_Y, $q(a)$, $U(a,1)$ estimated by maximum likelihood. For identifiability, $q(1)$, $V(1)$ are taken to be zero. For stability, some constraints must be put on the $q(a)$: the current implementation takes the $q(a)$, a > 1, to change linearly with age.

[4] Contributed by Rob Fryer, FRS Marine Laboratory, Aberdeen, UK.

- Sensitivity / robustness

Good starting values can be difficult to find for a new stock: some iteration and experience is required. Poor starting values will either make the current implementation crash or will erroneously suggest that the starting values are optimal. However, once good starting values have been found, the implementation is robust to the addition of an extra year's data, etc. The same starting values were used for all the NRC data sets (except for the clean set).

The method works on the log scale, so zero indices must be replaced by some small positive value. Unity was used for the NRC data sets. This means that the method can only be sensibly applied to those age classes where zero indices do not often occur – typically the younger age classes. An option would be to group older age classes into a single plus group, but this has not been implemented yet.

Very large year classes can cause a problem, because they can unduly dominate the parameter estimates associated with recruitment (i.e. μ and $\sigma_{recruit}$). It is possible to reduce their impact on these estimates, but this is done manually following graphical inspection of standardised prediction errors.

- Inputs and outputs

Inputs:

- survey indices-at-age; can handle several surveys in sequence (e.g. if there is a change in q in the single survey at a particular time), but not in parallel. Missing survey data are accepted, at the cost of increased standard errors in estimates around missing years while not causing bias.
- if natural mortalities-at-age are provided, then (relative) fishing mortalities will be estimated, otherwise (relative) total mortality Z will be estimated.

Outputs:

- estimates of relative numbers-at-age with approximate coefficients of variation; these can not be combined across age classes (there is a separate scaling factor for each age class), so it is not possible to estimate (relative) biomass, etc; however, sensible proxies for stock biomass can be estimated;
- estimates of relative fishing mortalities-at-age with approximate coefficients of variation; these can be combined across age classes, so it is possible to estimate (relative) mean fishing mortalities for groups of age classes;
- evidence of persistent changes in fishing mortality, either overall, or as departures from separability.

- Implementation

- Fortran 90, using NAG routines.
- Took ~ 30 seconds to run NRC set 1 on a 1.8GHz, 524MB RAM laptop.

- Predictive ability

The method can predict both relative numbers-at-age and fishing mortalities-at-age (with approximate coefficients of variation) as far into the future as required.

- Relationship with commercial quantities

If natural mortalities-at-age are provided, then relative fishing mortalities-at-age are estimated (otherwise only relative total mortalities-at-age Z are estimated).

1.5. Year-class curve (YCC) method[5]

- Model description

A 'year-class' curve is a plot of log CPUE over age for a single year-class of a species. Marine fish caught in trawls typically show nearly linear year-class curves for ages that are fully selected. The usual model of mortality over time *t*, assuming no net migration to or from the stock, is

$$\frac{dN}{dt} = ZN$$

where *Z*, the instantaneous rate of total mortality, is here expected to have a negative value. [The absence of a minus sign before *Z* is unconventional in fisheries work but leads to Equation (2) having all terms positive, as is conventional for regression models.] Solving gives

$$N_t = N_0 \exp(Zt). \qquad (1)$$

We now assume that catch per unit effort (cpue, denoted *U*) is a constant proportion of *N*, i.e. $U = qN$ for all ages included in the analysis, and that *Z* represents a constant, average value over time. Then, taking natural logarithms of Equation (1), restricting attention to one year-class, *c*, substituting *age* for *t*, and adding a random error term, *e*, gives the basic model for a year-class curve:

$$\log U_{a,c} = \log(U_{0,c}) + Z\,age + e_{a,c} \qquad (2)$$

where $U_{0,c}$ is the cpue (or survey) index for age zero, *a* is the age-class, i.e. the age in years as an integer index, while *age* is age in years as a real number. *e* is assumed to be normally distributed around zero with residual variance ${\sigma_e}^2$. Additional linear terms may be added to equation (2) to allow for varying selectivity of the survey trawl with age, to allow for different RV (or commercial) fleets having different catchabilities, *q*, and to allow for gradual changes in *Z* over time. The latter is achieved using polynomials in *age* and *year* with a minimum of additional parameters so as to yield best precision of estimation with the available observed data.

Different series of cpue data are likely to estimate year-class curves with different precision depending on the season and area covered by the fleet, on the precision of age-reading and other practical aspects, and on how well the chosen model fits the data. Weighting of different data sets to reflect their precision with respect to the chosen model is therefore desirable. Cotter and Buckland (2004) suggest that the weighting estimated for each fleet's data set should be balanced with the reciprocal of the estimated residual variance specific to that fleet computed after the model is fitted, i.e. $\hat{w}_f \propto \hat{\sigma}_f^{-2}$. They describe how the method can be implemented using iteratively weighted least squares (IWLS) taking into account the d.o.f. contributed by each fleet to the estimates of each parameter. Usually, 2 or 3 iterations produce stable values. Additionally, using the fleet specific residual variances, the relative precision of the different fleets can be compared using *F* tests (Cotter 2001). Note that biased cpue series will produce biased weights (Quinn and Deriso 1999, p. 353). Fleets that appear exceptionally precise should be scrutinised to see whether biased sampling may be the cause, e.g. due to clustering of observations in restricted times or places (Cotter and Buckland 2004).

A year-class curve can be fitted repeatedly in a process called *forward validation* that is designed to find the most reliable model for predicting next year's cpue. Starting from an early year and proceeding forwards in the time-series, it finds the differences between the predicted log cpue and the observed log cpue for one year after the time domain of the data used to fit the model. The preferred model is the one whose mean difference is closest to zero, and for which the mean square of the differences is lowest. This is merely a simulation of a fish stock assessment working group making predictions each year for the coming year, then checking

[5] Contributed by John Cotter, CEFAS, UK.

them when the outcome is known. Full details of available models, fleet weighting, and forward validation to find the preferred model are given by Cotter et al. (2007).

- Sensitivity / robustness

Catchability must be constant over time but may vary among different surveys or fleets since intercalibration factors are automatically fitted if required. Changes to the design of an RV survey that might cause a change in catchability (e.g. a different vessel or gear) can be accommodated simply by treating it as a new fleet and fitting an extra intercalibration factor.

Only gradual changes of Z are allowed by using polynomials in *year* to a maximum degree of 3. This is intended to minimise the dangers of erroneously treating random measurement errors as trends in the year-class signal over time. However, if sudden, real changes in Z actually do occur from year to year, they might be overlooked.

Year-class curves can be fitted across fleets, or nested within. Over- and under-fitting can both caused biased estimates of parameters. Forward validation helps to eliminate such models because they tend to be poor at predicting beyond the observed domain. The AIC may also be used to help with finding the best model.

- Inputs and outputs

The basic input is a standard VPA-type tuning file (Darby and Flatman 1994). *YCC* software operates on a flat file having fleet, age, year, time-of-year, cpue, etc., so such a file may be utilised directly if preferred. Year-class curves are available as plots over time, one per year class. These allow the fitted model to be compared to the observed values to check that the fit is credible. Relative recruitments, and Z over age by fleet are also given, along with various other outputs.

- Implementation

Software to fit year-class curves with all the options described here is called *YCC*; it is written in R. The user is asked what terms are wanted in the model, whether terms are to be nested in fleets, whether to switch weighting on or off, whether to use forward validation, and about outputs. The latter may be obtained on screen, as text files, or as graphics. Diagnostics include prediction and residual errors over time, age, and year class. Some of these outputs are illustrated by Cotter et al. (2007). Run times are usually seconds but may increase to a minute or more when there are many fleets, iterative re-weighting, and a long period of forward validation. The model may fail to fit if there are more parameters than observed vectors of cpue-at-age. Missing values may either be omitted from the data set or coded as negative cpue.

- Predictive abilities

Predictions one year ahead of observed data is carried out routinely with forward validation. *YCC* produces tables of predicted cpue-at-age for the year after the final observed year together with prediction mean square errors.

- Relationship with commercial quantities/management indicators

Predicted cpue-at-age in terms of numbers may be converted to weights per unit of effort-at-age using a matrix of weights-at-age by year. These may in turn be converted to spawning stock biomass per unit of effort-at-age using a matrix of maturity-at-age by year. The software allows users to insert independent observed values for each year, if available. Year-class curves fitted to cpue for commercial fleets could provide predicted catches from predicted cpue multiplied by predicted commercial effort under different fishing scenarios. However, *YCC* offers no prediction of next year's recruiting year class.

Z is estimated numerically for each age and year from fitted curves (rather than from fitted parameters which may be individually biased, depending on the model). No assumptions are made about natural mortality, M. Fishing mortality, F, could be estimated if they were.

1.6 ALADYM simulation method[6]

- Model description

ALADYM (**A**ge-**L**ength B**a**sed **Dy**namic **M**odel) is an age-length based simulation model developed in the conceptual framework of dynamic pool models, following the predictive Thompson & Bell (1934) approach. The model is designed to predict, through simulations, the effects of different fishing pressure scenarios on a single population, in terms of different metrics and indicators. Removals are simulated on the basis of the total mortality rate modulated using harvesting pattern and a fishing activity coefficient. *Aladym* can work in absence of fishery-dependent data, although its predictive capability of real catch levels can be verified using information on commercial catches or fishing activity per month.

From the *Aladym* core model three complementary, but independent, tools have been derived:

A) the quasi-deterministic dynamic tool named *Aladym-r*;

B) the tuning tool *Aladym-z*;

C) the stochastic dynamic tool named *Aladym-q*.

Aladym-q adds to the same mathematical model of *Aladym-r* the capability to deal with the stochastic representation, modelling the uncertainty of estimates related to recruitment, growth and maturity through stochastic processes. This makes *Aladym-q* more suitable for estimating the probability associated to predicted metrics, indicators and reference points. *Aladym-z* has been developed as a specific tool, which starting from the observed values of *Z* and the description of the life and population traits is able to calculate values of total mortality which better approximates a given scenario.

The model is designed to simulate population dynamics of a given population accounting for differences by sex in growth, maturity and mortality. All the quantities are calculated as vectors with an associated time step Δt (time slice=1 month).

The population dynamics is formulated following the simultaneous evolution of several cohorts at month scale through the exponential population decline model, both in absence (1) and in presence (2) of fishing mortality:

$$\frac{dN}{dt} = -MN \qquad (1)$$

$$\frac{dN}{dt} = -ZN \qquad (2)$$

used respectively in the form (3) and (4):

$$N_{(t+\Delta t),j} = N_{t,j} e^{-M_{t,j} * \Delta t} \qquad (3)$$

$$N_{(t+\Delta t),j} = N_{t,j} e^{-(F_{t,j}+M_{t,j}) * \Delta t} \qquad (4)$$

where j indicates the cohort, t the time, Z, M and F the total, natural and fishing mortality respectively.

Initial numbers in the population are from estimates of recruitment independently obtained (e.g. from trawl surveys). The number of recruits entering the population in successive years can be a vector or is estimated from a stock-recruitment relationship (Beverton & Holt, 1957; Ricker, 1975; Shepherd, 1982; Barrowman & Myers, 2000), with random variations. The number of the events (on monthly basis) generating the offsprings is an input of the model.

The growth process is assumed according to a VBGF and a length-weight relationship; the maturity follows an ogive model.

The natural mortality can be constant for each age/length or a vector by age/length calculated outside the model. Alternatively, it is estimated inside the model from the Chen and Watanabe equations (1989).

The fishing mortality rate $F(L)$ is modelled for each cohort using the following general equation (Sparre and Venema, 1998):

$$F(L) = F_{max} \cdot S(\overline{L})$$

[6] Contributed by Maria Teresa Spedicato, COISPA, Bari, Italy

where F_{max} is the maximum fishing mortality and $S(\overline{L})$ the proportion of retained fish.
In *Aladym* the fishing mortality rate is calculated as follows:

$$F(L) = F_{max} \cdot S(\overline{L}) \cdot f_{act}$$

where maximum fishing mortality (F_{max}) is calculated as follows:

$$F_{max} = QZ_{input} - M_{\min}$$

using the input values of QZ (a Z proxy) and where M_{min} represents the minimum value that the M vector assumes. In addition, a fishing activity coefficient (f_{act}) is introduced in order to consider the possibility of a fishing ban or changes in fishing effort throughout time.
The value of QZ by sex can be assumed, as a first order approximation, numerically equal to the value of Z observed that is obtained from estimations outside the simulation model (e.g. from trawl-survey). A better approximation of QZ is obtained using the tool *Aladym-z*.

- Inputs

- von Bertalanffy growth parameters by sex with associated variability,
- length-weight relationship parameters by sex;
- maturity ogive parameters by sex ($L_{m50\%}$ and $L_{m25\%}$-$L_{m75\%}$ range);
- natural mortality by sex (a constant value or a vector);
- seed values (minimum, maximum, *ln*-mean and *ln*-standard deviation) of recruitment by sex;
- proportion of offsprings entering in the stock by month;
- stock-recruitment relationship parameters or a vector of recruit numbers by month both with associated variability;
- time elapsing from spawning to birth;
- sex-ratio (female/total) of offsprings;
- F_{max} by month or from the model;
- *QZ* by sex;
- selection ogive parameters (2 options) of the gear used by the fleet ($L_{50\%}$ and $L_{25\%}$-$L_{75\%}$ range, $D_{50\%}$ in case of the selectivity option 2);
- fishing activity coefficient by month (0, in case of absence of fishing activity).

In *Aladym-q* the following inputs are also provided:

- the number of realizations;
- the parameters of the defined *pdfs*.

- Harvest control rules

The simulation approach can be used as a tool to convert survey biological information and relative assessment into quantitative HCRs. The options implemented in the simulation model are based on the following aspects: total mortality, gear selectivity (size at first capture $L_{50\%}$ and selection range) and fishing activity (alone or in combination). These three are inputs that can be used to simulate different exploitation scenarios. The effects of HCRs (selectivity and fishing activity) are then analysed in terms of sustainability for the population in the long-term. For example, the ratio between the mean spawning stock biomass and the mean unexploited spawning stock biomass (*SSB/USSB*, output) is also estimated for each harvesting scenario.
A vector of yield (Y) by time is also simulated, estimating the catch (C) according to the following general equation (Gulland, 1969):

$$C_{\Delta t} = \int_0^{\Delta t} F \cdot N_0 \cdot e^{-Z \cdot \tau} d\tau = \frac{F}{Z} N_0 \cdot \left(1 - e^{-Z \cdot \Delta t}\right)$$

where Δt is the time to which the catch is referred.
Thus the catch (Yield) in the time interval (t, t+Δt) is computed in *Aladym* as (Sparre and Venema, 1998):

$$Y_{t,j} = \frac{F_{t,j}}{Z_{t,j}} \cdot N_{t,j} \cdot (1 - e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t}) \cdot \mathrm{w}_{\mathrm{age}}$$

- Assumptions and sensitivity

The basic assumptions of the model are:

a) natural mortality as estimated reflects the rate of decline of a population for all causes excluding fishing;
b) total mortality *Z* reliably reflects the decline of ages/sizes in the population, including the effects of different fishing gears;
c) the growth, the natural mortality, and the maturity parameters are assumed constant along the time;
d) given the small time interval (1 month) between cohorts the effect of the spreading of the lengths respect to the ages can be neglected.

The model behaviour is influenced by the consistency between the set of life-history parameters and population dynamics. The model results are thus expected to be particularly sensitive to the stock-recruitment relationship and natural mortality.

To summarise, the main features of the F-I methods reviewed are presented in Tables 1.a,b below.

Table 1. Categorisation of methods

a. Technical

Method	Approach	Estimation	Structure	Input*	Need M?	Output
BREM	Biomass, Random effects	Max-Lik	Two-stage (Rec /Tot), in mass	Indices by stage; 1 or 2 fleets	N	Relative R & Btot
LENSUR	Length, Lagrangian	NLLS	Length	Indices by L. class; 1 fleet; v.B. params		Annual F, relative N@L
SURBA	Separable Z	Weighted NLLS	Age	Age-disagg. indices; n fleets	N	Z, relative R, TSB, SSB & N@age; conf. limits on R & Z
TSA	State-space, Time Series	Kalman filter, Max-Lik	Age	Age-disagg. indices; 1 fleet	N, but can be used (by age & year if avail.)	Relative N@age & Z; Not B
YCC	Y-Class curve	GLS	Age	Age-disagg. indices; n fleets	N	Z, relative R, rel. Bt & SSB, rel. fleets' q, rel. precision of fleets, predicted indices
ALADYM	Simulation	-	Age-Length	v.B. params, Z, selec		multiple

* only those essential for fitting; not for derived quantities such as Btot or SSB

b. Management measures that can be informed

Method	TAC	Effort	Gear/Mesh	Time closure	Other
ALADYM	(y)	Y	Y	Y	
BREM	Y	Y	N	N	
LENSUR	Y	Y	Y		
SURBA	Y	Y	N (unless *M* is known)	N (unless multiple surveys from ≠ times of year)	
TSA	Y	Y			
YCC	Y	Y			

2. Testing procedure

2.1. Data Sets

In absence of a better alternative at the time, we resorted to the suite of data sets concocted for the US National Research Council rounds of tests during 1997. One advantage is that the outcome has been published (NRC, 1998), enabling the performance of other methods to be compared with that of the methods considered by that committee (which all made use of catch and/or catch-at-age data). The data were generated by an age-structured model, where a 15-age population was projected over some 40 years but data for only the last 30 years were retained. Details of the data generation are given in Chapter 5 and Appendix E of the NRC report, and the main features are summarised in Table 2 below. Each data set is a single replication of a combination of stochastic processes[7]. A special comment applies to data set 3, which involves a change in survey vessel (and a near doubling of survey q), a feature that was not explicitly disclosed to the FISBOAT analysts initially and was a clear violation of a basic assumption in their method; however, given the knowledge of a step change in q, all methods are able to deal with this situation and most authors repeated the analysis with each period treated as a distinct survey (run labelled "set 3.2" hereafter), which resulted in improved performance. Also note that data set 5 simulates a case with very low exploitation rate (Yield/Biomass ratio in Table 2).

Table 2. Specifications of the simulated data sets (expanded from NRC 1998).

Set	Population trend	Age at 50% selectivity	Misreporting	Survey q	CV survey q	M	Mean Y/B
1	Depletion	Lower later	0.97-1.03	Constant	0.3	0.18-0.27	0.19
2	Depletion	Lower later	0.68-0.72	Constant	0.3	0.18-0.27	0.12
3	Depletion	Lower later	0.97-1.03	Higher later	0.3	0.18-0.27	0.12
4	Depletion	Constant	0.97-1.03	Constant	0.3	0.18-0.27	0.21
5	Recovery	Constant	0.97-1.03	Constant	0.3	0.18-0.27	0.07
6	2-way trip	Constant	0	Constant	0 (clean set)	0.2	0.15

Since some NRC sets are rather tough, a "clean" set (labelled # 6) was added where survey q has been strictly constant, and indices at age measured without error. This was also generated with an age-structured model comprising 15 age groups, and twenty years of data were output. Methods that break down on this easy set would clearly require some hard work.

The data sets were circulated to methods' authors in advance of a project workshop. The main information that was provided is the matrix of survey indices by age and year. Weights at age, natural mortality (average for the NRC sets, where M varied randomly) and maturity ogive were also provided, in case some methods would need these data, but no information about catches and effort by the fishery was given. It was proposed that analysts focus on the following outputs for comparisons: time series of recruitment (preferably in number); time series of total biomass and, if possible, of total numbers; optionally, time series of SSB.

Clearly these data were not adequate to test length-structured models, such as LENSUR, for which specific test data have to be set up (preferably providing true states, i.e. not on real stocks). The testing framework was also inadequate to evaluate the ALADYM simulation model.

[7] The report of the 2007 Methods WG (ICES CM 2007/RMC:04, Section 2.1.2) may leave the impression that the test data were not corrupted with noise. We point out that the NRC sets 1-5 did include various elements of noise, with perhaps the most relevant for this test being a random logNormal error on the survey indices at age with a 30% CV. Only set 6 was 'clean'.

2.2. Performance metrics

The intention behind selecting the NRC test sets was that comparisons might be possible with the performance achieved by catch-based assessment methods as documented in the NRC report. Since the latter methods are deemed to provide absolute estimates of key management variables, the NRC Committee chose to evaluate the methods based on relative error statistics (i.e. [(estimated – true)/true], both estimates and truth being in absolute value). For F-I methods, however, a clear message from all authors is that these could only provide estimates of relative trends in population variables, and thus the statistics above could not be used. Alternatively, the following approach to a performance metric involving relative values was considered: for each quantity of interest, the time series of estimates, on the one hand, and of true values, on the other hand, are first normalised by subtracting the respective mean and dividing by their SE (years with NA estimates, which are specific to each method, are excluded from both series before computing mean and SE), which gives a common scaling; the mean over years (rather than the sum, to account for NA-related differences in time series' length among methods) of the squared deviations between normalised estimates and normalised truth is computed; the square root of that mean is taken as the summary statistic (kind of RMSE). Although this statistic is not readily interpretable to gauge the performance against standard criteria, it enables fair comparisons between the F-I methods (unfortunately the results of catch-based methods are only shown graphically and not tabulated in the NRC report, otherwise the same statistic could have been computed and both classes of methods compared on equal footing).

The biomass depletion rate, that is the estimate of biomass in the final year divided by that in the first year, as considered in the NRC tests should in principle be the same when based on absolute or relative estimates and was also retained as an indicator for comparisons (for those F-I methods yielding biomass estimates), together with the NRC mild criterion that the relative error compared to the true rate should be within $\pm 25\%$.

As a further aid to compare methods, the estimation CVs for recruitment and biomass (when the method is able to provide them) obtained for each data set were also tabulated.

3. Results of methods comparisons across sets

The relative performance of the F-I methods tested is summarised in Tables 3.a-e for each of the performance metrics described above. Graphical comparisons of the trajectories of estimates vs. the truth (both normalised) are also shown to gain more detailed insight into the behaviour of each method (Figures 1-4).

The first thing to note is that most methods did very well with the clean set #6 (only YCC showed some inconsequential deviations for recruitment estimates), which is reassuring: this validation test indicates that there is no inherent defect in the rationale of these methods, nor in the computer code.

These methods essentially behave as smoothers for noisy indices, and may miss quick transient changes in stock abundance. However, in their expected usage to evaluate "current" stock state by comparing present and historic estimates, none would have caused managers to be misled about the situation of the stock and actions to take in the last decade of the time series. For recruitment, the position of weak or strong year-classes is generally correct, although there are cases of either over-smoothing or over-reaction to the signal in the survey.

Like most VPA tuning methods, these F-I methods make the strong assumption that survey q (by age or stage) is constant over time, and it should not come as a surprise that estimates were badly biased in the tests with set 3.1 where the large step change in q was ignored. In normal circumstances, the assessors would be aware of such marked changes in the survey procedure and would adjust the treatment of their data accordingly, as exemplified by the runs redone as 3.2. Nevertheless, this test highlights the fact that F-I methods are strongly dependent on the quality of the survey, notably the consistency of the survey protocol, as they use no other source of information which might counterbalance poor survey data. In actual life, year-on-year variations in survey design (e.g. due to weather or logistic constraints) or gear rigging are common, and users of F-I methods should be alert that they must take them into account, however benign they may appear at first sight.

In contrast, the test indicates no particular problem with set 5, a case with very low exploitation rate (F << M) which may cause poor convergence of VPA based methods.

Overall, based on inspection of summary statistics and patterns in the plots, all the methods tried in this test perform quite similarly and could be used interchangeably, depending on availability and familiarity with the software. There is a small practical advantage in favour of *BREM* which does not require extensive age compositions. Moreover, *TSA* does not (yet) provide biomass trajectories, and the plots of *SURBA* estimates show occasional wiggliness in some batches of years.

As said earlier, it is not straightforward to compare the performance of the F-I methods with those of the tuned catch-based methods applied to the same data in the NRC tests, since estimates from the latter are not available in tabular form. Coarse comparisons with the biomass trajectories plotted in Appendix I of NRC (1998) indicate that catch-based methods tended to consistently over- or (most often) under-estimate relative to the truth, whereas F-I estimates wander about the true trajectory. Note in passing that with set 5, all catch-based methods under-estimated the true absolute biomass by a considerable amount, but may have preserved the relative trend. More direct, albeit not necessarily easier, comparisons can be made with the estimates of depletion rate for those NRC runs where only the survey data (not the commercial CPUE series not considered here) were used for tuning. F-I methods, notably *BREM*, perform comparatively well and were generally outperformed only by the most highly parameterised catch-based methods.

It must be kept in mind that this evaluation is contingent on, among other things, scenarios where the error in observation of the indices has a CV of 30%, a value which is considered reasonable for well-behaved surveys. If in reality these methods are applied to survey data with larger errors, across the series or in specific years, their reliability in advisory contexts will obviously be poorer.

There is also the limitation that this test is based on a single replication of a stochastic data generation, and that a proper evaluation would require summarising over many replicates – this is the task of another work package in the project. We note, however, that our protocol is the same as the one adopted by an eminent scientific committee.

Table 3. Performance statistics

a. RMS of normalised deviations for Recruits

Method \ Set	1	2	3.1*	3.2*	4	5	Clean
BREM	0.559	0.435	0.775	0.744	0.540	0.548	0.001
SURBA	0.481	0.466	0.752	0.725	0.462	0.495	0.121
TSA	0.556	0.441	0.747		0.486	0.536	0.039
YCC	0.504	0.781	0.621	0.542	0.722	0.461	0.361

* 3.1: set 3 assuming a single consistent survey; 3.2: survey split in two (before/after change in vessel).

b. RMS of normalised deviations for Biomass

Method \ Set	1	2	3.1	3.2	4	5	Clean
BREM	0.207	0.211	0.805	0.524	0.194	0.197	0.012
SURBA	0.402	0.500	0.930	0.892	0.434	0.564	0.031
TSA							
YCC	0.182	0.187	0.869	0.347	0.135	0.146	0.152

c. CV (in %) on Recruits estimates (average over years)

Method \ Set	1	2	3.1	3.2	4	5	Clean
BREM				62.3			
SURBA	18.7	21.7	22.7	15.5	20.7	18.1	3.0
TSA	13.4	18.8	15.9		16.8	16.2	0.05
YCC	44.2	10.5	10.3	15.3	11.0	8.4	23.5

d. CV (in %) on Biomass estimates (average over years)

Method \ Set	1	2	3.1	3.2	4	5	Clean
BREM	46.3	54.6	39.9	69.2	12.1	37.0	14.4
SURBA							
TSA*	9.5	11.7	11.2		11.9	10.9	0.04
YCC							

* CV of GM stock number over ages

e. Relative error (in %) in Depletion rate (Biomass in final year / in year 1)

Results in boldface meet NRC ±25% criterion

Method \ Set	1	2	3.1	3.2	4	5	Clean
BREM	**-22.6**	**-3.9**	193.1	121.2	**15.7**	40.6	**1.0**
SURBA	-30.8	31.4	80.2	77.6	-39.8	**-20.0**	**-5.0**
TSA							
YCC	**-20.1**	42.5	137.9	**-3.5**	**2.0**	32.7	**0.3**

4. Conclusions

Although rudimentary, and awaiting further evaluation in full-fledged management strategy evaluation simulations, this exercise indicates that the F-I methods developed for this project are promising in terms of usefulness and reliability as bases for management advice.

Their main advantage, and indeed their raison-d'être, is that they are not subject to uncertainties in the commercial catches which have caused growing concern and controversies about scientific advice based on VPA approaches in recent years. Moreover, the dependence on catch data is the main reason for the current one-year delay between "data year" and "assessment year", which attracts criticism by managers that response from scientists to their requests is too slow. Clearly, survey-based methods can resolve this timeliness issue, as availability of updated information on stock state is a matter of days after a survey is completed (some overhead is still needed for data auditing, construction of the total area index when this involves more elaborate treatments than just aggregating samples, and mostly for age reading for those F-I methods requiring detailed age compositions). Another bonus with all the methods reviewed here is that their fitting procedures do not require prior knowledge of the natural mortality coefficient, which is a crucial ingredient in many other assessment methods and perhaps the most challenging parameter to estimate (M may still be needed for derived quantities, such as extracting F if management specifically needs it). Finally, it can be seen as an advantage that the methods reviewed have few if any "tuning knobs" to fiddle with.

Evidently, there are a few drawbacks. One is that there is no hope to estimate absolute stock size (overall or for specific ages): all abundance estimates are to be treated as relative, with an arbitrary scaling coefficient (= survey q) between actual and estimated abundance. In itself, this is not necessarily an issue, and examples might easily be found in many areas where decisions of utmost importance to society are made in reaction to relative indicators. The problem with fisheries management in Europe merely arises because, decades ago, scientists successfully sold the idea that they had the skills to deliver advice in absolute terms and the "system" has been built-up on these premises. One consequence is that managers were never educated to make use of alternative flows of information, such as relative indicators coupled with reference points based on past states (if only as a cross-check of the traditional advice), and more seriously that scientists have never formalised and evaluated an advisory process based on such information, although many critics argue that allegedly absolute VPA estimates are effectively relative since they are scaled by input M's which are guessed rather than known. However, this is mostly a problem with the advisory system and it should not count against the performance of the F-I methods *per se*.

A more inherent limitation of F-I methods is that they only use one source of information, and are thus critically dependent on the quality of survey protocols and data. Perceived year-on-year changes in abundance, and ensuing effects on advised management decisions, are likely to be very fragile to inconsistencies in the conduct of surveys (dates, geographical coverage, gear, etc.), and the best professional standards must be adhered to in order to reduce biases. When survey programmes are directed at groups of species (e.g. IBTS), the design tries to achieve a compromise between the needs of various species, and there are often populations whose distribution is only partially covered; this potential bias has to be borne in mind when candidate species are selected for application of F-I methods (and in any case when interpreting the results for advice). Finally, despite the complaint by paymasters that surveys are by far the most costly item in the assessment process, the implication of basing management on F-I approaches may well be that more, rather than less, investment in surveys is required notably for those where the precision of indices is near the limit of acceptability. Although gaps in survey data do not technically impede estimation with the methods reviewed, it is obvious that the quality of assessments degrades quickly when gaps occur frequently, and that the "current" state of stocks cannot be appraised in those years when data are missing. As a rule, surveys should be annual to be usable safely in the deplorably polemical context of fisheries management.

Acknowledgements

This work was made possible with the financial contribution of the EC to the institutes participating in the FISBOAT project (research project DG-Fish, STREP n° 502572 (2004-2007)).

Dr Terry Quinn (University of Alaska, Juneau, USA) kindly provided the NRC simulated data sets.

References

Barrowman, N.J. and Myers, R.A. (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. *Can. J. Fish Aquat. Sci.*, **57**: 665-676.

Beare, D.J., Needle, C.L., Burns, F. and Reid, D.G. (2005). Using survey data independently from commercial data in stock assessment: An example using haddock in ICES Division VIa. *ICES Journal of Marine Science* **62**: 996–1005.

Beverton, R.J.H. and Holt, S.J. (1957). On the dynamics of exploited fish populations. U.K. Min. Agric. Fish., *Fish. Invest*. (Ser. 2), **19**: 533 pp.

Chen, S. and Watanabe, S. (1989). Age Dependence of Natural Mortality Coefficient in Fish Population Dynamics. *Nippon Suisan Gakkaishi*, **55**(2): 205-208.

Cook, R.M. (1997). Stock trends in six North Sea stocks as revealed by an analysis of research vessel surveys. *ICES Journal of Marine Science* **54**: 924–933.

Cook, R.M. (2004). Estimation of the age-specific rate of natural mortality for Shetland sandeels. *ICES Journal of Marine Science* **61**: 159–164.

Cotter, A.J.R. (2001). Intercalibration of North Sea International Bottom Trawl Surveys by fitting year-class curves. *ICES Journal of Marine Science* **58**: 622-632 [Erratum, Ibid. 58:1340].

Cotter, A.J.R. and Buckland, S.T. (2004). Using the EM algorithm to weight data sets of unknown precision when modeling fish stocks. *Mathematical Biosciences* **190**: 1-7.

Cotter, A.J.R., Mesnil, B. and Piet, G. (2007) Estimating stock parameters from trawl cpue-at-age series using year-class curves. *ICES Journal of Marine Science* **63**: 234-247.

Darby, C.D. and Flatman, S. (1994) Virtual population analysis: version 3.1 (windows/DOS) user guide. CEFAS, Lowestoft, UK.

Deriso, R.B., Quinn, T.J.II and Neal, P.R. (1985). Catch-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences* **42**: 815–824.

Fryer, R.J. (2002). TSA: is it the way? Appendix D in Report of Working Group on Methods of Fish Stock Assessment. ICES CM 2002/D:01.

Fournier, D. (2005). An introduction to AD MODEL BUILDER version 7.0.1 for use in nonlinear modeling and statistics. Available from http://otter-rsch.com/admodel.htm.

Gudmundsson, G. (1986). Statistical considerations in the analysis of catch-at-age observations. *Journal du Conseil International pour l'Exploration de la Mer* **43**: 83–90.

Gudmundsson, G. (1994). Time series analysis of catch-at-age observations. *Applied Statistics* **43**: 117-126.

Gulland, J.A. (1969). Manuel des méthodes d'évaluation des stocks d'animaux aquatiques. Première partie- Analyse des populations. *Manuels FAO de science halieutique*, N. 4 : 160 pp.

Hilborn, R. and Walters, C.J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman and Hall, New York.

Johnson, S.J. and Quinn, T.J. II (1987). Length frequency analysis of sablefish in the Gulf of Alaska. *Technical Report UAJ-SFS-8714*, University of Alaska, School of Fisheries and Science, Juneau, Alaska. Contract report to Auke Bay National Laboratory.

Needle, C.L. (2002b). Preliminary analyses of survey indices for whiting in IV and VIId. Working Document WD2 to the ICES Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak, Copenhagen, June 2002.

Needle, C.L. (2002d). Survey-based assessments of whiting in VIa. Working Document WD1 to the ICES Working Group on the Assessment of Northern Shelf Demersal Stocks, Copenhagen, August–September 2002.

Needle, C.L. (2003d). Survey-based assessments with SURBA. Working Document to the ICES Working Group on Methods of Fish Stock Assessment, Copenhagen, 29 January – 5 February 2003.

Needle, C.L. (2004a). Absolute abundance estimates and other developments in SURBA. Working Document to the ICES Working Group on Methods of Fish Stock Assessment, IPIMAR, Lisbon 10–18 Feb 2004.

Needle, C.L. (2004b). Data simulation and testing of XSA, SURBA and TSA. Working Paper to the ICES Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak, Bergen, September 2004.

NRC, 1998. *Improving fish stock assessments*. National Academy Press, Washington, D.C., 177 p. (Appendix E describes the data generation; Appendix I shows plots of biomass trajectories).

Oehlert, G.W. (1992). A note on the delta method. *American Statistician* **46**: 27–29.

Patterson, K.R. and Melvin, G.D. (1996). Integrated Catch At Age Analysis Version 1:2. *Scottish Fisheries Research Report*. FRS: Aberdeen.

Pope, J.G. and Shepherd, J.G. (1982). A simple method for the consistent interpretation of catch-at-age data. *Journal du Conseil International pour l'Exploration de la Mer* **40**: 176–184.

Quinn, T.J.II and Deriso, R.B. (1999). *Quantitative Fish Dynamics*. Oxford University Press, Oxford.

Ricker, W.E. (1975). Computation and interpretation of biological statistics of fish population. *Bull. Fish. Res. Bd. Can.*, **191**: 382 pp.

Seber, G.A.F. (1982). *The Estimation of Animal Abundance*. 2nd edn, Griffin, London.

Shepherd, J.G. (1982). A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. *Journal du Conseil International pour l'Exploration de la Mer* **40** (1): 67-75.

Sparre, P. and Venema, S.C. (1998). Introduction to tropical fish stock assessment. Part 1, manual. *FAO Fish. Techn. Pap*., 306 (1) Rev. 2: 407 pp.

Thompson, W.F.and Bell, F.H. (1934). Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. *Rep. Int. Fish. (Pacific Halibut) Comm.* **8**: 49 p.

Trenkel, V.M. (2006). Combining acoustic and DEPM survey indices in the biomass random effects model for stock assessment. Working Document to the ICES Working Group on Acoustic and Egg Surveys for Sardine and Anchovy in ICES areas VIII and IX (WGACEGG), November 2006. 13 pp.

Trenkel, V.M. (2007). A biomass random effects model (BREM) for stock assessment using only survey data: application to Bay of Biscay anchovy. ICES CM 2007/O:03.

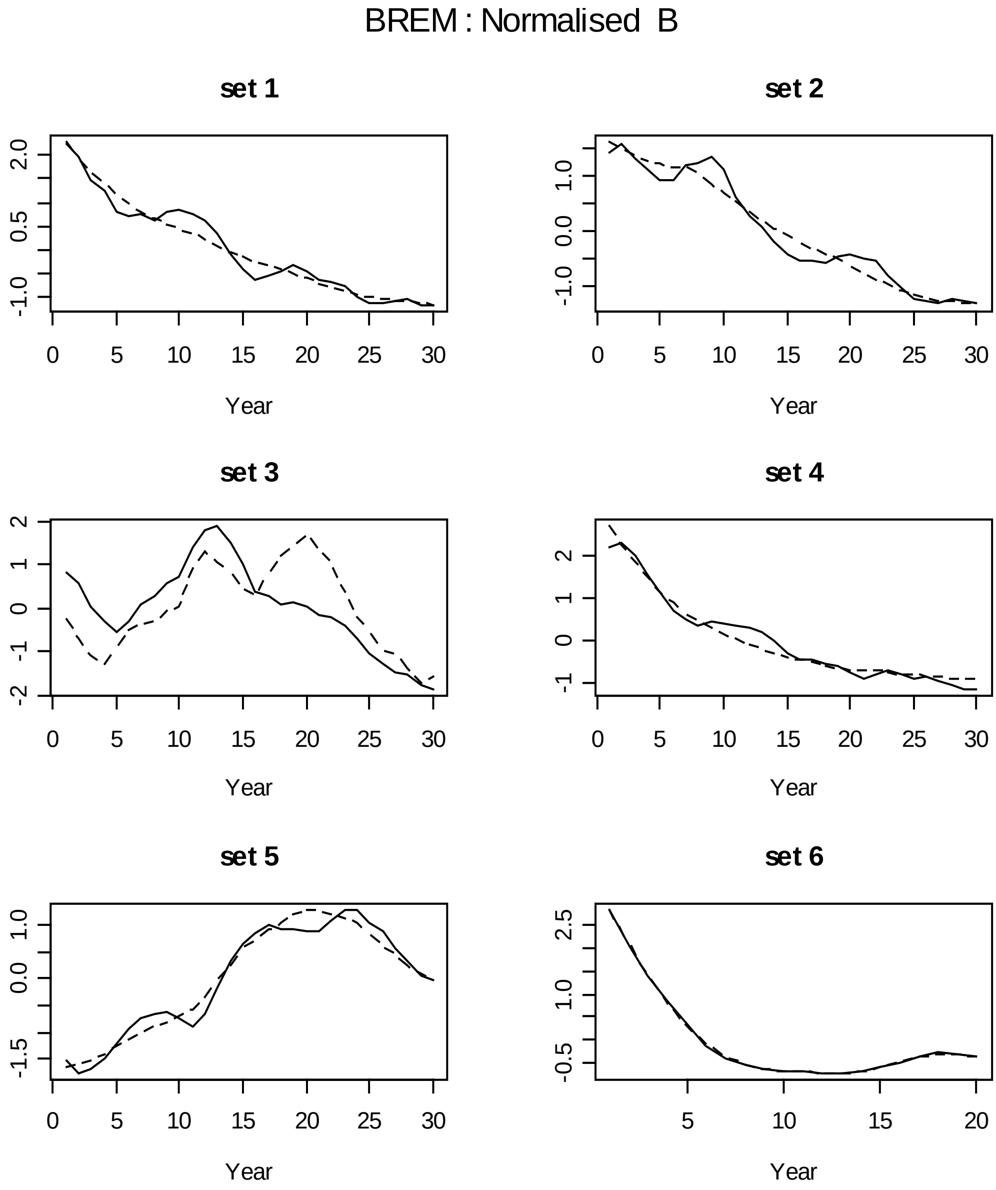


Fig. 1.a. BREM: Comparison of normalised series of biomass estimates (dashed) vs. truth (solid).

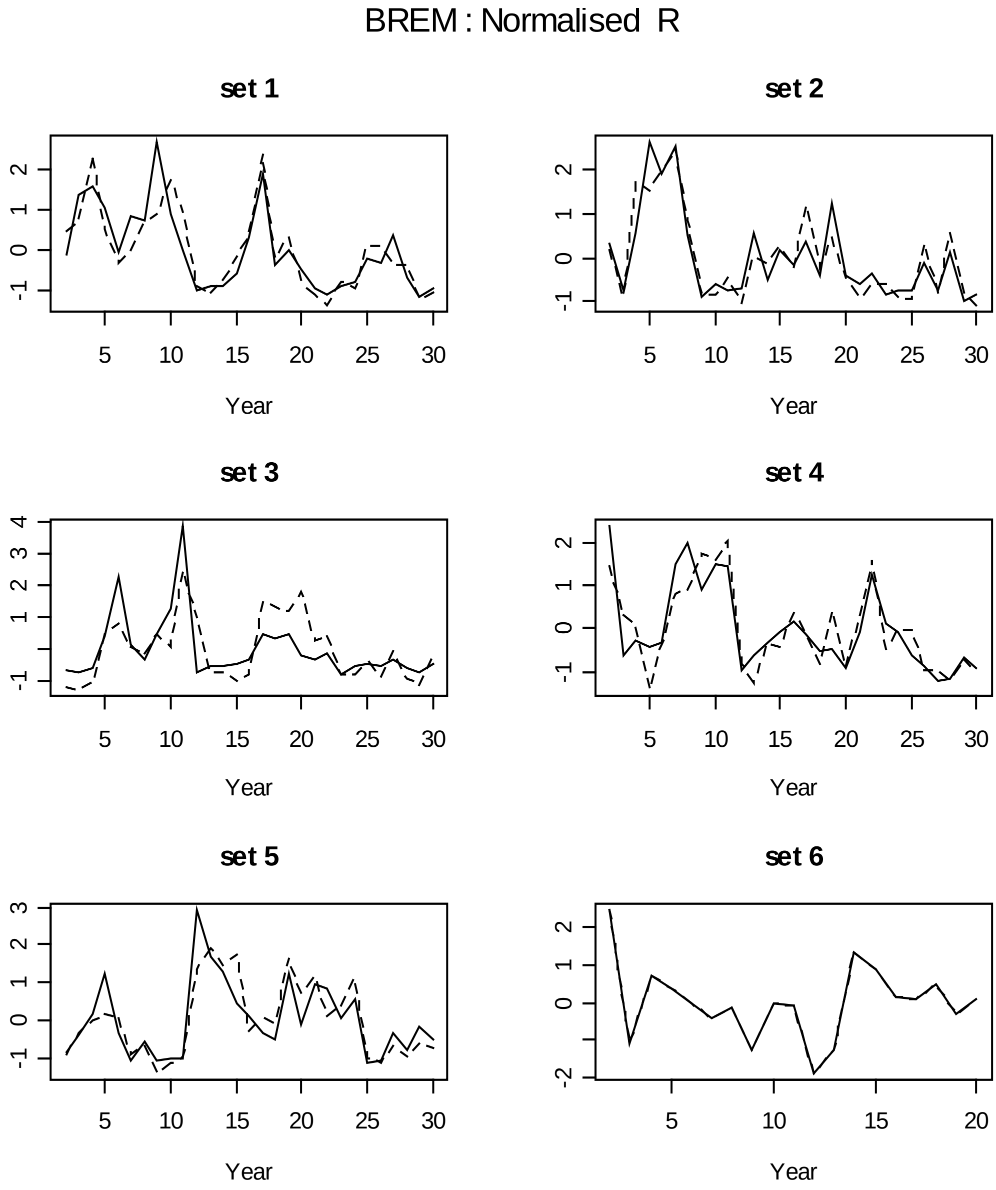


Fig. 1.b. BREM: Comparison of normalised series of recruitment estimates (dashed) vs. truth (solid).

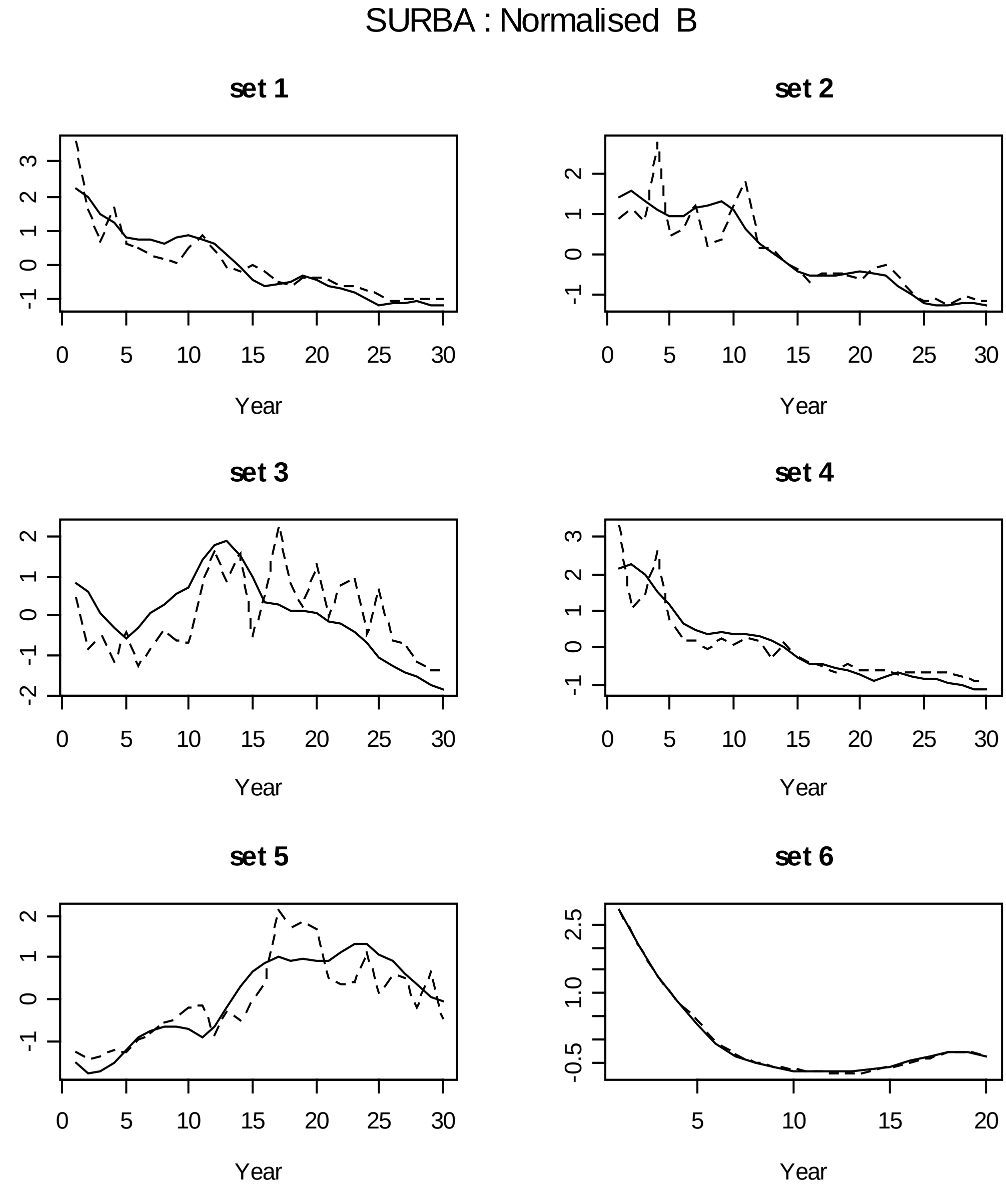


Fig. 2.a. SURBA: Comparison of normalised series of biomass estimates (dashed) vs. truth (solid).

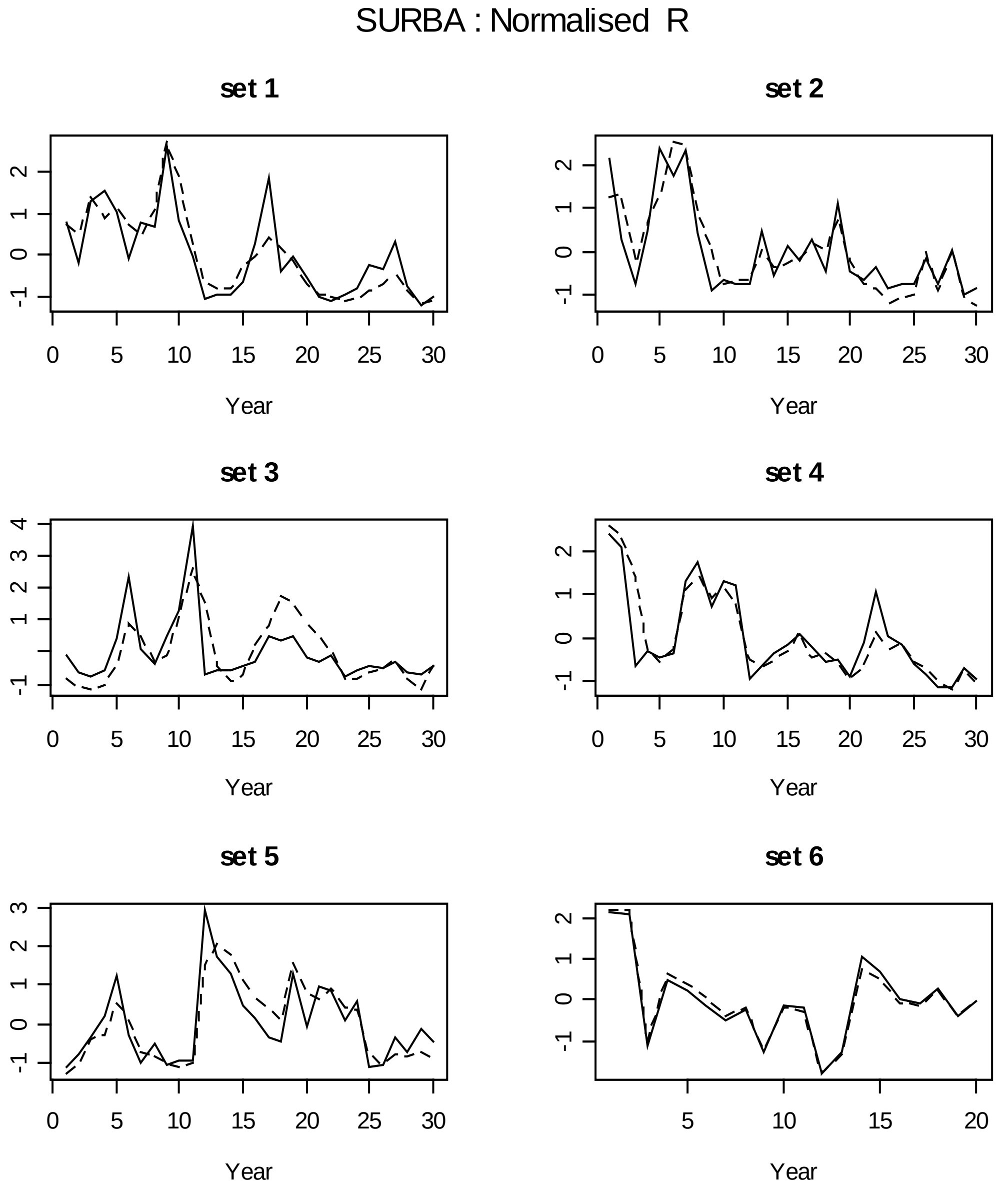


Fig. 2.b. SURBA: Comparison of normalised series of recruitment estimates (dashed) vs. truth (solid).

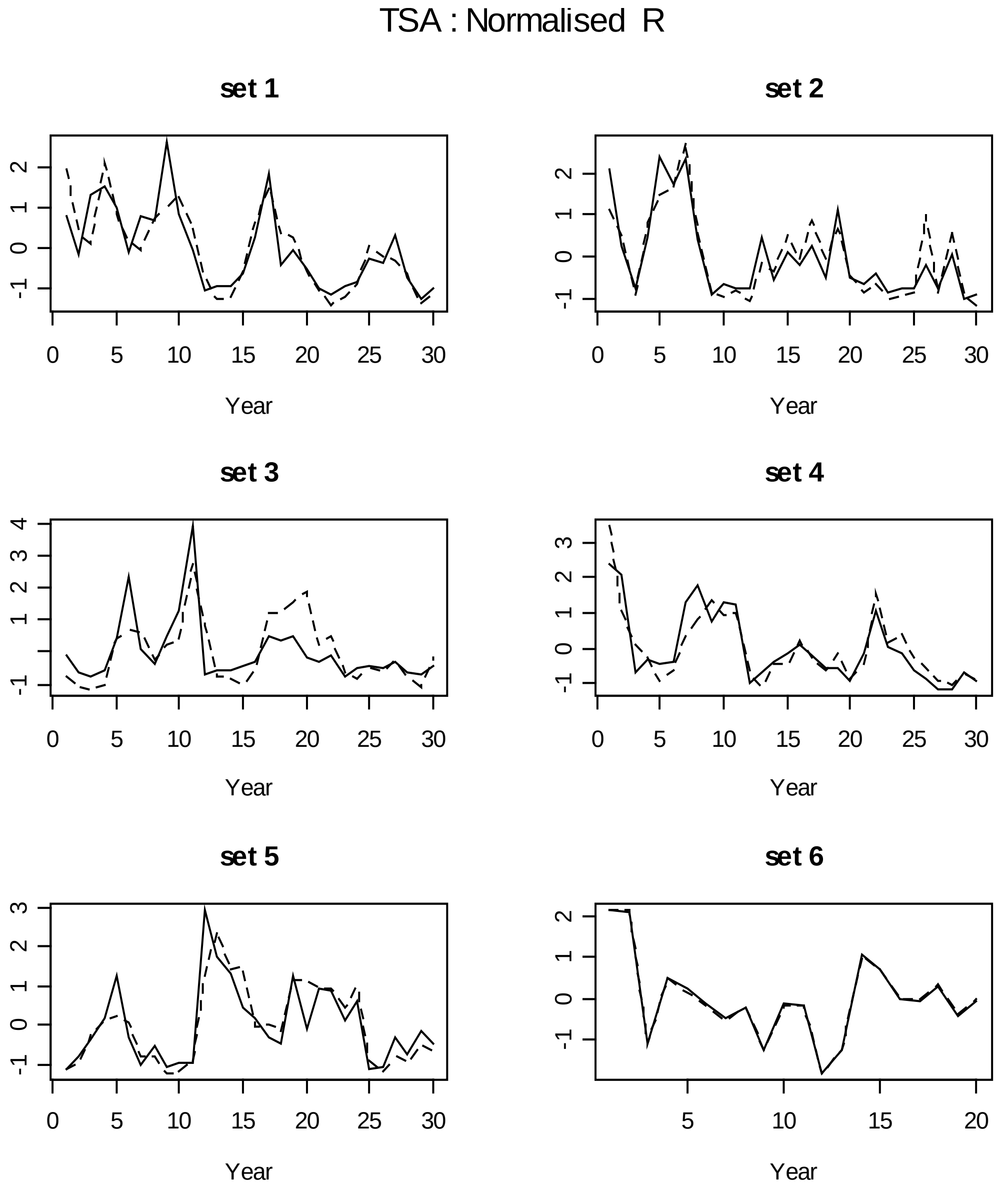


Fig. 3. TSA: Comparison of normalised series of recruitment estimates (dashed) vs. truth (solid).

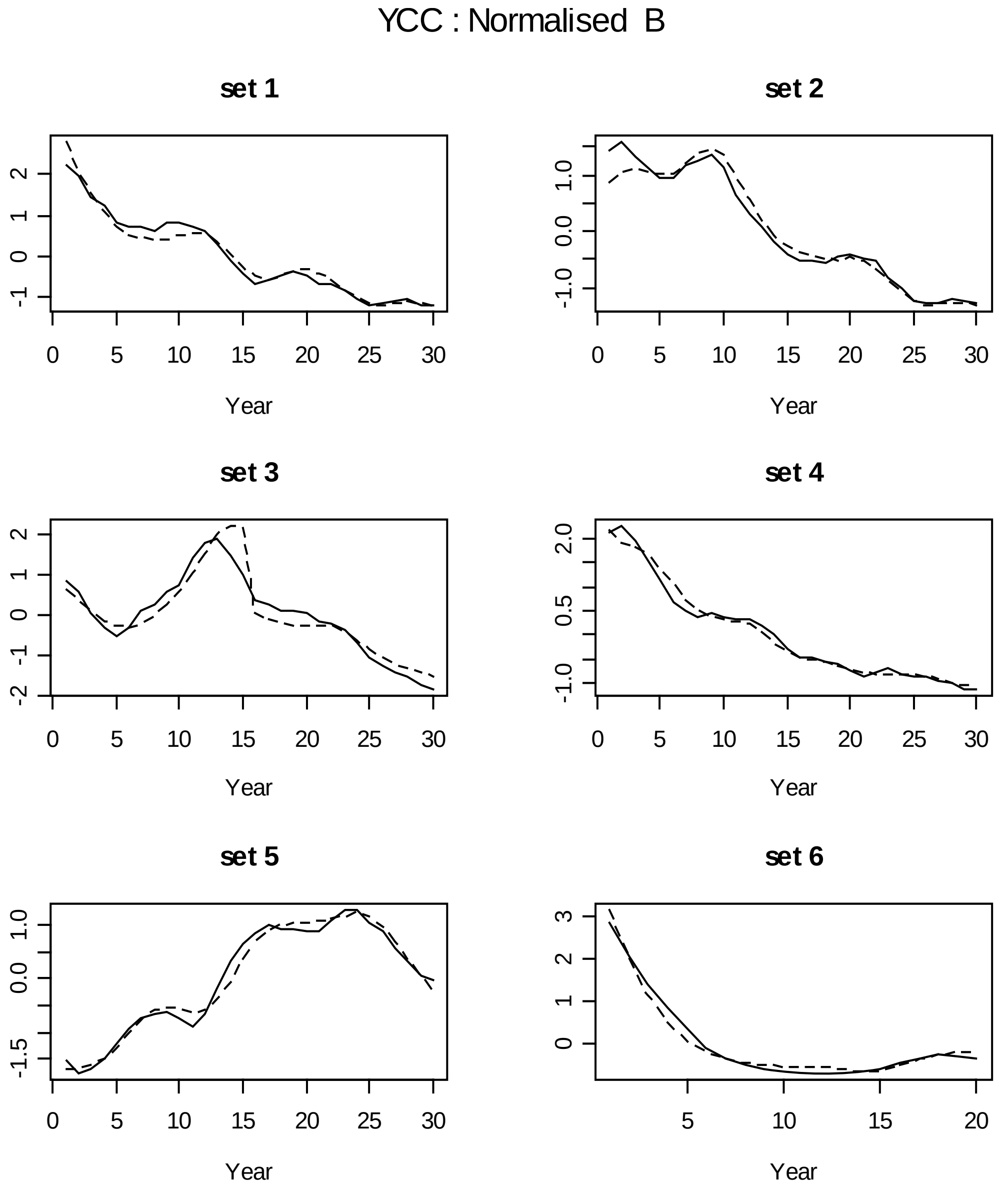


Fig. 4.a. YCC: Comparison of normalised series of biomass estimates (dashed) vs. truth (solid).
NB: set 3 = split survey (3.2)

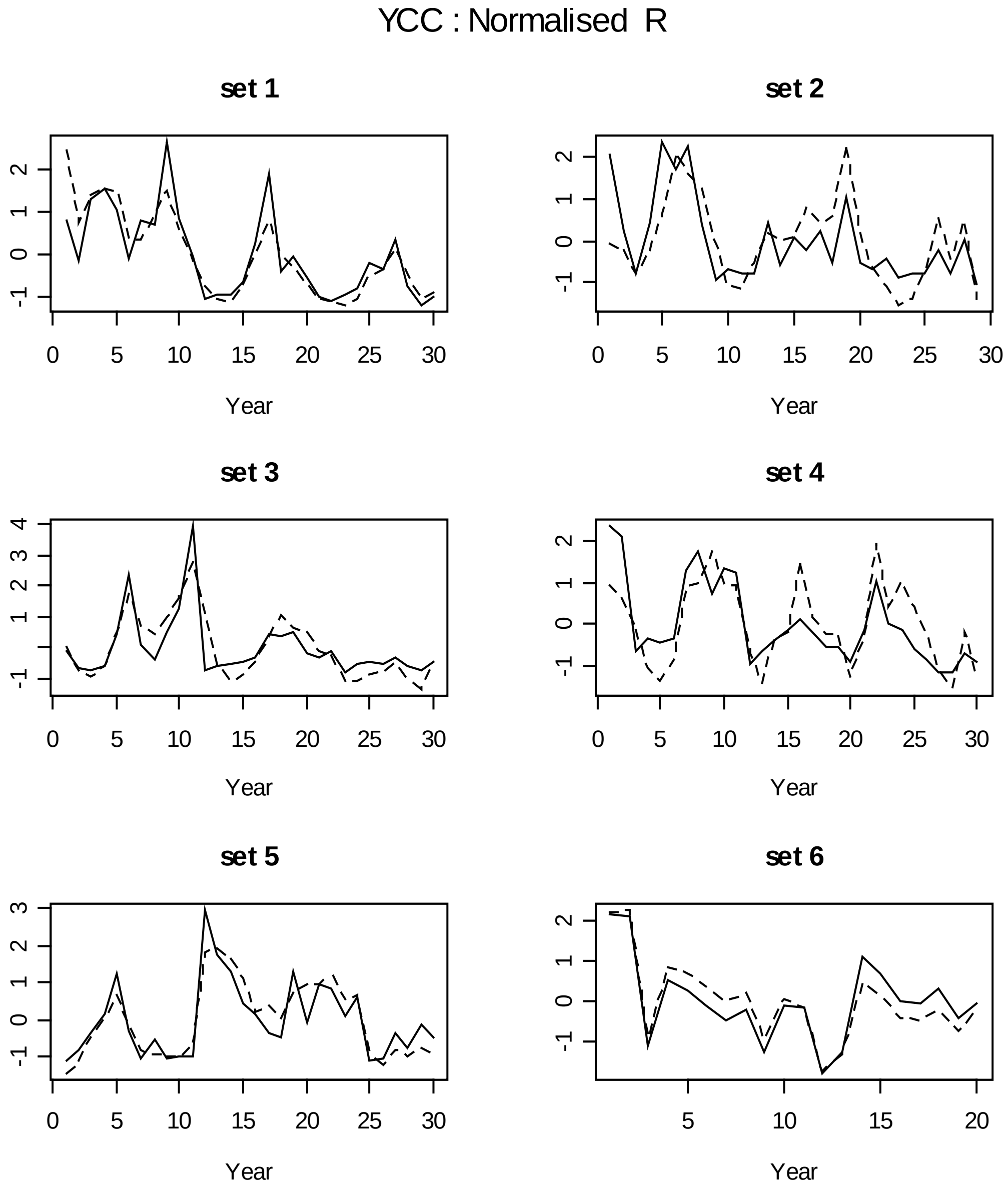


Fig. 4.b. YCC: Comparison of normalised series of recruitment estimates (dashed) vs. truth (solid).
NB: set 3 = split survey (3.2)

Document 4

FISBOAT Manual on Simulation Tools (FLR)

R. Hillary (Imperial College)

1 Introduction

This document is a generic manual on the specifics of both the relevant specifications of the observation error model, the biological operating model and the ideas behind the work done to demonstrate the coding of harvest control rules and visualising results as well as a gateway to the many courses, tutorials and case-study specific working examples of the simulation framework.

In terms of which deliverables this document covers, it covers the relevant elements of D2B.2, D4.1, D4.2, D4.3, D4.4 and D4.5.

The document is organised as follows:

1. The mathematical specifications of the biological operating model, along with details of worked examples of how to parameterise it with standard ICES WG data;
2. The mathematical specifications of the observation error model, and the details of the courses given and examples of how to use the observation error model given the biological operating model results and other information;
3. How one can define candidate harvest control rules based upon survey-specific data or stock assessment information, to then be used to define survey-based management strategies;
4. A review of the visualisation tools employed throughout the time of the project with particular attention given to dealing with probabilistic information and how to relate such information to relevant stakeholders.

The FISBOAT simulation software is all written in the FLR framework:
`http://www.flr-project.org/doku.php`.

This is an open source framework for fisheries modelling developed in the R language. The FLR framework is being used for a variety of EU projects and having such a common framework has both benefited the work in FISBOAT and *vice versa* for the other projects. All software is available on the FLR project's website and all of the tutorials and case-study examples have been run from this Wiki website so constant reference will be made to specific locations on this site.

2 The biological operating model

The FISBOAT project has a specific biological operating model, which differs substantially from the population dynamics model applied in the many VPA-type stock assessment models used in the ICES arena. This is, in part, due to the nature of some of the case-studies in the project (some lack catch-age data) and also so as to be able to generate a more sensible model of the case-study stocks' dynamics. Here are is the basis behind the biological OM:

- Time-steps are yearly and 'seasonal' - multiple within-year periods allowed;
- Harvesting model is single-fleet: defined by a supplied selectivity; effected via harvest rates, defined by the ratio of catch to total biomass;
- Three modes of stock-recruitment allowed: stochastic SRR relationship (B-H, Ricker, Hockey stick); stochastic (geometric) mean recruitment; bootstrap option, given a recruitment series.

2.1 Population dynamics

For the initial numbers, the model assumed that the population is at exploited/unexploited equilibrium, where the initial equilibrium harvest rates are defined by $\bar{h}_{a,s}$, at age a, in season s. If recruitment begins at a_r, in season s_r, then, for $s = 1, ..., s_r - 1$, $N_{1,a_r,s} = 0$, and $N_{1,a_r,s_r} = R_0$, where R_0 is a known parameter. For $s = s_r + 1, ..., S$, where S is the number of seasons:

$$N_{1,a_r,s} = N_{1,a_r,s-1}(1 - \bar{h}_{a_r,s-1})e^{-M_{1,a_r,s-1}}, \tag{1}$$

where $M_{y,a,s}$ is the natural mortality. For ages $a = a_r + 1, ..., A^+ - 1$, where A^+ is the plus-group, and $s = 1$ we have that

$$N_{1,a,1} = N_{1,a-1,S}(1 - \bar{h}_{a-1,S})e^{-M_{1,a-1,S}} \tag{2}$$

and for $a = A+$

$$N_{1,a,1} = N_{1,a-1,S}(1 - \bar{h}_{a-1,S})e^{-M_{1,a-1,S}} + N_{1,a,S}(1 - \bar{h}_{a,S})e^{-M_{1,a,S}}, \tag{3}$$

and finally, for $s = 2, ..., S$ and $a = a_r + 1, ..., A^+$

$$N_{1,a,s} = N_{1,a,s-1}(1 - \bar{h}_{a,s-1})e^{-M_{1,a,s-1}}. \tag{4}$$

This takes care of the initial population dynamics, and, for the remaining years $y = 2, ..., Y$ the dynamics are defined as follows:

For the seasons before recruitment, $N_{y,a_r,s} = 0$; we will explain the stock-recruit process in more detail later on; for the seasons following recruitment, the dynamics for $a = a_r$ are

$$N_{y,a_r,s} = N_{y,a_r,s-1}(1 - h_{y,a_r,s-1})e^{-M_{y,a_r,s-1}}. \tag{5}$$

For season one, the dynamics for $a = a_r + 1, ..., A^+ - 1$ are

$$N_{y,a,1} = N_{y-1,a-1,S}(1 - h_{y-1,a-1,S})e^{-M_{y-1,a-1,S}}, \tag{6}$$

and for the plus group:

$$N_{y,a,1} = N_{y-1,a-1,S}(1 - h_{y-1,a-1,S})e^{-M_{y-1,a-1,S}} + N_{y-1,a,S}(1 - h_{y-1,a,S})e^{-M_{y-1,a,S}}, \tag{7}$$

and for ages $a = a_r + 1, ..., A^+$, in seasons $s = 2, ..., S$, we have

$$N_{y,a,s} = N_{y,a,s-1}(1 - h_{y,a,s-1})e^{-M_{y,a,s-1}}. \tag{8}$$

The harvest rate, $h_{y,a,s}$ is defined to be

$$h_{y,a,s} = \kappa_{y,a,s} \times H_{y,s}, \tag{9}$$

where $\kappa_{a,s}$ is the selectivity function, and $H_{y,s}$ is the ratio of catch to total exploitable stock biomass. An assumption of separability can be made in the selectivity if required, by simply fixing it equal for all years. The SSB is defined in the standard manner, but with the inclusion of mortality before spawning in each of the seasons.

2.2 Stock-recruit behavior

If a stock recruit relationship is to be used, then the recruits are defined as follows:

$$N_{y,a_r,s_r} = \mathcal{F}(\boldsymbol{\theta}, SSB_{y-a_r,s_{\text{ssb}}}) \times e^{\zeta_y}. \tag{10}$$

In Eq. (10), $\mathcal{F}(\boldsymbol{\theta}, \cdot)$ is the particular stock-recruit function (Beverton-Holt, Ricker, hockey-stick), and $\boldsymbol{\theta}$ are the parameters. Notice that the delay used to relate spanners to recruits is implicitly defined by the minimum age in the model, and s_{ssb} is the spawning season. Note also, for reasons of common sense, that if $a_r = 0$, then recruitment **cannot** occur before spawning. The final term in Eq. (10) is the stochastic recruitment multiplier, defined as follows:

$$\zeta_y = \rho\zeta_{y-1} + \xi_y \tag{11}$$

where $\xi_y \sim N(0, \sigma_r)$ and ρ is the auto-correlation coefficient for the recruitment multipliers.

If fixed geometric mean recruitment, $\widehat{R}$, is requested, then

$$N_{y,a_r,s_r} = \widehat{R} \times e^{\zeta_y}, \tag{12}$$

but if a recruitment time-series is supplied, then the model simply resamples with replacement from this recruitment time-series to generate the model recruitments, but with no stochastic error term applied - with multiple simulations, the bootstrapping procedure should realise the recruitment uncertainty.

2.3 Parameterising the biological operating model

The main focus of the presentation given initially at the Pasaia meeting was to show how to parameterise the biological operating model using the sorts of data that come from the ICES stock assessment working groups: recruitment, ssb, fishing mortality and so on. Data such as weight-at-age, natural mortality, maturity and so on required by the model are all the same types of data available from the ICES working groups. For the biological operating model, we need the stock-recruit parameters (actual parameters, variance and structure or just mean recruitment and variance) and the selectivity pattern over time for the relevant fishery. The FLR framework allows one to estimate the stock-recruit parameters for a wide variety of stock-recruit models and there are many ways to estimate a selectivity pattern from fishing mortality-at-age. A complete example of how this can be done in practice can be found here:

`http://www.flr-project.org/doku.php?id=courses:fisboat_ijmuiden:old_hand:om`

2.4 Converting from age to length

The operating model simulates population indices-at-age, but there are some proposed harvest control rules and assessment methods that deal with length data. To this end, a conversion method was written into the FLFisboat package, called FLal and a guide to how to use the package can be found here:

`http://www.flr-project.org/doku.php?id=pkg:flfisboat`

Age-to-length conversion is the easier of the two 'directions' - for a given growth model, and whatever the estimation error model and variance, the expected length, $\bar{\ell}$, for a given age, a, and a given growth model, $\mathcal{G}(\cdot)$, is as follows:

$$\bar{\ell} = \mathcal{G}(a). \tag{13}$$

Given this expected length-at-age, we also must define a partition of lengths: $\ell_1, ..., \ell_N$, for the length-based return object. Given this length partition, the conversion process simply finds the bin containing $\bar{\ell}$, for the given age, and the value of the data by age, year etc. is added to the length entry defined by the *lower* value of this length bin. If, for some reason, $\bar{\ell}$ values fall outside the defined bins, then by default plus and minus groups are requested - a sensible choice of length bins, given the age range and growth model, should avoid the need for this option, and it can be switched off if required.

When converting from length to age, the fact that there are in fact a distribution of lengths - for any given age - must be taken into account, as we are, as yet, unaware of a perfectly estimated growth model. The methodology for this part of the conversion process is as follows:

We first assume that the inverse growth model, $\mathcal{G}^{-1}(\cdot)$, exists and can be calculated. For our given length, $\widehat{\ell}$, what we do is to compute the expected age, given the inverse growth and error model:

$$\bar{a} \;=\; \int_{\mathcal{L}} \mathcal{G}^{-1}(\ell)\pi(\ell \,|\, \widehat{\ell})d\ell \;=\; E^{\pi(\ell \,|\, \widehat{\ell})}\Big[\mathcal{G}^{-1}(\widehat{\ell})\Big], \tag{14}$$

and, for a normal error model with standard deviation σ, we have that

$$\pi(\ell \,|\, \widehat{\ell}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ell - \widehat{\ell})^2}{2\sigma^2}\right). \tag{15}$$

The integration step in Eq. (14) is numerically approximated. We allow for two error models: normal and lognormal. For the normal case, grid extrema are set-up at $\pm 2\sigma$ from $\widehat{\ell}$ (but if the lower is less than zero we set it to length zero) and an evenly spaced length grid of 100 points is defined between these two. For the lognormal case, we set these extrema at $\exp(\pm 2\sigma)$, so both encompass the 95 percentile of the error distribution being used. We then, after normalising, use this discrete approximation of $\pi(\ell \,|\, \widehat{\ell})$ to compute the value of $\bar{a}$. If, for example, $\widehat{\ell} > L_\infty$ for the von Bertalanffy model, then $\bar{a}$ is set at the maximum age, and, again, plus and minus groups act by default. The rounding up or down of the calculated value of $\bar{a}$ is as one would expect, in that if $\bar{a} - \mathsf{floor}(\bar{a}) < 0.5$, $\bar{a} = \mathsf{floor}(\bar{a})$; if not, then $\bar{a} = \mathsf{ceiling}(\bar{a})$. A simple example of how to generate length-based survey data can be found here:

http://www.flr-project.org/doku.php?id=courses:fisboat_aberdeen:thurs_morn

The biological operating model can be found in the FLFisboat package, which is a sub-package of the FLR framework that depends on the core FLR package, FLCore. The FLFisboat package can be found at the following location, along with relevant documentation and the required packages:

http://www.flr-project.org/doku.php?id=pkg:flfisboat

3 The observation error model

The second main part of the software framework was to develop an observation error model, which could simulate the known types of survey observations commonly available and used in ICES stock assessment and elsewhere. The observation error model takes data objects from the biological operating model, and together with information on catchability, observation error level and structure and so on and simulates observations that can be then used in the survey-based stock assessment methods and/or in developing survey-based harvest control rules.

The observation error package is also an FLR-based package, and is called FLOE. It is obviously of great use to non-FISBOAT EU projects that use the FLR framework but was developed by this project, and is a clear indication of how the FISBOAT project work is immediately being disseminated into the fisheries field as a whole. The FLOE package is split into two main parts:

1. FLObsIndex: the part of the package which simulates standard survey-type observations, such as the IBTS-type trawl surveys, acoustic surveys and recruitment surveys, under a wide range of observation error, catchability and bias regimes.

2. FLprop: this part of the package can be used in conjunction with observations from the FLObsIndex simulations or as a stand-alone method. Its purpose is to simulate correlated error in simulations (ageing error or year effects, for example).

3.1 The FLObsIndex suite of methods

The first thing to cover is the mathematical and statistical nature of the package. Speaking generically, we assume that any simulated index, $\widehat{I}$, can be expressed in the following way:

$$\widehat{I} = I^{\beta} \times q \times b\,(\times \,\text{or}\, +)\epsilon, \tag{16}$$

where I is the 'true' population variable being observed; q is the catchability, which in truth could be a composite of a number of factors; b is the bias in the observations; and ϵ is the error term, which could be additive or multiplicative. The β parameter is the index-to-abundance power coefficient. The error term should be defined by the particular observation error model being applied and/or assumed. This basic structure is sufficient to generate a very wide range observations - certainly enough to cover what is normally used as relative/absolute abundance tuning data. All the relevant parameters can vary by age, year and season, meaning that a very wide range of regimes can be simulated. Error structures permitted are normal (additive), log-normal and gamma (multiplicative), with a simple lag-1 correlation effect permitted in the normal and log-normal error regimes.

The FLR webpage for the FLOE package can be found here:

`http://www.flr-project.org/doku.php?id=pkg:floe:flobsindex`

Numerous tutorials and courses have been run that show in detail how the relevant methods of the FLObsIndex package can be used. The first of which was a course given in Aberdeen in May 2006:

`http://www.flr-project.org/doku.php?id=courses:fisboat_aberdeen`

and at the Management Strategy Evaluation course given in Ijmuiden, in the Netherlands in November 2006:

`http://www.flr-project.org/doku.php?id=courses:fisboat_ijmuiden:old_hand:surv`

3.2 The FLprop suite of methods

This part of the package was designed to simulate the more complex forms of error seen in fisheries data-at-age/length. These forms of error normally correlate across age or length - the prime example of course is ageing error - and are rarely, if ever, accounted for in all forms of stock assessment, be they fisheries dependent or independent. The idea is that, given proportions data, p, at the FLQuant (the basic data array object in the FLCore package) resolution, if we assumed that these data are multinomially distributed, then after applying the *logit* transformation to the data:

$$\widehat{p} = \ln\left(\frac{p}{1-p}\right), \tag{17}$$

the values $\widehat{p}$ will be normally distributed. We allow for the inclusion of structured noise (to simulate sampling/ageing error) but only along one dimension at a time. We use multi-variate normal noise to introduce the error, but only along either age, length or year dimensions. The noise can be structured with correlation at age or at year, for example, but not currently for both. The reason for this is that this type of noise represents a matrix, and matrix distributional theory is both extremely complicated and not as well developed as vector-based theory. Even allowing for this restriction, this should be more than enough freedom to explore many possibilities.

Given proportions data-at-age as an example $p_{y,a}$ (from catch-at-age information, for example), and assuming we can obtain a suitable estimate of the mean, $E^y(p_{y,a}) = \boldsymbol{\mu}$, and the variance-covariance matrix-at-age, $Var^y(p_{y,a}) = \Sigma$, we need to be able to translate this variance matrix into logit-space, as we simulate multi-variate normal noise in the logit-transformed proportions data in the simulations.

The delta method states that, given some differentiable transformation $\phi(p_{y,a})$, then the covariance matrix of this transformation, Σ^ϕ, can be expressed as follows:

$$\Sigma^\phi_{ij} = \frac{\partial\phi}{\partial p_i} \times \Sigma_{ij} \times \frac{\partial\phi}{\partial p_j}, \tag{18}$$

where the partial derivatives of $\phi()$, with respect to the proportions, are evaluated at $\boldsymbol{\mu}$.

Given the logit transformation:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right), \tag{19}$$

we have that the required adjustment to the known covariance matrix, Σ, is as follows:

$$\Sigma_{ij}^{\text{logit}} = \frac{1}{p_i(1-p_i)} \times \Sigma_{ij} \times \frac{1}{p_j(1-p_j)}, \tag{20}$$

where $p_i = \mu_i$.

The theory we use works as follows: we define a suitable adjusted covariance matrix, Σ^{logit}, which represents either the quant or the year dimension in the FLQuant of interest, and we generate the required number of simulations, assuming that

$$\widehat{p}_{\aleph} \sim MVN(\widehat{p}_{\aleph}, \Sigma^{\text{logit}}), \tag{21}$$

where $\aleph$ represents all the other quant dimensions not specified as the error covariate - in real terms, we add a multi-variate normal vector of mean zero and variance Σ^{logit} to every vector of age/length/year data in the FLQuant, depending on the covariate specified.

Once this has been done, we then apply the reverse logit transformation:

$$p = \frac{e^{\widehat{p}}}{1 + e^{\widehat{p}}}, \tag{22}$$

and re-normalise p to finally have our simulated proportions data. For a complete example of how to use this part of the package to simulate ageing error in catch-at-age data go to the following URL:

`http://www.flr-project.org/doku.php?id=courses:fisboat_ijmuiden:old_hand:agerr`

4 Assessment methods

FLR clearly offers a framework within which the assessment methods developed in WP3 can be incorporated, and in the next section we will see where these methods can fit into the management strategy evaluation process. Ongoing work to achieve the incorporation of these methods is happening and clearly offers the use of these methods to a wider audience of potential users.

5 Designing & defining harvest control rules

A key facet of the project is the designing of suitable harvest control rules (HCRs from now on) using survey-based data or assessment indices, as they are key elements of any survey-based management strategy. There can be no one single package that encompasses all possible HCRs - there are an infinite array of possibilities. By distributing an array of examples and the relevant documentation, what was done was to supply enough examples and the specifics of how they can be effected within the simulation framework so as to enable other project partners to adapt the examples as required.

To outline the mathematical considerations that are relevant we provide to examples of harvest control rules that can be derived from either survey data directly or from survey-based assessment methods.

5.1 HCR using relative SSB trends

The example harvest control rule we apply is one designed , and is based on year-to-year changes in abundance indices. The HCR essentially is a TAC adaption scheme which changes the TAC from year to year, relative to the TAC from the year before and the status of the stock, given the information in the abundance indices of interest. The HCR can be parameterised as follows:

$$\Delta_{y-1} = \frac{I_{y-1}}{I_{y-2}}, \tag{23}$$

$$TAC_{y+1} = \Delta_{y-1} \times TAC_y. \tag{24}$$

where the delay effect of time $y-2$ affecting the TAC in $y+1$ is present because of the delay that is intrinsic to the observable effect of changing the TAC on the stock, for this particular management system. Here, I_y represents the trend in spawning stock biomass (relative or absolute) and can either be from a survey directly, or have been estimated by a survey-based assessment method. The basic idea is that the TAC will increase if the spawning stock index is considered to be increasing, and will be decreased if the spawning stock index is decreasing.

5.2 HCR using trends in total mortality

A secondary option is how to define a HCR based upon trends in total mortality, **not** fishing mortality, as this is a potential output of both an age-based survey or age-based survey assessment method. The idea was to have two precautionary reference total mortality values, one on the juvenile portion of a stock, and another on the adult/mature portion of a stock, and to adapt the TAC relative to the trends seen in total mortality seen on both these elements of the stock. The basis for this was to try and mirror aspects of the actual decision rule applied to North Sea herring (a case study in the project), but based on total not fishing mortality levels on the juvenile and adult sections of the stock. Given our precautionary total mortality levels, $Z_{\mathrm{pa}}^{\mathrm{juv}}$ and $Z_{\mathrm{pa}}^{\mathrm{adu}}$, the HCR is defined as follows:

$$TAC_{y+1} = TAC_y \times \min\left(\frac{Z_{\mathrm{pa}}^{\mathrm{juv}}}{Z_{y-1}^{\mathrm{juv}}}, \frac{Z_{\mathrm{pa}}^{\mathrm{adu}}}{Z_{y-1}^{\mathrm{adu}}}\right). \tag{25}$$

The HCR defined in Eq. (25) is designed to be a precautionary "traffic light" kind of HCR. By this we mean that there will only be an increase in TAC if both levels of Z are below their precautionary levels; if either is above or below then the TAC will be reduced. If both are below their precautionary levels, then the smallest increase in TAC will be allowed. If both are above their precautionary levels then the largest decrease in TAC will be applied. The estimated values of Z can again be derived directly from surveys or are estimated from the survey data using an age-based survey-specific assessment method.

The practical implementation of both of these HCRs within the simulation framework can be found at the following URL:

`http://www.flr-project.org/doku.php?id=courses:fisboat_ijmuiden:old_hand:hcrs`

6 Visualisation tools

The final section of the manual relates to the visualisation and communication of the simulation results, to aid both the users and the intended stakeholder audiences to which these results will hopefully be presented.

6.1 Graphics within FLR

The default method of plotting FLR-based data objects is to use the lattice graphics package in R. All data objects in FLR are based on the FLQuant data object, which is a six dimensional array that permits the storage and modelling of complex data objects. As an example of what the lattice package can do, Figure 6.1 shows a lattice-plot of the fishing mortality-age for the North Sea herring Working Group estimates.

Using such plots is very useful to the user - particularly when doing exploratory data analysis and when parameterising the biological and observation error models. As an example, looking at trends in fishing mortality over time allow us to look for potential trends in selectivity over time, and whether we should be including such effects in our operating model.

In terms of plotting the results of management scenarios for the user to see the probabilistic dynamics of future key stock indices, under a given management scenario, the lattice package (using the bwplot() method) allows one to plot such quantities - see Figure 6.1.

6.2 Summary plots for stakeholders

Communication of results is obviously a vital part of any process that involves numerous stakeholders with different interests and levels of understanding. To this end, after consultation with the various members of the project as to what would be a useful type of summary plot, we designed such a plot. The idea was to have a plot that encompasses the relevant indices of interest for all stakeholder parties. For example, managers and

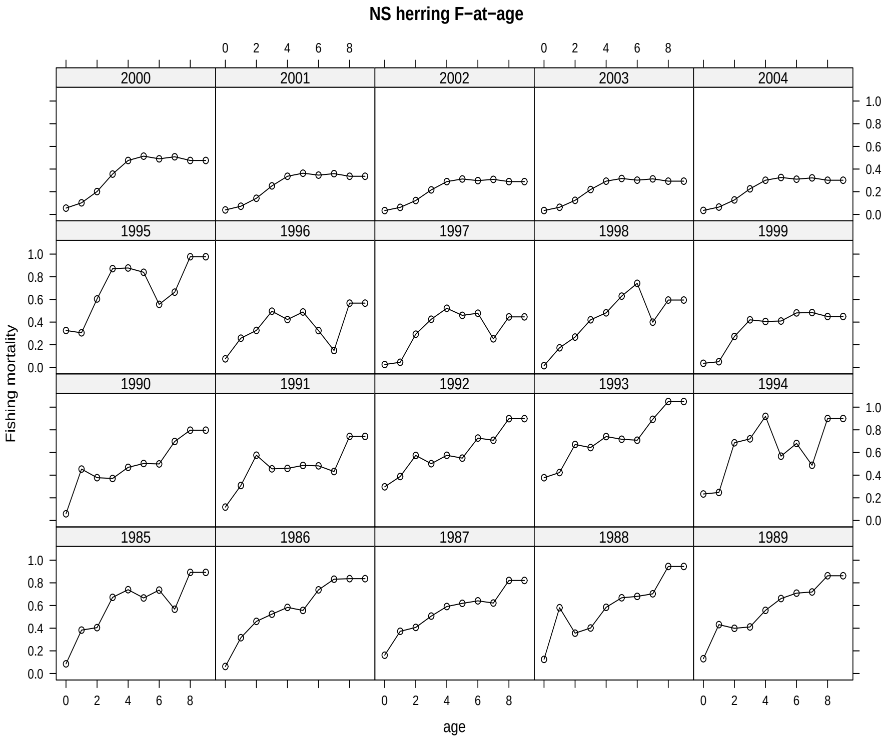


Figure 1: *Lattice plot of the estimated fishing mortality-at-age for the North Sea herring stock.*

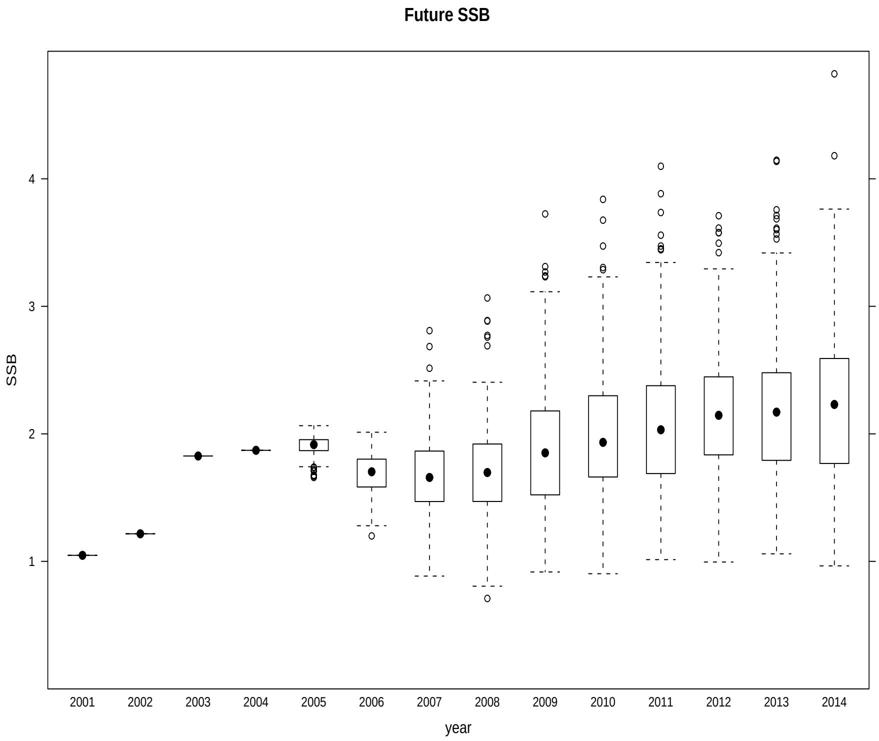


Figure 2: *Boxplot of the past and future SSB for the North Sea herring stock under a particular management scenario.*

commissioners will perhaps be more interested in the stock status indicators (SSB, average exploitation rate etc.) whereas fishermen will perhaps want to see projected catch and the inter-annual variation in the catch, as well as stock status. The result was a plot that is on a single page, but contains the following, for a specified time window:

- SSB dynamics;
- Recruitment dynamics;
- Mean (age-average) harvest rate - age range can be specified;
- Catch biomass;
- Inter-annual CV in the catch;
- Probability that SSB is less than a specified limit point.

Figure 6.2 shows an example plot for a particular management scenario run for the North Sea herring case study work.

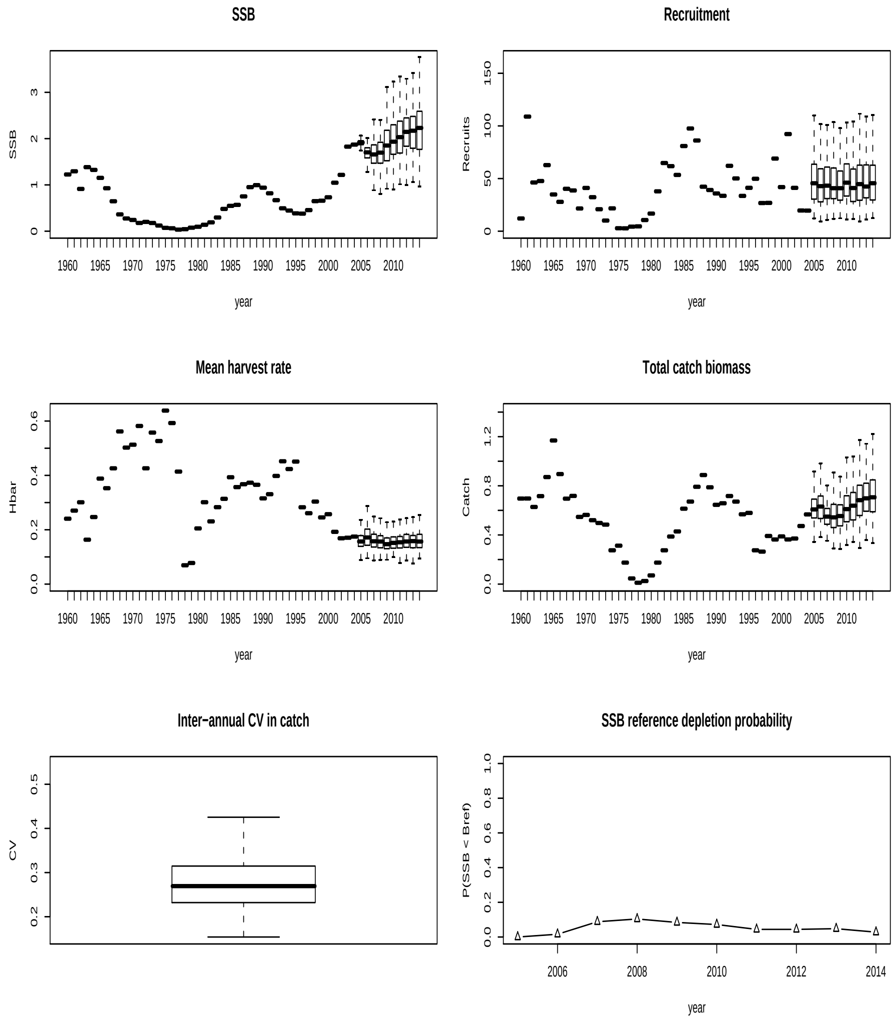


Figure 3: *The FISBOAT management scenario summary plot which is documented and a feature of the FLFisboat package.*

Document 5: Simualtion Evaluation with FLR - applications

Evaluation of the performance of survey-based assessment and fishery management approaches using FLR simulation evaluation case studies

Deliverable 5.3 for EC research project:

Fishery Independent Survey Based Operational Assessment Tools (FISBOAT),

DG-Fish, STREP n° 502572 (2004-2007)

Co-ordinator: P. Petitgas, IFREMER, Nantes, France

Authors:

Panayiota Apostolaki (CEFAS)
Leire Ibaibarriaga (AZTI)
Gjermund Bøthun (IMR)
Hans Bogaards (IMARES)
Marine Pomarede (IMP.COLL)

1. INTRODUCTION

The development of time-series of indicators from fish surveys and fishery-independent sources has been highlighted as an alternative option to using fishery dependent data to assess the status of a stock. Indicators that could be used for this purpose have already been presented in previous sections of this report. Although it is important to identify new types of data or uses of routinely collected data that can make stock assessments less dependent on certain types of information, such as fishery-dependent data, it is equally important to have a clear picture of the benefits and shortcomings of formulating management advice using these data. To this end, simulation evaluation methods can be very useful since, using simulated data (so allowing the level of error or bias to be controlled), we can evaluate the merits of different types of assessment approaches and management options. For this reason, a simulation evaluation framework was developed within the FISBOAT project and was used to evaluate the ability of survey-based assessment procedures to capture changes in population biology and test the sensitivity of the assessment procedure to uncertainties in survey estimates in order to explore alternative survey designs and guide the adoption management strategies.

The simulation-evaluation framework includes a population dynamics model, an observation model, an assessment model and a harvest rule model, which can be either model-based (using the results of assessment models) or model-free (based on the survey indices directly). A detailed description of all components of the FLR framework is presented in another FISBOAT document. Here, we present the results of the analysis conducted using this framework. The section that follows describes the case studies (reference species) chosen as representative ones for this part of the analysis. Evaluation of survey-based methods with respect to robustness, precision, capability to capture stock trends and data requirements, fishery yield, and stock risk requires the adoption of criteria that will determine the performance of each approach. Thus, the third section of this report provides the performance statistics that have been chosen for this purpose. The section after that presents a summary of the results of the analysis conducted under each case study. A general discussion of those results is provided in the last part of this document. Interested readers are encouraged to refer to supplementary material (individual case study reports annexed) for a detailed description of the parameterisation of the framework and the results of the analysis under each case study.

2. REFERENCE SPECIES

a. Herring

Herring (*Clupea harengus*) is one the most important commercial species taken in the North East Atlantic. While the fishery dates back at least to the middle ages, it expanded in the 19th century to respond to the need of industrialised cities. During the 20th century, the rapid development of the industrial fishing of herring led to a collapse in the 1970s with recovery made more difficult by juvenile bycatch by the sprat industry.

Following the two severe declines (up to 1977 and up to 1997), the North Sea herring became the first stock in the North Sea managed through the implementation of the precautionary approach. ICES classifies the stock *as "being at risk of having reduced reproductive capacity and at risk of being harvested unsustainably"*. The lower biomass reference point (B_{lim}), below which there is an aggravated risk of low recruitment, is set at 800 000 tonnes and triggers an emergency plan until the upper reference point – B_{pa}, set at 1 300 000 tonnes – is reached.

b. Bay of Biscay anchovy

The Bay of Biscay anchovy (*Engraulis encrasicolus*) is an important species for the Spanish and French fleets. Two direct surveys, Acoustics and Daily Egg Production Method (DEPM), are conducted in spring every year to assess the state of the stock. Based on these direct population estimates and on data from the commercial catches, the integral assessment of the stock is conducted by ICES in the Working Group on the assessment of Mackerel, Horse mackerel, Sardine and Anchovy (WGMHSA).

Currently the biological reference points for the stock, B_{lim} and B_{pa}, are set at 21 000 and 33 000 t respectively. Although there is no management plan developed, the stock has been traditionally managed by a fixed annual TAC (Total Allowable Catch) of 30 000 or 33 000 t.

Since 2002 the stock is at very low levels, being in 2005 the lowest of the historical series. After the failure of the fishery in spring 2005 the fishery has been closed successively for the second half of 2005 and 2006. In 2007 only experimental fishing with spatio-temporal restrictions has been allowed and the STECF has advised that any fishery reopening should not be considered until June 2008, when the results from the spring surveys become available.

As anchovy is a short lived species, the population is very dependent on the yearly incoming recruitment. Therefore, knowing the recruitment level beforehand can be very helpful for the development of any management plan. Currently, various juvenile surveys aiming at estimating recruitment and better understanding the recruitment process are being conducted but their results are not yet used for any management advice.

c. North East Artic Cod

The NEA cod is economically important for several countries, e. g. Norway and Russia. The stock is an important predator in the Barents Sea ecosystem. There have been observed large variation in growth rate, mean weight at age, maturity and degree of cannibalism. These fluctuations have been linked to water temperature, food supply and abundance of cod and capelin (Annon 2006). ICES describe the stock as overexploited in terms of fishing mortalities in relation to highest yield and agreed target. The SSB have been increasing since 2000 until 2004 and decreased in 2005. The catches have increased since 2000 until 2005.

d. North Sea Cod

Since the 1970s the stock of North Sea (NS) cod (*Gadus morhua*) has been decreasing. ICES classifies the stock as "being at risk of being harvested unsustainably". Since the late 1990s, several cod recovery plans have been adopted with the aim to increase the spawning stock biomass (SSB) of NS cod above the precautionary limit (B_{pa}) of 150×10^3 tonnes (t). However, stock assessment models have estimated a continuing decline since, SSB being well under the 70×10^3 t limit (B_{lim}) below which the stock is expected to suffer reduced reproductive capacity.
Although official catches (reported landings and estimated discards) are at an all-time low of around 35×10^3 t in the past years, surveys indicate that year classes are depleted faster than one would expect from these catches. This points to unaccounted removals, which are assumed to originate mostly from illegal fishing activities.

Management of NS cod traditionally rests on harvest control rules (HCRs) that target fishing mortality. While survey data are used to calibrate VPA-type assessment models, estimates of fishing mortality are still dominated by official catch figures. Consequently, estimated trends may be misleading whenever official catches are not representative of the true catches, a situation that readily applies to NS cod.

3. PERFORMANCE STATISTICS

In order to summarise the results from the simulations and evaluate the performance of the different harvest control rules, a series of performance statistics need to be defined and calculated. These performance statistics have to be related to the management objectives set for each of the stocks. Thus, in general terms, at least two groups of performance statistics can be distinguished: the ones related to the state of the stock and the ones related to the yield. In what follows we present a series of possible performance statistics within each of these groups:

a. Performance statistics related to the state of the stock
    - Probability that spawning stock biomass (SSB) is below some biomass reference point, B_{ref} , at least once in the series
    - Probability that SSB is below some biomass reference point, B_{ref} (all years and iterations)
    - Probability that SSB is below some biomass reference point, B_{ref} in the final year
    - Average or median number of years necessary to get SSB above B_{ref}

b. Performance statistics related to yield
    - Average or median catch over years and iterations
    - Standard deviation of the average catch or the average of the standard deviations
    - Average or median percentage of interannual change in catch and/or in total allowable catch (TAC)
    - Standard deviation of the percentage of interannual change in catch and/or TAC
    - Probability that actual catch is below TAC (i.e. that the TAC is above the exploitable biomass level)
    - Probability that the fishery is closed (i.e. that the TAC is zero)

4. SUMMARY OF RESULTS

Under each case study, the parameterisation of the framework aimed to simulate key characteristics of species dynamics and exploitation, and of survey data collection and was based on historical data (see individual case study reports annexed). In this context, simulation of sources of uncertainty in fishery-independent abundance indices considered the effects of errors in the calculation of those indices based on the information available for each of the reference species. The error term considered in the calculations for herring incorporated ageing error variability while a general error term was included in the simulations in the Bay of Biscay Anchovy, North East Artic Cod, and North Sea Cod reference cases. Different levels of uncertainty were used to test the robustness of the assessment results and management advice. The simulated abundance indices were either used directly as an input to a harvest control rule or were first processed using a stock assessment model. The stock assessment models and the general form of the harvest control rules considered have already been described in another FISBOAT document (FLR tools). The specific HCRs applied in each reference case are described before the description of the main results.

a. Herring

Harvest control rules

Four versions of the HCRs were used in the calculations for herring. The first model-free harvest control rule implemented is based on observation index:

$$TAC_{y+1} = TAC_y \left(\frac{I_y}{I_{y-1}} \right) \qquad (1)$$

The observation index, I_y, is the output of the observation error model and for this calculation, it tracks changes in SSB. The value of the TAC for the next year will depend on the TAC of the current year and the ratio of the observation index of the current year and the observation index of the previous year. If the ratio is larger than one this means there is an increase in the observation index which can be interpreted as an increase in the SSB of the stock so catches can be raised proportionally.

The second model-free harvest control rule implemented in this work is based on relative SSB trend from the acoustic survey data:

$$TAC_{y+1} = TAC_y \left(\frac{\overline{SSB}_y}{\overline{SSB}_{y-1}} \right). \qquad (2)$$

Relative SSB trends are calculated for all years from the beginning of the simulated survey (1990 in this study) until the current year of the simulation:

$$\overline{SSB}_y = \sum_{a=1}^{9}\left(I_{a,y}m_{a,y}w_{a,y}e^{-Mspwan_y M_{a,y}}\left(1-Hspawn_y H_{a,y}\right)\right) \quad (3)$$

The index used in the calculation is the output of the observation error model, the same as the one used in the first harvest control rule. Changes in the TAC of the subsequent year will be proportional to the ratio of the apparent relative SSB trend in the current year and the one of the preceding year.

The third harvest control rule is, again, a model-free one. For this harvest control rule we need an age non-aggregated index as 2 age groups (0 to 1 and 2 to 6) are distinguished in the calculation of the TAC:

$$TAC_{y+1} = TAC_y * \min\left(\frac{Z_{[0,1]}^{PA}}{Z_{y-1,[0,1]}}, \frac{Z_{[2,6]}^{PA}}{Z_{y-1,[2,6]}}\right). \quad (4)$$

The variation of the TAC of next year will be inversely proportional to changes in total mortality estimated along cohorts from the survey ($Z_{y-1,[0,1]}$ and $Z_{y-1,[2,6]}$ for age groups from 0 to 1 and from 2 to 6, respectively). Values of the total mortality rate at age at precautionary level (Z_{PA}) for each age group ($Z_{[0,1]}^{PA}$ and $Z_{[2,6]}^{PA}$) derive from other parameter values defined by ICES (i.e. B_{pa}, see section 2a). It is a harvest control rule that prioritises the precautionary approach as the variation in the TAC will depend on the smaller of the two ratios.

The last harvest control rule implemented is a model-based as uses outputs of the YCC (Cotter et *al*. 2004) assessment.

$$TAC_{y+1} = TAC_y * \min\left(\frac{Z_{[0,1]}^{PA}}{Z_{y-1,[0,1]}^{YCC}}, \frac{Z_{[2,6]}^{PA}}{Z_{y-1,[2,6]}^{YCC}}\right). \quad (5)$$

It is quite similar to the previous one. The major difference is that the harvest rate is estimated by the assessment and not directly from the survey.

Results

The biological model was parameterised using ICA results and ICES working group settings. Data provided by FRS Aberdeen begin in 1960. Those data contain stock numbers at age, fishing mortality and biological information required by the operating model. The population dynamics model starts in 1960 and runs to 2006.

The model showed that without an observation error in the survey the first two HCRs led to an increase in the SSB and catches and a reduction in the interannual variability in catches. Although the use of a Z based HCR like the third one also gave results that indicated a potential beneficial effects on the stock the picture was less clear than with the previous two HCRs. The inclusion of observation error in the survey did not affect the qualitative results of the model but increased the uncertainty in the data making any benefits from the application of the HCRs less evident. Despite that, the application of the SSB based HCRs continues to support an increase in SSB and catches. On the other hand, the third HCR (Z-based) appears to provide greater protection to the stock leading to increases in the stock size which are more significant that those predicted with the previous two HCRs. However, that comes with a reduction in catches which nevertheless are maintained above the minimum catches observed in the past. The model based HCR (YCC) also favours stock rebuilding at the cost of reduced caches. However, although the third HCR was able to respond to increases in SSB by increasing the catches (even though slowly) the latter HCR kept catches at very low levels even after considerable increases in SSB.

The introduction of low level of noise (random value between 0 and 5%) in the acoustic survey does not make noticeable changes in the outputs of scenarios using SSB-based harvest control rules contrary to those based on Z. In fact, for the SSB-based HCRs, the shape of the curve of the SSB and the averaged value of the SSB in the final years are similar. With a HCR based on Z, the noise modifies significantly the shape of the SSB curve: dome-shaped without noise and continuous increase with noise. Also, catch increases very quickly without noise while they start by decreasing when noise is taken into consideration. SSB-based HCRs seem to be more robust to the inclusion of noise in the acoustic survey than the Z-based HCR as outputs do not change significantly with the noise.

When testing the sensitivity of the results to the stock-recruitment functions used, noticeable changes can be observed but they are linked with the HCR used and the taking or not into account of noise. Again, changes are more profound with Z-based HCRs.

b. Bay of Biscay anchovy

Harvest control rules

Four HCR's defining the TAC based on fishery independent information have been tested and compared to the current constant TAC management. Only age-aggregated abundance indices were considered for this species and no assessment procedure was applied. So, all the HCRs tested are SSB-based and model-free, i.e., based directly on the SSB observations from the surveys ($I_y = S\hat{S}B_y$). The simplest of this type of HCRs is:

$$TAC_{y+1} = \frac{S\hat{S}B_y}{S\hat{S}B_{y-1}} TAC_y \,. \qquad (6)$$

A variant of this HCR can be obtained by simply adding a restriction of +/- 20% on the inter-annual variation allowed to the TAC:

$$TAC_{y+1} = \min\left\{ \max\left\{ \frac{S\hat{S}B_y}{S\hat{S}B_{y-1}}, 0.8 \right\}, 1.2 \right\} TAC_y \,. \qquad (7)$$

The potential benefits of including a recruitment index ($\hat{R}_y$) for setting the TAC was analysed by considering the following HCR:

$$TAC_{y+1} = \frac{S\hat{S}B_y}{S\hat{S}B_{y-1}} \frac{\hat{R}_y}{\hat{R}_{y-1}} TAC_y \,. \qquad (8)$$

Some of the indicators and corresponding methods developed from the surveys in the indicator-approach in WP5 have been tested in their ability to trigger an alarm, even when the abundance indices from the surveys do not indicate a downwards trend in the population level. In order to test how an alarm triggering indicator (i.e. a binary index that takes the value 1 when the alarm is triggered and 0 otherwise) could help improving the performance of the HCRs the following HCR was considered:

$$TAC_{y+1} = \begin{cases} \alpha \dfrac{S\hat{S}B_y}{S\hat{S}B_{y-1}} TAC_y & \text{when } \hat{A}_y = 1 \\ \dfrac{S\hat{S}B_y}{S\hat{S}B_{y-1}} TAC_y & \text{when } \hat{A}_y = 0 \end{cases} \tag{9}$$

That is, when the indicator triggers an alarm (i.e. $\hat{A}_y = 1$) immediate action is taken and the TAC is reduced automatically by a fraction α.

The alarm-triggering binary index, $\hat{A}_y$, has been simulated as follows: When the true population biomass is below B_{lim}, the probability that an alarm is triggered is of 0.9, i.e.:

$$P(\hat{A}_y = 1 \mid SSB_y < \mathrm{B}_{\mathrm{lim}}) = 0.9\,, \tag{10}$$

and when the true population biomass is above B_{lim}, the probability that a false alarm is triggered is of 0.05, i.e.

$$P(\hat{A}_y = 1 \mid SSB_y \geq \mathrm{B}_{\mathrm{lim}}) = 0.05\,, \tag{11}$$

corresponding to a relatively good indicator, with low type I and type II errors.

Results

The operating model for the Bay of Biscay anchovy has been parameterized based on the results from the Integrated Catch-at Age (ICA, Patterson and Melvin 1996) from the latest WGMHSA (ICES 2006), in which the population is structured in 6 age classes (from 0 to 5+) covering the period from 1987 to 2005.

The traditional management procedure for this stock sets the TAC constant at 30 000 t. Thus, this option was tested before any other HCR was used. The calculations showed such a management approach results in a high risk for stock collapse even if the stock before the application of this TAC is not overexploited.

The performance of the first of the HCRs presented above appears to depend on the starting conditions and especially the starting TAC. Starting from the current stock situation with a very low TAC then the rule will result in stock rebuilding and a slow increase in catches. However, if we keep the fishery closed for two years and then start with the usual TAC of 30 000 t and a less depleted stock, then the rule tends to favour small changes in TAC and thus any increase in population size is achieved much slower than in the former case. A similar dependency on the initial conditions is observed with the remaining HCR tried.

For the second HCR, the restrictions on inter-annual variability in the TAC decreased slightly the catch levels while increased the population levels. The use of the recruitment index in the HCR (third HCR) allowed adjusting the TAC with a better knowledge of the situation of the stock in the next year. This led to larger catches while keeping the depletion probability of the stock low (below 20%). The last HCR of those shown above (i.e. the one incorporating the alarm triggering indicator) with a highly risk-averse reduction factor appears to provide better protection against stock depletion than the other rules.

c. North East Artic Cod

Harvest Control Rules

Three HCRs were considered that used a PID-controller as described in (Bogaards 2007); two Z-based (a model-free and a model-based one) and one SSB-based. The HCRs were parameterised based on the results of deterministic calculations and then were used in stochastic calculations.

The HCR is of the form (see Bogaards (2007) for details) :

$$TAC_{y+1} = RT * f(\mu_y) * TAC_y \qquad (12)$$

Where

$$RT = \begin{cases} 1, \text{ if signal in biomass } (B) \\ -1, \text{ if signal in mortality } (Z) \end{cases}$$

And the response from the HCR is:

$$f(\mu_y) = \mu_y = K_P e_y + K_I \sum e_i + K_D (e_y - e_{y-1}) + 1$$

And the error is:

$$e_y = (signal_y - ref.point)/ref.point$$

Results

The model was conditioned using historical data from ICES form the period of 1984 to 2005. The last year with historical data is 2005, so the simulation starts in 2006.

The HCR based on information in Z is very sensitive to noise in the survey index as well as uncertainty in key biological parameters such as natural mortality. The stock did not reach equilibrium with this rule but showed a steady increase to sizes much greater than those observed in the past. The catches on the other hand were reduced and the average values remained below the recent catches. A Z-based HCR with a signal that was analysed using a stock assessment model (YCC) was also not able to control the stock in a better way than that observed with the previous HCR. There can be many explanations for this; we may have not chosen the best regression model within YCC, or the best signal from the assessment. An assessment tool can smooth the signal from a survey but if the signal is weak such analysis might reduce the value of that signal.

The chosen SSB-based HCR was able to stabilise the stock at an average level within the observed biomass and was able to control the stock even with moderate levels of noise in the survey index. However, the risk that the stock will collapse remained considerable indicating that a strict control and monitor of the stock might be needed if such a rule is chosen for the management of this species.

d. North Sea Cod

Harvest Control Rules

Three types of HCR were investigated; one based on SSB and the other two based on Z, one model-free and the other based on a linear trend estimated by YCC all of which were of the general form (see Bogaards 2007 for details or the individual case study report for North Sea cod in annex):

$$\begin{cases} TAC_{y+1} = \exp\{u_y\} TAC_y \\ u_y = K_P\, e_y + K_I \sum_{z=y-\delta}^{y} e_z + K_D\left(e_y - e_{y-1}\right) \end{cases} \tag{13}$$

In the SSB-based HCR, only a moving target (previous years' index) was evaluated. In the Z-based HCRs, both moving targets (previous years' index) and fixed targets (absolute mortalities or stable mortality) were evaluated. Tuning was done according to modified Ziegler-Nichols settings, in order to obtain a smooth response in the control signal.

Results

The operating model was parameterized using ICES WG estimates as of 2006 (data for 1963-2005, plus group at age seven).

The SSB-based HCR chosen as the most appropriate one hardly ever resulted in stock collapse and led to stock increase above the reference level within a few years of its introduction. Increased survey noise reduced the potential of the HCR to prevent stock decline and reduce interannual variability in catches. Nevertheless, even with high levels of noise the HCR performed. It also appeared to be robust to low levels of misreporting. Calculations with varying forms of a Z-based HCR that used year-class curve analysis (YCC) also showed that a parameterisation could be found that can eliminate the risk of stock collapse. However, the average annual catch supported by this rules was much smaller than that found when the SSB-based rule was used while its robustness to misreporting was lower than that of the SSB-based rule. The chosen tuning method was not able to produce a satisfactory control signal when the Z-based model-free HCR was used and therefore, this HCR was not considered further. The results with the other two rules depended upon the assumptions about the biology of the species (i.e. stock recruitment relationship, etc.)

5. DISCUSSION

For herring, HCRs based on SSB index tend to be better options than the one based on Z. It is worth noting though that the way Z is used in the harvest control rule tends to limit the increase or decrease of the TAC. This is because the TAC for the following year depends on the minimum of the two Z ratios (age class 0:1 and age class 2:6). Thus it is smaller than if all age classes were fused. More Z-based HCRs might also need to be considered for a thorough evaluation of the potential of such rules to provide robust management advice. Nevertheless, the analysis showed that it is possible to manage the North Sea herring stock using only fishery independent data with low levels of noise in the survey. It highlighted though the fact that the efficiency of a management approach depends not only on the choice of the HCR but also on the level of understanding of key biological processes (e.g. the stock-recruit function), and the assessment method.

For Bay of Biscay anchovy, only SSB based and model free HCRs were evaluated. For a short lived species as this, variations in Z mainly reflect the changes on the yearly incoming recruitment, so that Z-based HCRs are considered inappropriate. SSB-based HCRs proved to be useful to modulate the TAC. However, when the stock is depleted, more restrictive management measures seem to be needed until the population recovers within safe biological limits. Additional survey indices, such as the recruitment index and the alarm triggering indicator, resulted to be useful to improve the performance of the HCR: the former, in terms of maximizing catch while keeping the population within safe biological limits, and the later, in terms of reducing the stock collapse probability. In general terms, and as expected, the more precise the survey indices were, the better the performance of the HCR was.

For North East Artic cod, the results showed that if one allows for large inter-annual variation in TAC it is possible to manage the stock without data from the fisheries providing that a biomass index is available. However, the effectiveness of a HCR is very sensitive to variations in the recruitment. The Z-based HCRs on the other hand appear to be very conservative favouring increases in stock size. Although this might not be desirable in the context of this study, the results show that such rules can also be of value especially in cases in which the recovery of the stock receives priority.

For North Sea cod, it was found that model-free estimates of Z could not serve as the basis for a sensible HCR, nor did YCC-based control with a fixed target. HCRs based on survey SSB yielded higher average annual catches, but also a higher inter-annual catch variability, as compared to the more conservative HCRs based on YCC. The latter appear more robust to changes in survey measurement error, but survey-based HCRs were comparatively more robust to misreportings. The study showed that, in principle, it is possible to obtain excellently performing HCRs based on survey-derived information only. However, performance of HCRs is strongly dependent on the type of stock-recruit relation used. Hence, instead of aiming to maximize catch based on modelling assumptions it is better to strive for a robust management, that performs reasonably well given the uncertainties in recruitment of North Sea cod.

The scenarios considered in the four case studies showed that very conservative HCRs are needed to control a system with high variability regardless of whether that comes from external factors (e.g. high misreporting or uncertain survey results) or from the stock itself (e.g. yield is sustained by a couple of year classes, etc). This study assumed that only one index was available per year which, however, was unbiased; Availability of more indices may provide more information so that we are able to control the stock within the uncertainties in the model including potential bias in the indices.

The results of this analysis showed that survey-based indices could support fishery management when the appropriate HCRs are in place. However, they cannot eliminate the complexity introduced into fisheries management by limited knowledge of stock and fishery behaviour (complications that conventional fishery management also has). The benefits with using the survey based indices is that scientists can use the knowledge acquired through exercises as those presented here to modify and improve their survey design to provide case-specific indices based on the characteristics of the system and the requirements of the management approach chosen. The fact that catch data are not necessary to provide management advice also means that typical problems related to fishery dependent data such as catch misreporting can be avoided. The fishery models used in this analysis are quite simple and do not allow evaluation of the performance of the HCRs in terms of the fishery and economical and social issues. Although such aspects could be equally important, their consideration in the calculations was beyond the scope of this analysis.

References

Annon 2006, "Northeast Arctic Cod," in *Report of the ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems, 2006. ICES Advice. Books 1 - 10. 3, 89 pp.*, ICES, pp. 29-38.

Bogaards, H. 2007. FISBOAT FLR loop: document on harvest control rules. Working document to the FISBOAT project. http://www.ifremer.fr/drvecohal/fisboat/.

Cotter, A. J. R., L. Burt, et al. (2004). "Are stock assessment methods too complicated?" <u>Fish and Fisheries</u> **5**(3): 235-255.

ICES. 2006. Report of the working group on the assessment of mackerel, horse mackerel, sardine and anchovy (WGMHSA), 5-14 September 2006, Galway, Ireland, ICS CM 2006\ACFM:36. 601 pp

Patterson, K. R. and Melvin, G. D. 1996. Integrated catch-at-age analysis, version 1.2. Scottish Fisheries Research Report 38.

Document 6: Simulation Evaluation with ALADYM – Methods

*****

Suite of 3 documents on ALADYM tools:
Aladym-r, Aladym-q, input spreadsheet description

Contribution to the FISBOAT project
(EU FP6 STREP n° 502572)

*****

The quasi-deterministic tool Aladym-r

SIMULATING POPULATION DYNAMICS. ALADYM MODEL (V 08)

Lembo G., A. Abella, F. Fiorentino, S. Martino and M.T. Spedicato (SIBM)

INTRODUCTION

ALADYM (**A**ge-**L**ength B**a**sed **Dy**namic **M**odel) is an age-length based simulation model developed within the conceptual framework of dynamic pool models, following the predictive Thompson & Bell (1934) approach.

The model is designed to predict, through simulations, the effects of different fishing pressure scenarios on a single population, in terms of different metrics and indicators. Removals are simulated on the basis of the total mortality rate modulated using harvesting pattern and a fishing activity coefficient. *Aladym* can work in absence of fishery-dependent data, although its predictive capability of real catch levels can be verified using information on commercial catches or fishing activity per month.

From the *Aladym* core model three complementary, but independent, tools have been derived:

A) the quasi-deterministic dynamic tool named *Aladym-r*;
B) the tuning tool *Aladym-z*;
C) the stochastic dynamic tool named *Aladym-q*.

The core *Aladym* model is described in this chapter together with *Aladym-r* and *Aladym–z*, while *Aladym-q* is described in the following one.

General assumptions

The basic assumptions of the model are:

- natural mortality as estimated reflects the rate of decline of a population for all causes excluding fishing;
- total mortality *Z* reliably reflects the decline of ages/sizes in the population, including the effects of different fishing gears;
- the growth, the natural mortality, and the maturity parameters are assumed constant along the time;

- given the small time interval (1 month) between cohorts the effect of the spreading of the lengths respect to the ages can be neglected.

DERIVATION

The quasi-deterministic dynamic tool *Aladym-r*

General framework

The model is designed to simulate population dynamics of a given species accounting for differences by sex in growth, maturity and mortality. All the quantities are calculated as vectors with a time step Δt (time slice=1 month).

An operational framework of the *Aladym-r* model is in fig.1. The step A) regards the *input and initialization.* In order to generate an unbiased initial population, the number of runs specified by the user (e.g. 100) is performed in this step, randomly varying the recruitment, the growth and the size at maturity parameters according to the values and distributions specified by the user. The user can choose among the following distribution type: log-normal, normal, gamma and uniform, for the parameter t_0 a uniform distribution is associated by default.

Two populations are generated: the exploited (where total mortality is acting) and the unexploited one (where natural mortality only is acting).

The obtained initial populations enter in the *start loop* (or *seed run*) (step B in fig 1), where the dynamics is formulated following the evolution of several cohorts at monthly scale. Here the number of recruits entering in the population is generated from a stock-recruitment relationship. Alternatively, it is given as an input vector. In both cases a uniform variability on the obtained number of recruits can be set by the user. The *start loop* runs for a number of years that is a multiple of the two sex life-spans. This step aims to eliminate the artefacts in the initial population due to the use of an equilibrium model in the initialization step. After this phase, the *simulation loop* starts and runs along a period required by the user (step C in fig. 1) generating the outputs (step D in fig. 1).

Model components

Growth

The growth process is modelled using a VBGF.

$$L_{age} = L_\infty \cdot \left(1 - e^{-K\cdot(age - t_0)}\right).$$

For each age (time step Δt = 1 month) length is calculated using the input parameters L_∞, K and t_0.

The average length in the time interval (t, $t+\Delta t$) is calculated as:

$$\overline{L}_{age} = L_\infty + \frac{(L_{age} - L_{age+\Delta t})}{K \cdot \Delta t}$$

The weight at average length, for each age, is calculated from the length-weight relationship in the form: $W_{age} = a\overline{L}^b_{age}$; with a and b as input parameters.

Population

The population dynamics is formulated following the simultaneous evolution of several cohorts at month scale through the exponential population decline model, both in absence (1) and in presence (2) of fishing mortality:

$$\frac{dN}{dt} = -MN \qquad (1)$$

$$\frac{dN}{dt} = -ZN \qquad (2)$$

used respectively in the form (3) and (4):

$$N_{(t+\Delta t),j} = N_{t,j} e^{-M_{t,j} \cdot \Delta t} \quad (3)$$

$$N_{(t+\Delta t),j} = N_{t,j} e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t} \quad (4)$$

where *j* indicates the cohort, *t* the time, *Z*, *M* and *F* the total, natural and fishing mortality respectively. (Notice that in any formula where *j*, *age* and *t* are present, it is assumed that *age* represents the age of the cohort *j* at time *t*).

<u>Maturity</u>

Maturity *Mat* is a function of the length *L* and is calculated following an ogive model (Quinn and Deriso, 1999):

$$Mat(L) = \frac{1}{1 + e^{-r(L - Lm50\%)}}$$

where *r* is the ogive slope and $L_{m50\%}$ is the length at which 50% of fish matures.
The proportion of mature fish at age is computed as:

$$Mat_{age} = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{m75\%} - L_{m25\%}}\right) \cdot \left(L_{m50\%} - \bar{L}_{age}\right)}}$$

where the maturity range $L_{m75\%}$-$L_{m25\%}$, is related to the ogive slope.

<u>Biomass</u>

The biomass (B_j) and the spawning stock biomass (SSB_j) of the cohort *j* at time *t* are respectively computed as:

$$B_{t,j} = N_{t,j} \cdot w_{age};$$

$$SSB_{t,j} = N_{t,j} \cdot w_{age} \cdot Mat_{age}$$

Analogously, the unexploited biomass (UB_j) and the unexploited spawning stock biomass ($USSB_j$) of the cohort *j* at time *t* are calculated as:

$$UB_{t,j} = UN_{t,j} \cdot w_{age};$$

$$USSB_{t,j} = UN_{t,j} \cdot w_{age} \cdot Mat_{age}$$

<u>Initial recruitment and stock recruitment relationship</u>

During the step A) (fig. 1) the initial number of individuals in the population are from estimates of recruitment independently obtained from e.g. trawl surveys or other sources.
These numbers randomly selected for each of the e.g. 100 runs (see also the *general framework* paragraph) are used to initialize the population.
Successively (step B and C in fig. 1), the number of individuals entering in the population can be a vector or is estimated from one of the following user selected stock-recruitment relationships:

Beverton & Holt (1957):

$$R = \frac{S}{(a + bS)};$$

Ricker (1954):

$R = a \cdot S \cdot e^{(-bS)}$;

Shepherd (1982):
$R = a \cdot S / \left[1 + (S/c)^b\right]$;

Barrowman & Myers (2000):
$$R = \alpha \cdot \min(S, S^*) = \begin{cases} \alpha \cdot S \ \ if \ \ S < S^* \\ \alpha \cdot S^* \ if \ \ S \geq S^* \end{cases}$$

$$R = \begin{cases} \alpha \cdot S & if \ \ S \leq S^* \cdot (1-\delta) \\ \alpha \cdot (S - \dfrac{(S - S^* \cdot (1-\delta))^2}{4\delta \cdot S^*}) & if \ \ S^* \cdot (1-\delta) < S < S^* \cdot (1+\delta) \\ \alpha \cdot S^* & if \ \ S \geq S^* \cdot (1+\delta) \end{cases}$$

R and *S* represent the number of recruits and spawners respectively, whilst *a*, *b*, *c*, α, δ, S^* are the model's parameters. Uniformly distributed random variations can be applied by the user to the number of offsprings (from vector or stock-recruitment relationship).
The number of the events (on monthly basis) generating the offsprings is an input of the model.
The population of spawners generating the recruits is calculated summing up the number of individuals of the different age classes of the different cohorts occurring in the population one or more (depending on the species biological features) months before the offsprings are produced. Thus this quantity is calculated as follows:
$SSN_t = \sum_j SSN_{t,j}$;
where $SSN_{t,j}$ represents the number of mature females at time *t,* of the cohort *j*:
$SSN_{t,j} = N_{t,j} \cdot Mat_{age}$.

Mortality

The natural mortality can be constant for each age/length, or a vector by age/length calculated outside the model and used as input. Alternatively, it is estimated inside the model from the Chen and Watanabe equations (1989):

$$M(t) = \begin{cases} \dfrac{K}{1 - e^{-K \cdot (t - t_0)}} & t \leq t_M \\ \dfrac{K}{a_0 + a_1 \cdot (t - t_M) + a_2 \cdot (t - t_M)^2} & t \geq t_M \end{cases}$$

where:

$$t_M = -\frac{1}{K} \ln\left|1 - e^{Kt_0}\right| + t_0$$
$$a_0 = 1 - e^{-K \cdot (t_M - t_0)}$$
$$a_1 = K \cdot e^{-K \cdot (t_M - t_0)}$$
$$a_2 = -\frac{1}{2} K^2 \cdot e^{-K \cdot (t_M - t_0)}$$

Two are the parameters of the Chen and Watanabe model, t_0 and K. The asymptotic length (L_∞) is not necessary, but t_0 cannot be equal to 0 (otherwise the parameter t_M cannot be defined). The quantities a_0, a_1, a_2 and t_M cannot be strictly considered as parameters of the model, as they depend from t_0 and K. The parameter t_M represents the age beyond which the contribution of the fish of a given cohort can be considered negligible. If parameters are consistent the relationship between age and natural mortality shows a “bath tube” shape.

The fishing mortality rate $F(L)$ is modelled for each cohort using the following general equation (Sparre and Venema, 1998):

$$F(L) = F_{max} \cdot S(\bar{L})$$

where F_{max} is the maximum fishing mortality and $S(\bar{L})$ the proportion of retained fish.

In *Aladym* the fishing mortality rate is calculated as follows:

$$F(L) = F_{max} \cdot S(\bar{L}) \cdot f_{act}$$

where maximum fishing mortality (F_{max}) is calculated as follows:

$$F_{max} = QZ_{input} - M_{\min}$$

using the input values of QZ (a Z proxy) and where M_{min} represents the minimum value that the M vector assumes. As an alternative option, F_{max} can be also a user selected input to be set for each month. In addition, a fishing activity coefficient (f_{act}) is introduced in order to consider the possibility of a fishing ban or changes in fishing effort throughout time.

The value of QZ by sex can be assumed, as a first order approximation, numerically equal to the value of Z observed that is obtained from estimations outside the simulation model (e.g. from trawl-survey). A better approximation of QZ is obtained using the tool *Aladym-z* (see a later paragraph).

In the model the probability of selection $S(\bar{L})$ of the cohort j is calculated at time t from one of the two following user selected relationships:

$$S(\bar{L}) = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{75\%} - L_{25\%}}\right) \cdot \left(L_{50\%} - \bar{L}_{age}\right)}};$$

or

$$S(\bar{L}) = \frac{1}{1 + e^{\left(\frac{2 \cdot Ln3}{L_{75\%} - L_{25\%}}\right) \cdot \left(L_{50\%} - \bar{L}_{age}\right)}} \cdot \frac{1}{1 + e^{\left(\frac{-2 \cdot Ln3}{D_{25\%} - D_{75\%}}\right) \cdot \left(D_{50\%} - \bar{L}_{age}\right)}};$$

where $L_{50\%}$, $L_{75\%}$ and $L_{25\%}$ are the selectivity parameters and $D_{50\%}$, $D_{25\%}$, $D_{75\%}$ the de-selection parameters of the model.

The total mortality Z at time t for the cohort j is thus computed as:

$$Z_{t,j} = F_{t,j} + M_{t,j}$$

that is the value acting on the population in the model computations.

The biomass of individuals of the cohort j at time t death for all causes ($BP_{t,j}$) is computed as:

$$D_{t,j} = N_{t,j} - N_{t+\Delta t,j} = N_{t,j} \cdot \left(1 - e^{-Z_{t,j} \cdot \Delta t}\right),$$

$$BP_{t,j} = D_{t,j} \cdot w_{age};$$

while the biomass of those death for all causes excluding fishing ($BND_{t,j}$) is computed as:

$$BND_{t,j} = \frac{M_{t,j}}{Z_{t,j}} \cdot N_{t,j} \cdot (1 - e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t}) \cdot \mathrm{w}_{\mathrm{age}}.$$

Harvest control rules

The simulation approach can be used as a tool to convert survey biological information and relative assessment into quantitative HCRs. The options implemented in the simulation model are based on the following aspects: *QZ*, gear selectivity (size at first capture $L_{50\%}$ and selection range) and fishing activity (alone or in combination). These three are inputs that can be used to simulate different exploitation scenarios. The effects of HCRs (selectivity and fishing activity) are then analysed in terms of sustainability for the population in the long-term. For example, the ratio between the mean spawning stock biomass and the mean unexploited spawning stock biomass (*SSB/USSB*, output) is also estimated for each harvesting scenario.

A vector of yield (*Y*) by time is also simulated, estimating the catch (*C*) according to the following general equation (Gulland, 1969):

$$C_{\Delta t} = \int_0^{\Delta t} F \cdot N_0 \cdot e^{-Z \cdot \tau} d\tau = \frac{F}{Z} N_0 \cdot \left(1 - e^{-Z \cdot \Delta t}\right)$$

where Δt is the time to which the catch is referred.

Thus the catch (Yield) in the time interval (t, $t+\Delta t$) is computed in *Aladym* as (Sparre and Venema, 1998):

$$Y_{t,j} = \frac{F_{t,j}}{Z_{t,j}} \cdot N_{t,j} \cdot (1 - e^{-(F_{t,j} + M_{t,j}) \cdot \Delta t}) \cdot \mathrm{w}_{\mathrm{age}} .$$

SOFTWARE

Aladym is written in the R language and licensed as open source under GPL2.

The data and parameters feeding the model can be easily entered using an excel data sheet.

The results of the simulation are stored into three Export files (.din for inputs, .dou for outputs, .RData for the R workspace) and saved in the same directory where R is started using the basename of the input sheet.

To give an idea of the running time, *Aladym-r* requires about 25 seconds (assuming 40 years of start loop and 20 years of simulation) with a Intel (R) Pentium (R) personal computer with a processor of 1.70 GHz and 1 GB RAM.

The tool *Aladym-z* requires about 2.6 hours (assuming 40 years of start loop and 20 years of simulation) with a Intel (R) Pentium (R) personal computer with a processor of 1.70 GHz and 1 GB RAM.

The software can be downloaded from the Fisboat web-site, where also a detailed description of the input sheet for user help is available.

INPUTS

Input parameters to the *Aladym-r* model are:

- von Bertalanffy growth parameters by sex with associated variability,
- length-weight relationship parameters by sex;
- maturity ogive parameters by sex ($L_{m50\%}$ and $L_{m25\%}$-$L_{m75\%}$ range);
- natural mortality by sex (a constant value or a vector);
- seed values (minimum, maximum, *ln*-mean and *ln*-standard deviation) of recruitment by sex;
- proportion of offsprings entering in the stock by month;
- stock-recruitment relationship parameters or a vector of recruit numbers by month both with associated variability;
- time elapsing from spawning to birth;
- sex-ratio (female/total) of offsprings;
- F_{max} by month (option 2) or from the model (option 1);

- *QZ* by sex;
- selection ogive parameters (2 options) of the gear used by the fleet ($L_{50\%}$ and $L_{25\%}$-$L_{75\%}$ range, $D_{50\%}$ in case of the selectivity option 2);
- fishing activity coefficient by month (0, in case of absence of fishing activity).

OUTPUTS

The outputs automatically produced by the simulations of *Aladym-r* can be summarised in the following items.

Export data file (.dou):

1. exploited and unexploited population by sex, per month and age;
2. exploited and unexploited biomass by sex, per month and age;
3. exploited and unexploited population of females belonging to the spawning stock per month;
4. total mortality Z calculated by the model for females, males and the whole population in each month and year of the simulation as follows (Sinclair, 2001):

$$Z_t = \frac{1}{\Delta t} \ln \left(\frac{\sum_{j=1}^{\infty} N_{t,j}}{\sum_{j=2}^{\infty} N_{t+\Delta t,j}} \right);$$

5. exploited and unexploited biomass per month;
6. exploited and unexploited spawning stock biomass per month;
7. ratio between exploited and the unexploited spawning stock biomass per month;
8. average length and age of exploited and unexploited populations per month;
9. average length and age of exploited and unexploited spawning populations per month;
10. yield in tons per month;
11. average length and age of catches per month;
12. fishing mortality per month calculated as;

$$F_t = \frac{1}{\Delta t} \ln \left(\frac{\sum_{j=1}^{\infty} N_{t,j}}{\sum_{j=2}^{\infty} N^F_{t+\Delta t,j}} \right),$$

where $N^F_{t+\Delta t,j}$ is the number of survivors at the time $t+\Delta t$ under the hypothesis that only fishing mortality is acting;

13. biomass of natural losses and total biological production per month.

Plots per year of the outputs listed from items 4 to 13 are also produced.

Some other outputs are also made available to the user:

1. average length at age and age by sex;
2. natural mortality at age/length by sex;
3. weight at age/length by sex;
4. proportion of mature individuals at age/length by sex.

These outputs help the user to check the results obtained from the sub-models, in particular those related to the VBGF, the length-weight relationship, the natural mortality, and the maturity.

PRACTICAL GUIDELINES

The *Aladym* core model does not make any fixed or hidden (to the user) assumption about the values of the parameters describing the behaviour of the equations on which the model itself is built.

The user is allowed to (and need to) input all the parameters involved: whilst this makes the model highly flexible in adapting to different species/environments it loads the user with the responsibility to validate each single value and to assess the coherence as a whole.

Very few checks are foreseen at the moment to supervise the consistency of the data supplied: often is a critical analysis of the results which spots such consistency. The checks guarantee the positivity of F_{max}, of length at t_0 and a sex ratio between 0 and 1.

The model is extensively based on closed form solution to the dynamical equations it solves, thus two key options, both related to the early phase, are available for tuning: the 'Multiplier of Life-span' which controls the amount of years that had to be simulated in order to cancel the artefacts from the equilibrium model used to initialise the population; the 'Number of Run for seed randomization' which sets the number of samples to be taken in order to derive the average values for the growth and population parameters. For both parameters the rule is: bigger is better, however the default values (1, 100) are a reasonable choice.

One of the parameters highly influencing the behaviour of the model is *QZ* which, however, does not have an immediate counterpart but can be naively associated to the total mortality *Z*. A specific tool (*Aladym-z*) has been developed which starting from the observed values of Z and the description of the life and population traits is able to calculate the values of *QZ* which better approximates the given scenario.

Starting from the *Z_observed*, *Aladym-z* iterates the model modifying, in each run, the amplitudes of the *QZ* waveforms and it stops when the Least Square convergence criteria are met.

Sensitivity

The extensive number of the simulation run performed evidenced that the model behaviour is influenced by the consistency between the set of life-history parameters and population dynamics. The model results are thus expected to be particularly sensitive to the stock-recruitment relationship and natural mortality.

Strengths/weaknesses

In *Aladym* the following points can be considered the strength ones:

- the model is designed to work in absence of fishery-dependent information;
- the model is built using separated components that give it enough flexibility to account for the use of different equations;
- the model allows the population dynamics to evolve in a very detailed time scale, thus permitting to analyse fluctuations within the year;
- the detailed time scale allows modelling the effects of the harvest control along the year;
- the model allows a input natural mortality varying by age/length, thus being able to account for species exploited also at early phase.

The following points can be considered as the weak ones:

- the model does not account for environmental changes, such as those related for example to temperature variations, or food availability;
- the life-history traits that are used for modelling the population dynamics (e.g. growth, natural mortality, maturity) are assumed stable along the time and no density dependent, only the direct effects of fishery on the population are considered;
- the model does not include spatial behaviour components;
- harvesting scenarios based on the control of the total catches are not foreseen;

- the user should be aware of the range of validity of the sub-model parameters such as those related to the stock-recruitment relationships.

REFERENCES

Beverton R.J.H., Holt S.J. (1957). On the dynamics of exploited fish populations. U.K. Min. Agric. Fish., *Fish. Invest.* (Ser. 2), 19: 533 pp.

Barrowman N.J., Myers R.A. (2000). Still more spawner-recruitment curves: the hockey stick and its generalizations. *Can. J. Fish Aquat. Sci*., 57: 665-676.

Chen S., Watanabe S. 1989. Age Dependence of Natural Mortality Coefficient in Fish Population Dynamics. *Nippon Suisan Gakkaishi*, 55(2): 205-208.

Gulland J.A. (1969). Manuel des methods d'évaluation des stocks d'animaux aquatiques. Première partie-Analyse des populations. *Manuels FAO de science halieutique*, N. 4 : 160 pp.

Quinn T.J.II, R.B. Deriso (1999). *Quantitative fish dynamics*. Oxford University Press, Oxford, U.K.: 542 pp.

Ricker W.E. (1954). Stock and recruitment. *J. Fish. Res. Bd. Can.*, 11: 559-623.

Shepherd J.G. (1982). A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. *Journal du Conseil, Conseil Internationale pour l'Exploration de la Mer*, 40 (1): 67-75.

Sinclair, A. F. (2001). Natural mortality of cod (*Gadus morhua*) in the Southern Gulf of St Lawrence. ICES Journal of Marine Science, 58: 1-10.

Sparre P., Venema S.C. (1998) - Introduction to tropical fish stock assessment. Part 1, manual. *FAO Fish. Techn. Pap*., 306 (1) Rev. 2: 407 pp.

Thompson W.F., Bell F.H. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. *Rep. Int. Fish. (Pacific Halibut) Comm.*, (8): 49 p.

Working documents

Lembo G, M.T. Spedicato (2006a). ALADYM – Age-Length based dynamic Model. *In*: Petitgas P. (coordinator). Fisboat Activity Report, Appendices: 113-114 p.

Lembo P. and M.T. Spedicato (2006b). ALADYM Age-Length Based Dynamic Model. MEDITS Coordination Meeting, Kavala, Greece, April 2006.

Lembo G., Spedicato M.T. (2006c). Red mullet assessment in the GSA 10 using Aladym simulation model. Sub-Committee Stock Assessment SAC-GFCM, FAO Headquarters, 11-14 September 2006. 6 pp.

Lembo G., Spedicato M.T., Abella A., Fiorentino F, Martino S. (2006d). Aladym description (version 06). Fisboat Meeting, IJmuiden 06-10 November 2006, 11 pp.

Lembo G., Spedicato M.T. (2006e). Aladym: un modello previsionale di valutazione e gestione per le risorse della pesca nel Mediterraneo. 5° Convegno Nazionale delle Scienze del Mare, promosso ed organizzato da CoNISMa in collaborazione con A.I.O.L., S.I.B.M. e S.It.E. Viareggio, 14-18 Novembre 2006. abstract.

Lembo G., Martino S., Abella A.J., Fiorentino F. and M.T. Spedicato. 2007. *ALADYM* (**A**ge-**L**ength B**a**sed **Dy**namic **M**odel): a stochastic simulation tool to predict population dynamics and management scenarios using fishery-independent information. Scientific Advisory Committee-Sub-Committee Stock Assessment. Workshop on trawl survey based monitoring fishery system in the Mediterranean, Rome, Italy, 26-28 march 2007. 6 pp.

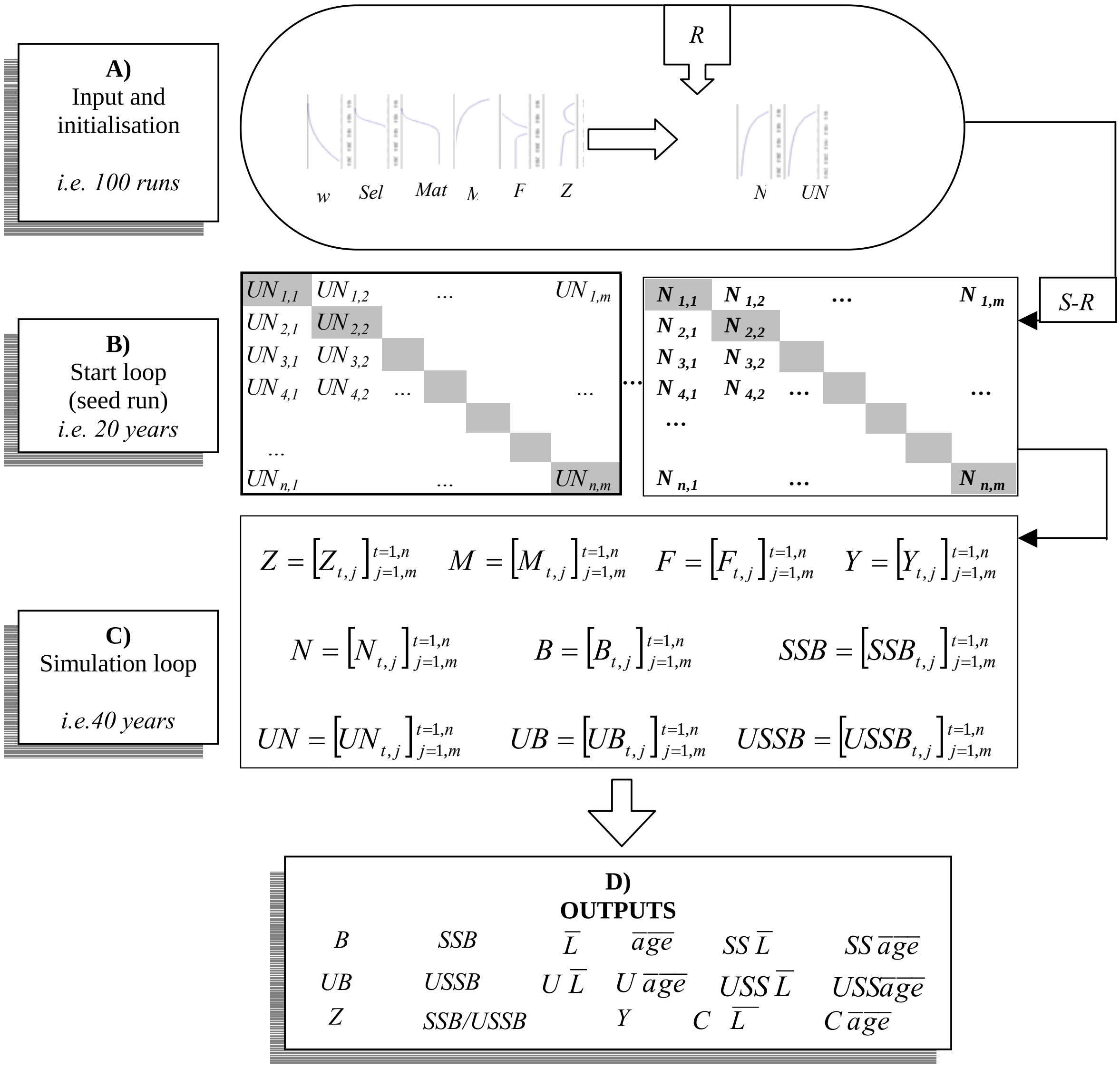


Fig. 1 – Scheme of the *Aladym-r* tool. *R*=recruitment; *w*=individual weight; *Sel*=selectivity; *Mat*=maturity; *M*=natural mortality; *F*=fishing mortality, *Z*=total mortality; *N*=exploited population, *UN*=unexploited population, *B*=exploited biomass, *SSB*=exploited spawning stock biomass, *UB*=unexploited biomass, *USSB*=unexploited spawning stock biomass, *S-R*=stock-recruitment relationship, *Y*=yield, *t*=time, *j*=cohort.

The stochastic dynamic tool *Aladym-q*

USING ALADYM FOR ESTIMATING MODEL-BASED INDICATORS AND SEEKING REFERENCE POINTS BY SIMULATION

Lembo G., A. Abella, F. Fiorentino, S. Martino, and M.T. Spedicato (SIBM)

INTRODUCTION

Aladym-q adds to the same mathematical model of *Aladym-r* the capability to deal with the stochastic representation of some input parameters, in order to evaluate the corresponding distribution functions of the output variables using a MonteCarlo approach. This feature aims to build a procedure to help identifying indicators and/or reference points associating them a confidence interval.

DERIVATION

The stochastic dynamic tool *Aladym-q*

The stochastic dynamic model defined as *Aladym-q* follows the same basic formulations as *Aladym-r.* The main difference consists in modelling the uncertainty of estimates related to the initial recruitment, growth and maturity traits of the population through stochastic processes.
Moreover, a uniform distribution is applied to the number of recruits generated by the stock-recruitment relationship. In addition, probability distribution functions (*pdf*) selected by the user are applied to the growth parameters K and L_∞, and to the maturity parameters. This makes *Aladym-q* more adaptable for estimating the probability associated to metrics, indicators and reference points. An operational framework of the *Aladym-q* is in fig. 2.
The step AA) regards the *input and initialization.* Given the parameters of the identified *pdfs* a first random realization is made in this step. Then the population evolves in the steps BB) and CC). These steps are reiterated for a number of realizations, sampling at each run a new set of parameters from the *pdfs*. In the output step *pdfs* and cumulative *pdfs* are generated, the latter calculated according the following general formulation:

$$f(X) = P(X < x) = \int_{-\infty}^{x} pdf(\chi)\,d\chi$$

SOFTWARE

Aladym is written in the R language and licensed as open source under GPL2. The data and parameters feeding the model can be easily entered using the same excel data input sheet as *Aladym-r*. The differences regard the number of realizations to be performed (user selected and mandatory for *Aladym-q*) and the parameters of the *pdfs* associated to growth and maturity, that for *Aladym-q* operate also in the simulation loop.
The results of the simulation are stored into three Export files (.din for inputs, .dou for outputs, .RData for the R workspace) and saved in the same directory where R is started using the basename of the input sheet.
To give an idea of the running time, using a Intel (R) Pentium (R) personal computer with a processor of 1.70 GHz and 1 GB RAM, *Aladym-q* might requires 572 seconds for 100 realizations, ~1.5 hours for 1000 realizations and about 17 hours for 10000 realizations (assuming 40 years of start loop and 20 years of simulation).
The software can be downloaded from the Fisboat web-site, where also a detailed description of the input sheet for user help is available.

INPUTS

As regards the *inputs*, besides those already mentioned for *Aladym-r*, *Aladym-q* requires:

- the number of realizations;
- the parameters of the *pdfs*.

OUTPUTS

The *outputs* automatically produced by the simulations of *Aladym-q* can be summarised in the following items.

Export data file (the quantities are related to each realization):

1. exploited and unexploited biomass in tons per month;
2. exploited and unexploited biomass of spawners in tons per month;
3. ratio between exploited and unexploited spawning stock biomass per month;
4. *Z* calculated by the model combined for sex per month and by sex for year;
5. annual *Z* calculated by the model per sex;
6. *QZ* (the input value) by sex;
7. average length and age of exploited and unexploited populations per month;
8. average length and age of exploited and unexploited spawner populations per month;
9. F per month;
10. yield in tons per month;
11. average length and age of the catches per month;
12. biomass of natural losses and total biological production in tons per month;

Plots of the *pdfs* and the cumulative (*cpdfs*) are interactively produced per year for the same items listed above.

Some other outputs are also made available to the user:

1. average number of recruits at each realization;
2. growth and maturity parameters by sex at each realization.

These outputs help the user to check the results from the sub-models related to the VBGF, the maturity, and the recruitment. In addition, they also allow to track at each realization the outputs with the related key-inputs.

PRACTICAL GUIDELINES

Same considerations as developed for *Aladym-r* hold for *Aladym-q*.

A new parameter is introduced for tuning the quality of the output pdfs: the number of realizations. This parameter should be set accounting for a trade-off between the running time and target confidence level. Experiments showed that values in the range from 1000 to 10000 give an error level varying from about 6-7 to ~1%. These confidence levels are well below the precision by which most of the input parameters are known.

As regards sensitivity and strengths/weaknesses of the models, similar consideration as developed for *Aladym-r* can be applied to *Aladym-q*, although the latter tool has the advantage of including stochastic effects in some of the key life-history traits. This stochasticity masks the effects due to uncertainty on the knowledge of input data and of their relationships.

REFERENCES

See previous chapter: Simulating population dynamics. Aladym model.

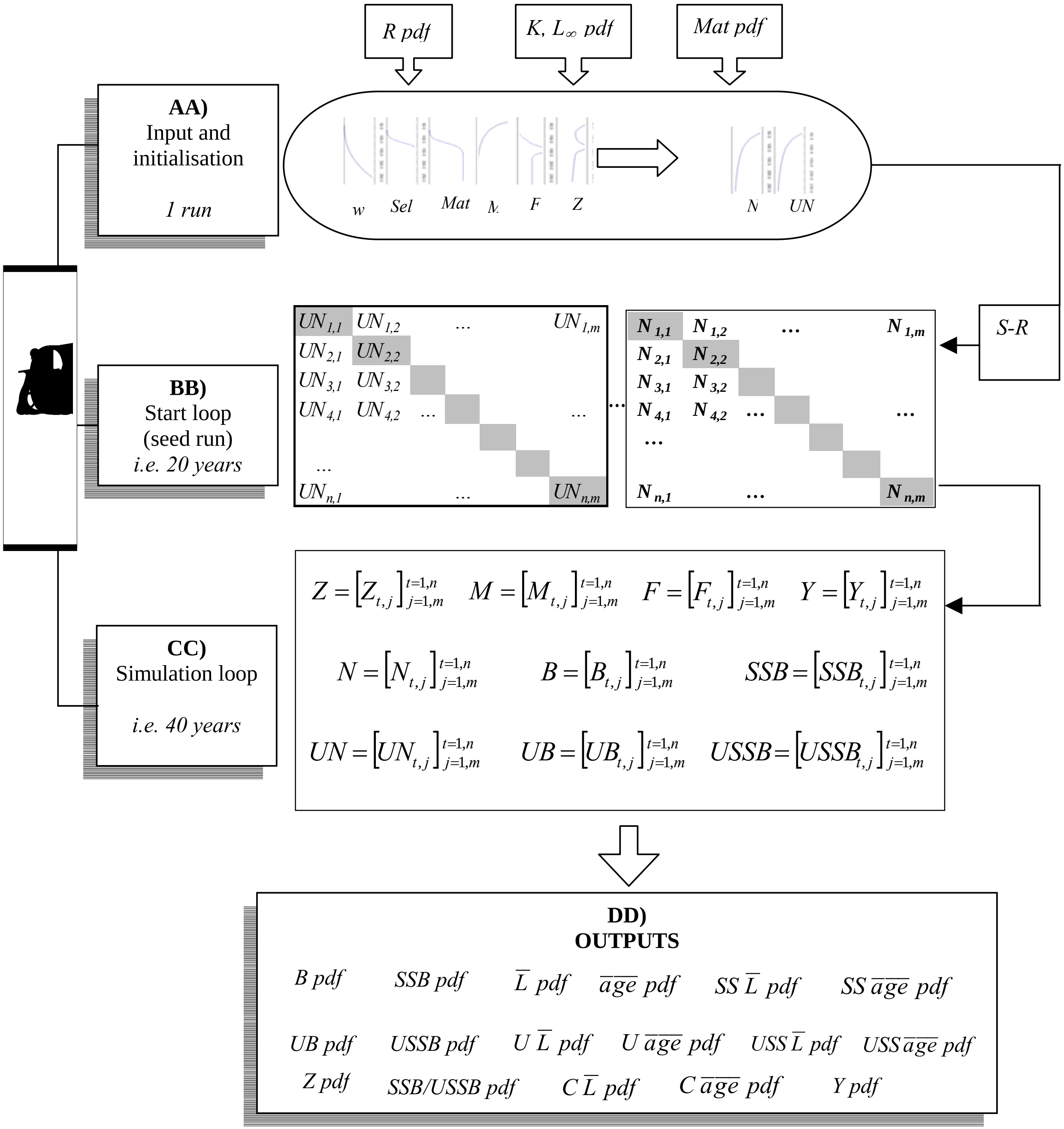


Fig. 2 – Scheme of the *Aladym-q* tool. *pdf*=probability distribution function; *K, L*∞ growth parameters, *R*=recruitment; *w*=individual weight; *Sel*=selectivity; *Mat*=maturity; *M*=natural mortality; *F*=fishing mortality, *Z*=total mortality; *N*=exploited population, *UN*=unexploited population, *B*=exploited biomass, *SSB*=exploited spawning stock biomass, *UB*=unexploited biomass, *USSB*=unexploited spawning stock biomass, *S-R*=stock-recruitment relationship; $\overline{L}$ = average length ; $\overline{age}$ = average age ; *SS*=exploited spawner's population; *USS*=unexploited spawner's population; *C*=capture in numbers; *Y*=yield, *t*=time, *j*=cohort.

The input sreadsheet for Aladym tools

ALADYM (V. 08) EXCEL INPUT SPREADSHEET DESCRIPTION

Lembo G., A. Abella, F. Fiorentino, S. Martino and M.T. Specicato (SIBM)

1. TECHNICAL ASPECTS

Aladym is written in the R language and licensed as open source under GPL2.
The parameterization of the model can be easily done using an excel spreadsheet for data input; this sheet is the same for the three tools:

D) the quasi-deterministic dynamic tool named *Aladym-r*;
E) the tuning tool *Aladym-z*;
F) the stochastic dynamic tool named *Aladym-q.*

2. CONVENTIONS

The cells are colour coded:

- Grey shaded cells: contains fixed value that at this time can not be modified. They are reserved for future use;
- Red cells: data are calculated according to other value present in the sheet. They are shown as help in inputting the model parameters;
- Green bold cells: contains the inputs and parameters of the model.

Do not move or eliminate or add cells, rows and columns.
No default value is provided for any of the parameter the model uses, so it is mandatory to fill the proper cells, on the basis of the user choices.
The species name and geographical area are for documentation only.

3. DESCRIPTION OF INPUTS

3.1 Control parameters

Cell E9 (fig. 1): to input the years to be simulated.
Cell F9 (fig. 1). The *start loop* of the model runs for a number of years that is given by the product of cells G9=F9×J17×J28 (see fig. 1 and 2). It is a *multiplier of life-span* of the two sexes and it should be chosen in order to have G9 an integral multiple of both the sex life span (G9$\propto$m.c.m. (J17, J28)).
Cell H9 (fig. 1). Contains the number of runs needed to stochastically initialise the population. Recommended value is 100 for *Aladym-r* and *Aladym-z*, while it should be set to 1 for *Aladym-q:* the software automatically does it.
Cell I9 (fig. 1). It is used only in *Aladym-q* and it should be set according to the desired error level (see document on Methods). Typical values ranges between 1000 and 10000.

	D	E	F	G	H	I
7 8	Time slice per year	Years to be simulated	Multiplier of Life-span (pre-simulation)	Years to be pre simulated	Number of Run for seed randomization	Number of Run for Histogram creation
9	12	**40**	**0.25**	70	100	1000
10	e.g. months/year if 12					

Fig. 1

3.2 Von Bertalanffy growth parameters by sex with associated variability

It is possible to use different growth parameters for males and females. The same is for life span, length-weight relationships and maturity parameters.

Cell H17-I17 (fig. 2a), cell H28-I28.Regarding the growth parameter t_0 the minimum and maximum values should be set, and the mean is calculated (Fig. 2a). The stochasticity of the parameter t_0 is uniform between these two values.

The variations of the parameter t_0 are used only in the initialization run.

It is important that the parameter t_0 is less than 0 in case the option of the natural mortality estimated inside the model by Chen and Watanabe equation is used (see below the natural mortality paragraph, option 2).

	G	H	I	J
15	t_o [years]			Lifespan [years]
16	Mean	Min	Max	
17	-0.1000	-0.11	-0.09	14

Fig. 2a

Cell Q14-U14, Q16-U16, Q18-U18, Q20-U20 (fig. 2b). The stochastic pattern of the growth parameters L_∞, and K (fig. 2b) can be chosen according to one of the options in fig. 2c. It is used in the initialization run of *Aladym* and in the several realizations of *Aladym-q*.

	P	Q	R	S	T	U
9	Parameters for the random runs					
10	Parameter	Distribution	Min	Max	A	B
11-12	Offspring R [none]	1	37,281,381	129,516,993	18.06	0.88
13-14	Male Growth K [$years^{-1}$]	3	0.234	0.286	0.260	0.0260
15-16	Female Growth K [$years^{-1}$]	3	0.132	0.162	0.147	0.0147
17-18	Male Growth $L_{infinity}$ [mm]	3	527.40	644.60	586.00	58.60
19-20	Female Growth $L_{infinity}$ [mm]	3	795.60	972.40	884.00	88.40
21-22	Male Maturity Ogive L50% [mm]	4	287.30	318.40	0.00	0.00
23-24	Male Maturity Ogive L75%L25% [mm]	4	39.30	43.50	0.00	0.00
25-26	Female Maturity Ogive L50% [mm]	4	324.90	360.00	0.00	0.00
27-28	Female Maturity Ogive L75%L25% [mm]	4	37.90	41.90	0.00	0.00

Fig. 2b

	P	Q	R	S
31	Legend			
32		Distribution	A	B
33	1	Lognormal	Mean ln(x)	Ds ln(x)
34	2	Gamma	Shape	Scale
35	3	Normal	Mean (x)	Ds (x)
36	4	Uniform	None	None

Fig. 2c

3.3 Length-weight relationship parameters by sex

Cells A22-B22, A33-B33. The parameters should be determined using measure of weight in grams and length in mm.

3.4 Maturity ogive parameters by sex

Q22-U22, Q24-U24, Q26-U26, Q28-U28 (fig. 2b). The length at first maturity $L_{m50\%}$ and the range $L_{m75\%}$-$L_{m25\%}$ are mandatory for females, as the current release do not use those parameters for males.

The stochastic pattern of $L_{m50\%}$ and $L_{m75\%}$-$L_{m25\%}$ can be chosen according to one of the options in fig. 2c. It is used in the initialization run of *Aladym* and in the several realizations of *Aladym-q*.

3.5 Natural mortality

Cells G22-H22 (fig. 3), G33-H33. There are three options:

1. a constant value of natural mortality for all the ages/length (option 1, fig. a) by sex, that has to be inputted by the user in the cell G22 and G33;
2. a vector of natural mortality by age/length and sex that is calculated inside the model according to the Chen & Watanabe (1989) equation (option 2, fig. 3);
3. a vector of natural mortality by age and sex provided by the user (option 3, fig. 3a, and fig. 3b). In this case the number of steps of the vector are computed (cells N9 and N10, fig. 3b). Each row of the column M and N of fig. 4 must be filled. The numbers in the N9 and N10 cells indicate how many steps must be filled in order to be coherent with the species life-span.

	G	H	I	J
19	Constant			
20	Mortality	Mortality Type	Legend	
21	[years-1]	[none]	1	M constant
22		2	2	Chen&Watanabe
23			3	From vector

Fig. 3a

	L	M	N
8		Mortality Vector Length	
9		Male	61
10		Female	97
11			
12			
13			
14		Mortality	
15		Male	Female
16		[years-1]	[years-1]
17	1		
18	2		
19	3		
20	4		
21	5		
22	6		
23	7		
24	8		
25	9		

Fig. 3b

3.6 Values (minimum, maximum, mean and standard deviation) of offspring

Cells Q12-U12 (fig. 2b). These values are used in the initialization run of *Aladym* to generate unbiased populations by sex. The stochastic pattern can be chosen according to one of the options in fig. 2c. *Aladym-q* will use the variability set by the user in the columns F and G of figure 5 (see next paragraph).

The model assumes a coincidence between recruits and offspring, thus the number of individuals should be referred to the 0 age group. In addition, a guess estimate of absolute recruitment is necessary.

3.7 Stock-recruitment relationship

Cells A38-D38 (fig. 4). Six options are foreseen to model the number of offspring entering in the simulation in each time step:

- Beverton & Holt model, option 1
- Ricker model, option 2
- Shepherd model, option 3
- from input vector, option 4
- from hockey-stick, option 5
- from quadratic hockey-stick, option 6

and the input choice is set in the cell D38. The parameters of the relationships are in the columns A, B, and C (row 38, fig. 4) and should be chosen considering the number of spawners expected in the population in each month.

	A	B	C	D	E	F	G	H	I
36	Stock-Recruitment Relationship Parameters			Recruitment	Legend (Recruitment Type)				
37	a (or ≅)	b (or S*)	c (or≡K)	Type	1	R=S/(a+b*S)		4	From Vector
38				4	2	R=a*S*exp(-b*S)		5	Hockey-Stick
39					3	R=a*S/(1+(S/c)^b)		6	Hockey-Stick quadratic

Fig. 4

In case the option 4 is selected, the number of offspring by month is entered in the column H of the fig. 5. The whole number of recruits should be inputted, because it is thereafter split by sex from the sex ratio value set in the column I (Fig. 5).

	A	B	C	D	E	F	G	H	I	J
63	Time slice Parameters		(total rows) =	480	+ seed run					
64										
65		QZ				Offspring Variability		Offspring	Sex Ratio	Fishing
66	Parameter	Male	Female			-%	+%		Female/Total	Coefficient
67		[years-1]	[years-1]			none	none	none	none	none
68	seed run	0.8	0.8			83.00	83.00	70613757	0.43	1.00
69	1 month	0.8	0.8			83.00	83.00	0.00	0.43	1.00
70	2 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
71	3 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
72	4 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
73	5 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
74	6 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
75	7 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
76	8 month	0.8	0.8			83.00	83.00	70613757	0.43	1.00
77	9 month	0.8	0.8			83.00	83.00	0	0.43	1.00
78	10 month	0.8	0.8			83.00	83.00	0	0.43	1.00

Fig. 5

The variability on the number of offspring generated by the stock-recruitment relationships (or from a vector) is set according to a uniform distribution. The range (±%) is given by the user in the columns F and G (fig. 5).

3.8 Proportion of offspring entering in the stock by month

Cells B58-G58, B60-60 (fig. 6). This input is used for splitting the total number of offspring generated by the stock-recruitment relationship, or inputted in the column H, according to several monthly recruitment pulses.

	A	B	C	D	E	F	G
56	Proportion of offsprings/month						
57	Total	January	February	March	April	May	June
58		0.00	0.00	0.00	0.00	0.10	0.40
59	1.00	July	August	September	October	November	December
60		0.35	0.10	0.05	0.00	0.00	0.00
61							

Fig. 6

3.9 Delay for SS calculation

Cell I60 (fig. 7). This input is used to model the time elapsing from the spawning event to the time offspring reach the length estimated from the VBGF for age 0. Typical value is 1.

	I
57	
58	Delay for SS calculation
59	[month]
60	1

Fig. 7

3.10 Sex-ratio

Column I68- (fig. 5). These values are the ratio of female offspring/total offspring. They are used to split the number of offspring generated by the stock-recruitment relationship (or from vector) in males and females population.

3.11 *QZ* (*Z proxy*) and Z-observed

Column B68-, C68- (fig. 5). These inputs are used for predicting the evolution of a given population.
The value of *QZ* can be assumed, as a first order approximation, numerically equal to the value of *Z* observed that is obtained from estimations outside the simulation model (e.g. from trawl-survey). *Z* estimates are commonly expressed on a year basis, thus if monthly estimates are not available, the same value can be set for 12 months. The possibility of different values by sex is foreseen. In the row 68 of fig. 5, the values (in B68 and C68) are referred to the seed run.
The value of *QZ* can be changed to simulate different exploitation scenarios and the consequences on the population and yields of different HCRs.
The input of *QZ* is mandatory for *Aladym-r* and *Aladym-q*.
Column W68-, X68- (fig. 8). A better approximation of *QZ* is obtained using the tool *Aladym-z.* To do this *Z-observed* (estimates from trawl survey or other source) must be also inputted. The two values *(Qz* and *Z-observed*) can be also the same. The input of *Z-observed* and *QZ* are both mandatory for *Aladym-z.* The annual *QZ* output of *Aladym-z* can be then used as input in *QZ* of *Aladym-r* and *Aladym-q*. If all the parameters and inputs of the model are consistent the differences between *QZ* and *Z-observed* should be small.

The operational suggestion is: 1) run first *Aladym-r*, 2) look and evaluate the results, 3) tune better the parameters of the model if necessary, 4) refine *QZ* with *Aladym-z*, 5) run again *Aladym-r* and then, for getting a confidence interval of your model-based indicators and reference points, run *Aladym-q*.

	W	X
66	Male Z Observed	Female Z Observed
67	[years-1]	[years-1]
68	0.8	0.8
69	0.8	0.8
70	0.8	0.8
71	0.8	0.8
72	0.8	0.8
73	0.8	0.8
74	0.8	0.8
75	0.8	0.8
76	0.8	0.8
77	0.8	0.8
78	0.8	0.8
79	0.8	0.8
80	0.8	0.8

Fig. 8

An alternative option consists in the input, besides *QZ*, of F_{max} if estimates of this parameter are available. In this case the user should choose the option 'from the vector' (column F in Fig. 9a) and input the value 2 in A41, then monthly values must be inputted in column Y of figure 9b.

	A	B	C	D	E	F	G
41	Select Fmax	1		1 = from the model		2 = from the vector	

Fig. 9a

	Y
66	Fmax
67	[years-1]
68	
69	
70	
71	
72	
73	
74	
75	
76	

Fig. 9b

3.13 Selection ogive parameters of the gear used by the fleet

Column P68-, V68- (fig. 10a). These values are used to shape the fishing mortality accounting for gear selectivity. These parameters are referred to the gear used by the commercial gear. The $L_{50\%}$ and the range $L_{75\%}$-$L_{25\%}$ are both necessary. Two options are available (fig. 10a), that model the selection process according to an ogive (option 1) or an ogive with a deselection process (option 2). In this case it is mandatory to input also the deselection size D50% (column S in fig. 10b), as the other parameters L75%, D75% and DSR are automatically calculated.

The values of selection parameters can be also changed to simulate different exploitation scenarios and the consequences on the population and yields of different HCRs.

	P	Q	R	S	T	U
57	**Select Gear**	2				
58	1	trawler-ogive	S(L)= 1 / (1 + exp((log(9) / (L75p - L25p) * (L50p - L)))			
59	2	trawler-ogive-des	S(L)=1/(1+exp((2*ln(3)/SR)*(L50%-L))*1/(1+exp((-2*ln(3)/DSR)*(D50%-L)			

Fig. 10a

	P	Q	R	S	T	U	V
66	**Parameter**	**L50%**	**Selection range (L75%-L25%)**	**D50%**	**L75%**	**D75%**	**DSR**
67		[mm]	[mm]	[mm]	[mm]	[mm]	[mm]
68	seed run	80.00	20	550	90	540	20
69	1 month	80.00	20	550	90	540	20
70	2 month	80.00	20	550	90	540	20
71	3 month	80.00	20	550	90	540	20
72	4 month	80.00	20	550	90	540	20
73	5 month	80.00	20	550	90	540	20
74	6 month	80.00	20	550	90	540	20
75	7 month	80.00	20	550	90	540	20
76	8 month	80.00	20	550	90	540	20
77	9 month	80.00	20	550	90	540	20
78	10 month	80.00	20	550	90	540	20
79	11 month	80.00	20	550	90	540	20
80	12 month	80.00	20	550	90	540	20

Fig. 10b

3.14 Fishing coefficient

Column J68- (fig. 5). This coefficient can be set at 0 in case of fishing ban simulation or in a range between 0 and 1 if changes in fishing effort would be simulated (1=no changes in the exploitation scenarios).
It can be also higher than 1, provided that the average along each year is 1.
The fishing coefficient can be also used for simulating modifications in the fishing mortality not due to the selection pattern.

3.15 Parameters of the pdfs for Aladym-q

Column P12-U28 (fig. 2b). Four different pdf distributions are foreseen: the normal, log-normal, gamma and uniform (see the legend in fig. 2c). The user should select the more appropriate for a given variable. The variables to be selected are listed in the column P of fig. 2b, in column Q there is the index of the chosen pdf and in the other columns the parameters needed to identify numerically the distribution.

4. Procedures

Each tool should be run from its own directory in order to allow the automatic preload of its history file, the Excel input sheet must be in a directory on a read/write filesystem (see example DataInputs.xls) as the output file will be written in the same directory.

The instruction to run the simulation is the same for all three tools:

source("src/main.r")

A windows should pop-up to allow the selection of the Excel file containing the input data.
The simulation begins and some information on the step that the tool is executing is displayed.

The results are automatically saved into three files which have the same basename of the Excel file and extension din (to store the input data in text format), dou (to store the results of the simulation in text format) and RData (to store the R workspace for further processing, binary).

The tool *Aladym-r* display also automatically the plots of the main variables calculated (averaged on a yearly timescale).

The tool *Aladym-q*, due to the large quantity of produced results offers an interactive session where the user selects the plot of interest; the session is started entering in the console:

PlotInt()

24 variables are listed, then the user should choose the number associated to the variable and successively the year of the simulation (1, 2…20). The graphs of the probability distributions and of the cumulative pdf related to the variable and to the year will be produced.

Document 7: Simulation Evaluation with ALADYM - Applications

Evaluation of the effects of different pressure and management scenarios by application of Aladym simulation model to case-study stocks

M.T. Spedicato[1], J.-C. Poulard[2], K. Radtke[3], Politou C.-Y.[4], G. Lembo[1], P. Petitgas[2]

[1]*COISPA Tecnologia & Ricerca, Bari, Italy*
[2]*IFREMER, Nantes, France*
[3]*SFI, Poland*
[4]*Hellenic Centre for Marine Research, Agios Kosmas, Greece*

Fishery Independent Survey Based Operational Assessment Tools (FISBOAT),
DG-Fish, STREP n° 502572 (2004-2007)
Co-ordinator: P. Petitgas, IFREMER, Nantes, France

Introduction

Aladym (**A**ge-**L**ength B**a**sed **Dy**namic **M**odel) simulation model, belonging to the group of dynamic pool models (Thompson & Bell, 1934), has been designed to predict, through simulations, the effects of changes of biological (e.g. size at first maturity, growth, recruitment), pressure (e.g. total mortality) and management (e.g. size at first capture, fishing activity) parameters on single fish population dynamics. The model uses the classical equations of the fish population dynamics. This is planned to evolve at a very detailed time scale (1 month), using vectors by age and size and accounting for differences by sex in growth, maturity and mortality. Also the natural mortality can be modulated as a vector by age-length. This is an important issue for Mediterranean fisheries, where fish populations are exploited starting at an early stage in the fish life (e.g. Caddy, 2006a). Removals are simulated on the basis of the total mortality rate modulated using a harvesting pattern and a fishing activity coefficient. The very detailed time scale implemented in the model allows to envisage management action considering biological process such as recruitment and growth in time, that means both along the year and during a transition phase from a 'current' to a new state.

Aladym contains also an harvest control rule thought to be particularly useful in specific situations, when fish are exploited at early stages, management is mainly based on a frame of technical regulations and a TAC regime cannot be implemented for most demersal fisheries ('passive management', according to the definition of Caddy, 2006b).

Thus the foreseen management options are based on gear selectivity and fishing activity that can act alone or in combination. The latter can be for example very useful to predict the effect on the population of a fishing ban or of changes in fishing effort along the year and throughout the whole time of simulation. Though based on non-TAC management context, the model can however give indications for TAC-management situations, when the effects of alternative measures would be explored.

Aladym works with fishery-independent information, although its predictive capability of real catch levels can be verified using data on commercial catches or fishing activity per month. This characteristic has the advantage of making the model useful also where good quality catch data are difficult to achieve for different reasons, including the problems in sampling extreme dispersed and diversified fishing sites and fleets.

Metrics characterising fish population attributes (here termed indicators) were simulated from the *Aladym* model, resulting in the construction of model-based time series of a variety of indicators of stock attributes. The effects of different management strategies were then analysed in terms of the sustainability of the population in the long-term, using indicators as the ratio between the mean exploited spawning stock biomass and the mean unexploited spawning stock biomass (E*SSB/USSB*). Also other candidate indicators as the ratio of ESSB vs EB (exploited biomass) and vs yield have been tested.

The project case studies scanned four different stocks across European waters in the demersal domains with different vital traits, stock histories and survey methodologies. The case studies were: red mullet in the central-southern Tyrrhenian sea, hake in the Bay of Biscay and in the Aegean sea, and cod in the Baltic sea.

The age-length based Aladym model has been thus used to test, through simulations, the consequences of changes of pressure parameters (mortality) and management strategy (e.g. fishing activity, size at first capture) on the fish population dynamics of the target stocks.

These effects have been estimated along 20 or 40-years simulations, depending on the species, analysing the changes of model-based population indicators, i.e. the total biomass, the spawning stock biomass, the biological production (all deaths removed from the population for natural and fishing causes). Consequences on model-derived vital traits indicators (i.e. average length of the population and of the spawning population) have been also evaluated, as well as changes regarding simulated yield and mean length of the catches. The analysis of population indicators (Petitgas et al., 2007) had showed for some of the considered species that Z increased during the last years of the observation period whereas old ages exhibited a decreasing abundance. Hence the importance of simulating scenarios for Z to assess the viability range of the population according to exploitation pressure. The aim of this study was consequently also to get reference values for Z and model-based indicators for a sustainable exploitation (i.e., durable and little variable in time). Finally, effects of exploitation strategies on a sustainability indicator as the ratio between exploited and unexploited spawning stock biomass (ESSB/USSB) were assessed.

Details on the Aladym model (documentation and code) and application to case studies can be found in another section of this report as well as on the Fisboat website.

Materials and methods

Red mullet (*Mullus barbatus*) in the central-southern Tyrrhenian Sea

Aladym model was applied to the red mullet case study following three complementary objectives that can be summarised as follows: 1) analysing the effects of the observed fishing pressure levels (yearly total mortality rates, 1994-2002) on the population; 2) investigating the influence of recruitment mode, i.e. dependent or independent from the parental stock (random recruitment and a stock-recruitment relationship were used); 3) evaluating the reaction of the population to different fishing pressure scenarios, based on changes of total mortality Z.

All these effects were explored using pairwise relationships of the output simulated metrics, selecting those suitable as reference points.

The inputs of the model, fully detailed in the case-study report, were mostly obtained from trawl-surveys (WP2a&b of Fisboat project; national trawl-surveys: Spedicato et al., 2003, 2004, 2006), except for the size at first capture (L_{50}) and selection range (SR) that were from selectivity experiments conducted in the area using a commercial trawl net (Lembo et al., 2002). Considering the early stage of exploitation of red mullet in the area and following an Aladym option, natural mortality variable by length/age was set according to the Chen and Watanabe model, while a guess estimate of longevity was obtained using the Taylor approximation. All the simulations were run for 20 years.

Hake (*Merluccius merluccius*) in the Bay of Biscay

Aladym model was applied to the hake in the Bay of Biscay to investigate the effects of different fishing pressure on the population and explore pairwise relationships of the output simulated metrics. The study period (1987-2004) was always simulated using a constant total mortality set at ~1.04 for the first 14 years and at a higher value (~1.8) from years 15 (2001) to 18 (2004). In the simulated scenario 'Z high', the total mortality remained at this high value (~1.8) during the following 22 years, while in the scenario "Z mean" the total mortality was set to ~1.04, i.e., the mortality value observed during the first part of the study period for the whole simulated period. Finally in the "Z low" scenario, the mortality value was fixed at the three quarters of the mortality used in "Z mean" scenario, i.e. 0.78 from the year 19 onwards.
The inputs of the model, fully detailed in the case-study report, were mostly obtained from trawl-surveys and from literature as regards some biological parameters (ICES, 1991; Jensen, 1996; de Pontual et al., 2006; Murua and Motos, 2006; Murua et al., 2006). Recruitment was assumed independent from the parental stock. All the simulations were run for 40 years.

Hake (*Merluccius merluccius*) in the Aegean sea

Population modelling was based on the total mortality indices estimated within WP2A of the Fisboat project in the study period (years: 1-9, corresponding to the available time series). The value of the last year (year 9; Z~0.9) was then projected forward for 11 years (years: 10-20), while for the final 20 years (years: 21-40) a lower value of total mortality was introduced (~0.78), to simulate long-term effects of a pressure reduction. This was combined with a decreasing of fishing mortality as obtained by modulating two factors: 1) the size at first capture that was increased (L_{50}=108 mm) from year 8 onward as a result of a new cod-end mesh opening (40 mm) and 2) the fishing ban of trawlers in June-September. In the harvesting pattern also a de-selection length was considered accounting for vulnerability/accessibility of the fish to the gear. This was based on the knowledge about the fishing grounds targeted by differently equipped fishing units, accounting for the distribution of adult hakes in the area, which inhabit the deeper waters (Anon., 2006).
The inputs of the model, fully detailed in the case-study report, were mostly obtained from trawl-surveys and from literature as regards some biological (Papaconstantinou & Stergiou, 1995; Papaconstantinou et al., 1998; Karlou-Riga & Vrantzas, 2001; Anonymous, 2002) and technical (Abella and Serena, 1998; Fiorentino et al., 1998; Petrakis et al., 2004) parameters. All the simulations were run for 40 years and output simulated metrics were analysed through pairwise relationships between pressure and population indicators.

Cod (*Gadus morhua*) in the Baltic sea

The Aladym simulation model was used to predict the effects of various fishing pressure scenarios. Input values (growth parameters, stock-recruitment relationship etc.) to the model were obtained from scientific surveys. Options implemented in the model were as follows: gear selectivity (from commercial fleet -unchanged in the simulation), fishing activity (changed according to fishing scenario considered – total fishing ban, periodical fishing ban etc.). Recruitment variability was assumed as +/-20% (on the basis of observed recruitment variations) and total mortality Z (first order approximation equal to the value of Z observed as obtained from research surveys - the outcome from WP2). In each HCRs scenario considered the sustainability of the Baltic cod (eastern stock) population in the long-term context was analyzed. Among the simulations performed the results of a few selected ones were retained for the objectives of the present study.

Results

Red mullet (*Mullus barbatus*) in the central-southern Tyrrhenian Sea

The relationships of fishing pressure parameters (Z and F) vs. population (biomass and spawning biomass of the exploited and unexploited population, related mean lengths and ages) and removal metrics (yield, biological production, mean length/age of catches) showed significant pairwise negative correlation (range: minimum -0.57, maximum -0.94), generally with a 2-years delay. This might be explained considering a cascade effects along cohorts combined with the growth rate of the species, that requires a time lag to be evidenced.
The impact of a high rate of total mortality is well evidenced in the evolution of the selected indicator, i.e ESSB vs USSB that falls down at very low values almost every 5-6 years, when the additive effects of harvesting along cohorts (peaks of Z) were combined with the characteristic of the life cycle, although the contribute of the latter is rather low (about 5% as computed for the unexploited population) compared to the former.
In addition, also the indicator ESSB vs EB was considered, given the advantage to be more easy to understand (which proportion should the biomass of spawners represent for a sustainable exploitation?) and the very high level of positive correlation (range: 0.91-0.99) with population metrics. The two indicators ESSB vs USSB and ESSB vs EB were retained for evaluating the effects of a Ricker stock-recruitment relationship on the population dynamics in comparison

with a recruitment pattern independent from the parental stock. The results highlighted that the ratio ESSB/USSB, under the hypothesis of independent recruitment, was about 50% of the ESSB/USSB (about 34% in case of stock-recruitment relationship with higher density-dependent effect), whilst the ESSB/EB ratio was about 130%.
Assuming a current average rate of total mortality of 2.4 and a stock-recruitment relationship characterised by a higher density-dependent effect, that is plausible according to the knowledge on the species (Levi et al., 2003), we tried to evaluate the effects of changing pressure, from -25% to +25% of the current value, using Aladym-q. Overall results highlight an alert in the current situation and the positive effect of reducing pressure on the population: the probability the ESSB vs USSB ratio would reach levels lower than 0.16, 0.22 and 0.3 is 0.05 at Z values corresponding to 2.8, 2.4 and 2.1.

Hake (*Merluccius merluccius*) in the Bay of Biscay
Among the indicator generated from the Aladym model three candidates: ESSB vs USSB, ESSB vs EB, and ESSB vs yield were selected as descriptor of the population response to different fishing pressure. The second and third ratios are probably more informative, as the adult part of the exploited population was compared to total biomass available or yield. The first ratio is more sensitive to changes in recruitment while the second one allows to illustrate the effect of the fishing pressure, making easier the proposition of a management action that would warrant a more stable catch level. These indicators are thus retained to evaluate the effects of the simulated Z scenarios.
The scenario "Z high" describes changes occurring in hake population if the fishing pressure remains at the level observed during the four last years (2001-2004) of the study period. All indicators exhibited long term decreasing trend although some improvement can be seen some years as a consequence of very good recruitment. Effects of good recruitment were of short duration. All the population production indicators (e.g. yield, biological production, exploited spawning stock biomass, length …) and the ratios ESSB vs EB (mean 0.25) or ESSB vs USSB (mean 0.02) were on average the lowest simulated and lower than ones of 1987-2004 period. The relationships of fishing pressure parameters (Z and F) vs. population (biomass and spawning biomass of the exploited and unexploited population, and related mean lengths) and removal metrics (yield, biological production, mean length of catches) were analysed for this scenario, that presented a more contrasted situation in term of mortality. Significant pairwise negative correlation were found (range: minimum -0.41, maximum -0.86), with 2-years delay for biomass and removal metrics vs Z and F and 3-years delay for vital trait indicators vs Z and F.
In the scenario "Z mean" the fishing pressure was maintained to the mean level prevailing during the first 14 years of the study period. This scenario can be considered as the continuation of the exploitation on the same conditions as during the period of study (ESSB vs EB mean 0.36; ESSB vs USSB mean 0.09). The catch levels observed and obtained by simulation were close. The last scenario, "Z low", allowed to increase on average all the population production indicators and the ratios ESSB vs EB and ESSB vs USSB that reached mean values of 0.41 and 0.17, respectively. The lowest catch was predicted the first year of the simulated period afterwards catches would be higher and less variable than in any other scenarios. Indeed the ratio ESSB vs yield reached an average the value of 1.04 that was almost twofold the value in the "mean" scenario (0.6) and 5 times the level of the "high" scenario (0.22).

Hake (*Merluccius merluccius*) in the Aegean sea
All the relationships regarding Z and population or removal metrics or vital traits had a delay of 3 years, while those regarding F had a delay of 2 years. This might be explained considering a cascade effects along cohorts combined with the growth rate of the species, that requires a time lag to be evidenced. Length-based indicators and population metrics resulted well correlated, except the length mean of the exploited spawning stock versus Z and F and the yield versus Z and F. All the relationships showed negative slopes, as expected. The beneficial results following the decreasing of the total mortality from years 20 onwards becomes evident after the year 20 with a continuous rising phase, due to the cumulative effects along cohorts. At this stage a new and more safe state seems to be reached, as evidenced by the indicator ESSB vs USSB ratio. This assumed values of about 0.07 at initial time and was gradually growing to 0.17 when the mortality was reduced. A similar pattern showed the ratio ESSB vs EB, which values were at level of about 0.4 at beginning and progressively increasing to 0.5, as a result of the beneficial effect on the population of a mortality diminution. More probable values of the ratio ESSB vs USSB, as estimated from Aladym-q, were in the range 0.07-0.12, both at the years 7 and 20, while at year 40 the more probable values were between 0.12 and 0.18.

Cod (*Gadus morhua*) in the Baltic sea
Model-based indicators obtained from the exploitation scenario corresponding to the *status quo* fishery showed low spawning stock biomass (SSB ~61-64 000 t), as compared to reference points (B_{lim}=160 000 t, B_{pa}=240 000 t), which is at present the case of Baltic cod. Therefore in further simulations options with fishing decrease were studied. Thus two months fishing ban for each year during spawning season (July-August) were applied to the *status quo* scenario. It resulted in a very slight increase of length indicators and SSB (~80 000 t) that however did not reach the B_{lim} level. Since ICES for several years has recommended total ban on Baltic cod, in another simulation 2 years total ban was set and after that fishing was continued with the intensity as in *status quo* scenario. The results revealed an increase in length indicators, but after 5 years metrics returned to their initial values. SSB exceeded B_{lim} in year 2 and 3 (~180 000 t) of the simulation but then returned to the value of 80 000 t. Simulation with two years ban and next applying fishing

mortality (F) reduced to 0.3 (as recommended by EU in multiannual plan for Baltic cod) gave much better results, since indicator values increased and remained on the same level in consecutive years. In addition, that exploitation allowed for SSB to rebuild and maintain B_{pa} target. However, exploitation strategy that assumes total ban on fishing might be hardly accepted by fishermen. Therefore, another strategy assuming gradual reduction of F by 10% each year until recommended F=0.3 was examined. Aladym simulation showed that stepwise F reduction would allow in 10 years perspective to obtain SSB equal to B_{pa}. Positive effect of gradual F decrease as compared to strategy implementing 2 years ban was higher average yield (by 5% in 20 years simulation), which would be welcomed by fishermen.
In addition, the Aladym series of simulations showed that the most adequate value of Z for the Baltic cod recovery should be equal to 0.5. Also it seems that Z=0.5 should assure safe stock exploitation in the future and therefore Z=0.5 could be considered as a reference value.

Discussion

The relationships of fishing pressure parameters vs. population and removal indicators showed significant pairwise negative correlation for the examined stocks, generally with 2-3 years delay, depending on the species and indicator. This might be explained considering a cascade effects along cohorts combined with the growth rate of the species, thus all requiring a time lag to be evidenced.
The state of the red mullet population in the central-southern Tyrrhenian sea has been evaluated by previous studies carried out within Samed (Anonymous, 2002) and Medits (Tserpes et al., 2002) projects. In both cases recommendations of reducing pressure and protect recruitment were formulated, although 'direct' indices of abundance did not show any trend, but the time series was short. The analyses conducted in this study underpin the identification of sign of deterioration in the red mullet population and provide converging evaluation with the comprehensive indicator approach based on 'direct' estimates performed in WP5 of the Fisboat project. Thus the outcomes of this study supports the usefulness of coupling evaluations based on indicators directly estimated on scientific survey data with those model-based that are aimed at understanding how changes of biological and pressure parameters affect fish population dynamics and which are the consequences of different management strategies.
As regards hake in the Bay of Biscay, in 2004 a recovery plan for the Northern stock followed up a previous emergency plan. Based on the most recent estimates of SSB and fishing mortality (WGHMM, 2006) ICES classifies the stock as being at full reproductive capacity and being harvested sustainably. SSB appears to have been very close to Bpa over the last 3 years, and F has been around Fpa since 2001. As the growth rate and thus the age determination and productivity of northern hake stocks are uncertain, absolute estimates of SSB and F have to be considered with caution. In the analyses conducted in this study under the hypothesis of a 'mean' scenario (Z lower and constant except for three years) signs of negative changes were identified following the period of mortality increase. Alternate positive and negative changes occurred also as consequence of recruitment fluctuations, becoming these effects more severe when coincident with the fishing pressure intensification. Aladym simulation results confirmed the conclusion of the "Indicator Approach", i.e. "Knowing the worrying state of the stock at the beginning of the EVHOE surveys and as no improvement occurred in recent years, on contrary some deteriorations of the indices for older age groups, it seems necessary to reduce the fishing mortality". In the case of 'high' scenario a continuous decrease, with some fluctuations, of the indicator ESSB/USSB was observed and likely the population still survived because the initial hypothesis was based on the independence of recruitment from parental stock. Thus, a potential option to warrant a sustainable exploitation of the hake population would be to target a value of Z ranging from "Z mean" to "Z low".
One may question why the catch levels observed and obtained by simulation are so close. Is it fluke or reality ? Our understanding is that the hake recruitment recorded over the eastern continental shelf of the Bay of Biscay during autumn groundfish surveys might supply the hake fishery beyond the VIIIab area. This would then imply that part of the F assessed by Aladym model is due to hake migration from VIIIab to neighbouring areas.
The application of the Aladym model on the Aegean hake gives the opportunity to explore further long-term effects of the management measures on the population. Recommendations from previous studies regarding the state of the hake stock in the Aegean Sea, using a non-equilibrium surplus production model fed with Medits data, have stressed an overexploitation condition and the needing of reducing the fishing pressure (Tserpes et al., 2007). The analysis conducted in the present study identifies signs of positive changes after the first 7 years (i.e. around 2000) as consequence of mesh increase. A (slight) reduction of fishing pressure on Aegean hake population would produce in the long-term a positive change, increasing of about ~50-60% the current levels of the ESSB/USSB sustainability indicator. Comparing the above results with those obtained from the indicator approach developed in WP5 for the Aegean hake, they seem to be in a quite good agreement, in particular the CUSUM analysis showed positive changes (mainly in the abundance) after 1998, which then led to a stable situation until 2003.
Single species management strategies assume that the productivity of a stock depends on its current size and reproductive potential. Thus, managers have to control fishing through actions directed at keeping the stock at an adequate level and protect spawners in order to durably obtain good yields. As regards Baltic cod, Aladym model simulations to predict the effects of various management scenarios showed the effectiveness of measures based both on fishing ban and gradual reduction of fishing pressure, allowing to compare the respective benefits in restoring safe levels of SSB for the Baltic cod population in the long-term.

References

Abella A.J. & Serena F, 1998 - Selettività e vulnerabilità del nasello nella pesca a strascico. Biol. Mar. Medit., 5 (2): 496-504.

Anonymous (2002) - Stock Assessment in the Mediterranean - SAMED. Final Report EU Project n° 99/047.

Anonymous, 2006. National Program for Fisheries Data Collection 2002-2006 (ER 1543/2000). Technical report. HCMR.

Caddy, J.F. 2006a. Practical issues in choosing a framework for resource assessment of Mediterranean and Black Sea fisheries. Joint meeting on Stock Assessment Methodology and Workshop on Black Sea Assessments of Pelagic and Demersal Fish Stocks Istanbul 8-10 March, Turkey. 33pp.

Caddy J.F. (2006b). The potential use of indicators, reference points and the traffic light convention for managing Black Sea fisheries. In Lembo G. (ed.). Selected papers presented at the Workshop on biological reference points. Rome, 20‒21 April 2004. *Studies and Reviews. General Fisheries Commission for the Mediterranean.* No. 83. Rome, FAO: 80 pp.

Cotter, J., Mesnil, B., Trenkel, V., Rochet, M.-J., Petitgas, P., Woillez, M., Rivoirard, J., Uriarte, A., Witthames, P., Spedicato, M.-T. and Lembo, G. 2007. Manual of indicators and methods for assessing fish stocks using only fishery independent survey based information. ICES CM 2007/O:27.

de Pontual H., Groison A.L, Piñeiro C. and Bertignac M. 2006. Evidence of underestimation of European hake growth in the Bay of Biscay, and its relationship with bias in the agreed method of age estimation. ICES Journal Marine Science 63: 1674-1681.

Fiorentino F., Zamboni A. & Relini G., 1998 - La selettività della rete a strascico in *Merluccius merluccius* sulla base delle esperienze riportate in letteratura. *Biol. Mar. Medit.*, **5** (2): 465-474.

Karlou-Riga C. & N. Vrantzas, 2001. Fluctuations of seasonal abundance index in the Saronikos Gulf. Proceedings of 10th Hellenic Congress of Ichthyologists, Chania, 18-20 October 2001, 105-108.

ICES, 1991. Report of the Working Group on Fisheries Units in Sub-areas VII and VIII. Int. Council Explor. Sea C.M. 1991/Asses:24, 215 pp

ICES. 2006. Report of the Working Group on the Assessment of Southern Shelf Stocks of Hake, Monk and Megrim (WGHMM), 9 - 18 May 2006, Bilbao, Spain. ICES CM 2006/ACFM:29. 792 pp.

Jensen, A. L. 1996 Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Can. J. Fish. Aquat. Sci. 53, 820–822.

Lembo G., Carbonara P., Silecchia T., Spedicato M.T. (2002) – *Prove di pesca a strascico con rete a doppio sacco per la valutazione della selettività dell'attrezzo e della qualità del prodotto*. I quaderni scientifici di Lega Pesca, Roma: 1-47.

Levi D., M.G. Andreoli, A. Bonanno, F. Fiorentino, G. Garofalo, S. Mazzola, G. Norrito, B. Patti, G. Pernice, S. Ragonese, G.B. Giusto, P. Rizzo 2003. Embedding sea surface temperature anomalies into the stock recruitment relationship of red mullet (*Mullus barbatus* L. 1758) in the Strait of Sicily. Scientia Marina, 67(1): 259-268.

Martin, I. 1991. A preliminary analysis of some biological aspects of hake (*Merluccius merluccius*, L. 1758) in the Bay of Biscay. ICES CM 1991/G:54, 31 pp.

Murua H. and Motos L. 2006. Reproductive strategy and spawning activity of the European hake *Merluccius merluccius* (L.) in the Bay of Biscay. Journal of Fish Biology, 69 (5): 1288–1303.

Murua H., Lucio P., Santurtún M. and Motos L. 2006. Seasonal variation in egg production and batch fecundity of European hake *Merluccius merluccius* (L.) in the Bay of Biscay. Journal of Fish Biology, 69 (5): 1304–1316.

Papaconstantinou C. & K. Stergiou, 1995. Biology and fisheries of eastern Mediterranean hake (*M. merluccius*). In: Alheit J. & T. J. Pitcher. Hake: Biology, fisheries and markets. Chapman & Hall, London, 149-180.

Papaconstantinou C., Petrakis G., Caragitsou E., Labropoulou M., Karkani M., Vassilopoulou V., Mytilineou Ch., Lefkaditou E., Siapatis A., Kavadas S., Chatzinikolaou P., Anastassopoulou A., Kapiris K., Terrats A., Dogrammatzi A., Bekas P., Christidis G. & Fourtouni A., 1998. Development of the Greek fisheries – Assessment of the demersal fisheries resources of commercial interest in the S. Aegean Sea. Technical Report. NCMR.

Petrakis G., R. Holst, A. Chilari & C. Alidromiti,. 2004. Selectivity estimation of trammel nets and bottom trawl cod-end. Final Report. In: Papaconstantinou C. (ed.): Development of an integratedmanagement system to support the sustainability of Greek fisheries resources, IMAS-Fish. HCMR, August 2004.

Petitgas, P., Poulard, J.-C., Radtke, K., Spedicato, M.-T., Ibaibarriaga, L., Politou, C.-Y., Korsbrekke, K., Deernberg, C., and Fernandes, P. 2007. Comprehensive indicator-based diagnostics of fish stocks using fishery-independent survey data: the FISBOAT report on case studies. ICES CM 2007/O:16, 33 pp.

Spedicato M.T., Carbonara P., Lembo G. (2003) – Valutazione delle risorse demersali dal Fiume Garigliano a Capo Suvero nel triennio 2000-2002. Relazione Finale GRU.N.D. (L.N. 41/82) U.O. 5, COISPA Tecnologia e Ricerca, Bari: 146 pp.

Spedicato M.T., P. Carbonara, T. Silecchia & G. Lembo. 2004. Strategie di gestione basate sulla biomassa di riproduttori. Simulazione di diversi scenari di prelievo della triglia di fango (*Mullus barbatus* Ll., 1758). *Seminario per l'identificazione di Reference Points. MIPAF-SIBM, 28-29 January 2004.*

Spedicato M.T. (coord.). 2006. GRUND 2006 – Relazione finale Sub-area Geografica (GSA) 10 – Medio e Basso Tirreno. EU Reg. 1543/00-Mipaf. 64 pp.

Tserpes G., Fiorentino F., Levi D., Cau A., Murenu M., Zamboni A., Papaconstantinou C. (2002) - Distribution of *Mullus barbatus* and *M. surmuletus* (Osteichthyes: Perciformes) in the Mediterranean continental shelf: implications for management. In: Abelló P., Bertrand J. A., Gil de Sola L., Papaconstantinou C., Relini G., Souplet A. (eds.), Mediterranean Marine Demersal Resources: the MEDITS International Trawl Survey (1994-1999). Sci. Mar., 66 (Suppl. 2): 39-54.

Tserpes G., J. Haralabous and C. Maravelias, 2007. A non-equilibrium surplus production model approach using Medits data. GFCM-SAC-Sub-Committee Stock Assessment. Workshop on trawl survey based monitoring fishery system in the Mediterranean, Rome, Italy, 26-28 march 2007. 4 pp.

Thompson W.F., Bell F.H. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. *Rep. Int. Fish. (Pacific Halibut) Comm.*, (8): 49 p.

Document 8: Comprehensive Assessments

Proposition On How To Create Comprehensive Stock Assessments Based On Several Assessments And A Combination Of Quantitative And More Qualitative Results.

Deliverable 5.2 for EC research project:

Fishery Independent Survey Based Operational Assessment Tools (FISBOAT),

DG-Fish, STREP n° 502572 (2004-2007)

Co-ordinator: P. Petitgas, IFREMER, Nantes, France

authored by
J. Cotter, CEFAS, Lowestoft, Suffolk (john.cotter@cefas.co.uk)
P. Petitgas,IFREMER, Nantes, France
A.J. Arbella, ARPAT-AREA MARE, Livorno, Italy
B. Mesnil, IFREMER, Nantes, France
V. Trenkel, IFREMER, Nantes, France
M-J. Rochet, IFREMER, Nantes, France
P. Apostolaki, CEFAS, Lowestoft, Suffolk
M.T. Spedicato, COISPA Tecnologia & Ricerca, Bari, Italy
M. Woillez, Ecoles des mines, Fontainebleau, France
J. Rivoirard, Ecoles des mines, Fontainebleau, France
C-Y. Politou, Hellenic Centre for Marine Research, Agios Kosmas, Greece

Contents

1. Introduction....................274
2. Using FISBOAT methods alongside existing methods of fish stock assessment..277
3. Proposition for comprehensive stock assessments using FISBOAT methods.......279
3.1 Introduction....................279
3.2 Tasks....................279
3.2.1 Agree management objectives....................279
3.2.2 Select biological and spatial indicators....................280
3.3.3 Select indicators of fishing pressure....................281
3.3.4 Select survey type and design....................282
3.3.5 Select reference period and reference values/trend directions for indicators 283
3.3.6. Select statistical methods relating to individual indicators....................284
3.3.7 Select statistical means of combining indicator results....................284
3.3.8 Select survey-based assessment method....................285
3.3.9 Agree management responses to good and bad assessment results.........285
4. Experiences with case studies....................287
4.1 Barents Sea cod....................287
4.2 North Sea cod....................287
4.3 North Sea herring....................287
4.4 Western (Biscay) hake....................287
4.5 Biscay anchovy....................287
4.6 Central Mediterranean red mullet....................288
4.7 Baltic cod....................288
4.8 Eastern Mediterranean hake....................288
Aegean Sea....................288
Ionian Sea....................289
4.9 Comments....................289
Table 1. FISBOAT project: Biological indicators....................291
Table 2. FISBOAT project: Spatial indicators....................292
Table 3. FISBOAT project: Methods for integrating and interpreting indicator series 293
Table 4. FISBOAT project: Fishery-independent assessment models....................295
Table 5. FISBOAT project: Details of case studies....................296
References....................298

1. Introduction

Many currently used fisheries assessment methods depend heavily on fisheries landings- or catch-at-age data[1]. The FISBOAT project (Fishery Independent Survey-Based Operational Assessment Tools) happened partly because of problems with fishery-dependent data:

- Landings data may not be available due to the geographic spread of landing points, or due to area closures.
- Landings data may be distorted by quota restrictions, poor markets, or mis-reporting.
- Landings data may be poorly representative of total catch because of discarding at sea.
- Landings data for stocks that have become of interest as others have declined may not be broken down into age groups as required for VPA-based stock assessments.

Concerns about the adequacy of official landings statistics in Europe have existed for a long time (Anonymous 1986). Additionally, stock assessments use data for landings- or catch-per-unit-of-effort (LPUE or CPUE, respectively) recorded for commercial fishing vessels. These data, of course, have all the same problems as official landings statistics but, in addition, may provide poor and misleading information about abundance because of fishers' adeptness at finding fishable concentrations of fish even when stocks are low (Rose and Kulka 1999).

Fishery-independent abundance indices can be obtained from trawl surveys carried out by research vessels, specially chartered fishing vessels, or commercial vessels operated in partnership with scientists (Armstrong *et al.* 2007). The vessel fishes the area occupied by the stock(s) of interest at least annually according to a fixed protocol with a standardised trawl, usually equipped with a small-mesh codend having a reasonably constant catchability for most size classes. Fish surveys can also be carried out using acoustic methods, and using plankton nets to catch eggs and drifting larvae. Both of these are likely to be supplemented by standardised trawling, to identify species during an acoustic survey, or to obtain samples of adults to estimate reproductive parameters during a planktonic egg survey. The objectives of FISBOAT were to develop tools based on any type of scientifically conducted survey in order to allow fishery-independent assessments, and to evaluate the ability of the developed tools to provide quantitative advice on management options.

Fish survey data have their own weaknesses. For a recent review of many aspects, see Anonymous(2004a; 2005a); for standardisation of trawl gear see Anonymous(2006); and for problems associated with the noise made by survey vessels see Mitson (1995). Trenkel (FISBOAT report) illustrates special problems when using survey data as the unique source of data for carrying out a stock assessment, namely mismatch of survey area and stock area, selective catching of size and age classes that are not representative of the size or age structure of the stock, and variation of survey catchability (Trenkel 2007).

A likely direction for reform of the CFP is away from the management of single-species stocks and towards management of whole ecosystems, the so-called

[1] Catch ≥ landings + discards, the latter being relatively important in many European fisheries.

Ecosystem Approach to Fisheries Management (EAF) or Ecosystem-based management (EBF) (Ward *et al.* 2002; Garcia and Cochrane 2005). The need for an EAF is now accepted by many scientists (Rice *et al.* 2005), non-governmental (Ward *et al.* 2002), and governmental organisations (Commission 2001a; FAO 2003). Now that so many fisheries are over-fished (Commission 2001b), such an approach holds promise that exploitation of commercial fisheries can be sustainable and without undesirable environmental side effects such as have been seen and cogently interpreted for the Black Sea (Daskalov 2007; Daskalov *et al.* 2007). An EAF also embraces social and economic needs (Commission 2001c).

Survey-based assessment methods appear to have a major role to play in an EAF. Surveys can provide time-series for numerous and varied indicators of the health of commercial fish stocks, non-target species of fish, as well as other ecological components (though the latter were not part of the FISBOAT project). Discussions are currently in progress for reforming the European Commission's Data Collection Regulation (EC 1639/2001) so that data are collected on non-target and ecological components of the sea as well as on commercial fisheries. These changes should permit the development of new time-series of indicators from fish surveys as well as from fishery-dependent sources. Surveys can also support single-species stock assessments of the type comparable with those currently carried out with fishery-dependent data. Survey-based stock assessment methods considered for Deliverable 3.1 of the FISBOAT project are summarised by Mesnil et al. (2007).

The purpose of the present section of the final report of the project is to assemble the results of FISBOAT research in order to indicate how a comprehensive stock assessment might be prepared when most of the available information about the stock is derived from one or more surveys. The text is kept general because every individual stock and ecosystem is likely to make its own special requirements that cannot all be dealt with in one document. We first comment on how survey-based indicators and methods might be applied in parallel with existing stock assessments. Information could thereby be added to improve precision and understanding, and such a parallel approach might assist a transition from one system of advice and management to another. The bulk of the text is devoted to a scheme for assessing stocks using only survey-based results; fishery-dependent data might be added to the analytical assessment but that possibility is not treated here because it was not considered under FISBOAT (except as part of the development of the FLR simulation framework which is general for all types of data). Finally, experiences from applying the FISBOAT methods to 8 case studies representative of a diverse range of European marine fisheries are summarised. These indicate, as might be expected, that the informativeness and applicability of the different indicators and methods depends heavily on the special circumstances of the fishery and the survey. The literature on ecological assessment of fisheries is extensive and well reviewed already so citations are restricted to papers that fill gaps in our research or that can help readers find the roots of the various ideas.

The biological and spatial indicators, and the methods for interpreting the results that were developed or considered during the FISBOAT project are documented in the accompanying FISBOAT report, the 'Manual of Indicators and Methods for assessing fish stocks using only fishery-independent, survey-based information' (Cotter *et al.* 2007). The biological and spatial indicators are summarised here in tables 1 and 2, respectively, and the methods in table 3, so that the present document can be read without necessarily referring to the Manual. The methods of stock assessment developed specifically for use with survey data are documented in another accompanying FISBOAT report, ‘A review of Fishery-Independent assessment models, an initial evaluation based on simulated data’ (Mesnil *et al.* 2007). They are summarised here in table 4.

2. Using FISBOAT methods alongside existing methods of fish stock assessment

Currently, scientific advice supporting the management of European marine fisheries is formulated by ICES fish stock working groups using well known models fitted to fisheries-dependent and/or fisheries-independent survey data. FISBOAT, on the other hand, developed software and documentation for carrying out diagnostics of fish stock abundance and status as well as advice on management strategies, using only fishery-independent information from research surveys.

Three categories of methods were developed under FISBOAT: (i) assessment models, (ii) monitoring procedures based on indicators of stock attributes, and (iii) simulation evaluation tools to investigate appropriate management strategies. The two first categories provide relative information concerning the state of a stock. The assessment models use survey-based indices of abundance only, while the indicator-based methods use a wide range of indices of abundance, vital traits, and spatial occupation. These tools can be tasked to assess whether management actions have effectively allowed the stock to go in the desired direction within limits having known statistical risks of false alarms and non-alarms. Complementary to the FISBOAT diagnostic tools are the FISBOAT simulation tools. Once the fish stock is diagnosed to have followed a particular evolution in its abundance and vital traits, the projection of the population into the future under a regime of defined harvest control rules can be simulated, allowing quantitative testing of the available management options. The simulation tools thus provide the basis for advising on management action.

The FISBOAT methods for carrying out, or supplementing stock assessments using only fishery-independent abundance indices (CPUE) from fish surveys are shown in table 4. They all performed satisfactorily with simulated stock data (Mesnil *et al.* 2007). The opportunity exists, therefore, to fit one or more of the FISBOAT models to survey-only data as support for current methods of stock assessment. For example, it may be possible to obtain useful results with surveys that were rejected for tuning purposes in this way. The options are to use LENSUR for length-structured assessments, BREM for modelling biomass, and TSA, SURBA, or YCC for age-structured assessments. YCC could also serve for screening data prior to using them in the main assessment model, or to estimate gradual changes in the apparent total mortality or survey catchability between ages and years. The FISBOAT model and the main assessment model (e.g. XSA) should not use the same survey data. Otherwise, a degree of corroboration could be expected just as a result of the common information and variability (Cotter *et al.* 2004).

FISBOAT also documented a range of indicators suitable for assessing the biological state of a stock and its geographic distribution, together with various methods for interpreting the resulting time-series of results. Use of these ideas alongside existing stock assessment methods is recommended to enhance understanding of the stock, especially as many past assessments have prioritised quantitative, rather than biological aspects. An assessment is likely to be more robust and biologically safe if it is based not just on abundance indices but also on a larger set of stock attributes including biological and spatial indices. For instance, in a particular year, the recruitment index may increase which, considering abundance indices only, could justify increasing TAC. But, if the spatial occupation continues to depart from its reference position, and biological indicators imply decline in the health of the stock despite the better recruitment, an increased TAC may be seen as excessively risky. Parallel use of FISBOAT ideas might also assist a smooth transition between existing single-species assessments and a more general, ecosystem-based approach to management.

The FISBOAT simulation tools were the ALADYM simulator and the FLR simulation evaluation loop. Both were based on simulating an underlying biological population model and testing different harvest control rules to investigate the domain of sustainability of the harvest. The FLR loop also allows the robustness of the harvest rules to be tested when errors and biases are present in the information about the stock, or when the rule is not complied with (simulating cheeting). Testing of different types of HCRs, e.g. protecting juveniles, area closures, etc., might also be possible though some may require re-formulation of the operating model. Use of these tools alongside existing stock assessment methods could enhance robustness in the advice against uncertainties.

3. Proposition for comprehensive stock assessments using FISBOAT methods

3.1 Introduction

The FISBOAT project concerned survey-based methods for assessing stocks. However, since its inception, the political impetus for an ecosystem approach to managing fisheries seems to have firmed, judging from the large numbers of publications appearing on the subject. The following proposition on how to create comprehensive stock assessments therefore assumes that all information is derived from surveys and that some sort of ecosystem approach is required. Readers should not infer from this a view that use of fishery-dependent data is necessarily wrong (Cotter and Pilling 2007), or that an ecosystem approach, whatever its definition ultimately turns out to be, is the only worthy way to manage a fishery.

Another assumption concerns the nature of the body managing the fishery. Political and organisational options for this task are discussed fully in the collection of papers resulting from the EC EFIMAS project (Motos and Wilson 2006). There are many possibilities ranging from the present 'command and control' system, through regional councils and co-management, to rights-based management. It is envisaged that the management body referred to in the following text may include members of the fishing industry, scientists, other interested professionals such as economists and sociologists, politicians, and non-governmental organisations. This is in conformity with recommendations concerning the EAF (Commission 2001b; FAO 2003). A consequence of envisaging a body drawn from such a wide range of skills is that scientific results about the fishery should be expressible simply and, preferably, visually.

3.2 Tasks

3.2.1 Agree management objectives

Managers need objectives. The top tier of objectives for managing a fishery are probably best decided by political processes outside the management body so that its time is not taken up with arguments among competing interests, e.g. for more fish to be harvested, or for more conservation. FAO (2003) offers suggestions for top level objectives:

- keep harvested species within ecologically viable stock levels by avoiding overfishing and maintaining and optimizing long-term yields;
- maintain habitats and populations of non-retained (by-catch) species with ecologically viable levels;
- keep impact on the structure, processes and functions of the ecosystem at an acceptable level;
- maximize net revenues; and
- support regional employment

Another possibility would be to

- achieve a significant reduction in the rate of biodiversity loss (Jennings 2005)

Fishing interests will probably wish to add explicit objectives for financial yields. They do not automatically translate into yields of fish since market prices tend to go up as quantities landed go down. Economic arguments of this sort could play an important role in ameliorating fishing pressures.

The top level objectives have to be translated into operational objectives for the work of the fishery management body. The operational objectives are likely to have to deal with a wide range of fisheries impacts and, under an EAF, should be directed towards achieving sustainability. Prioritisation of objectives is necessary to prevent an impractical proliferation of them (Jennings 2005). An example of an operational objective designed to meet the first political objective in the bulleted list above might be 'To maintain 10% of the stock at age 3 or older". Each operational objective should be directly addressable by scientific means, for example using one or more indicators of the type shown in tables 1 and 2.

3.2.2 Select biological and spatial indicators

Guidance on the selection of indicators for managing a fishery is provided by Jennings (2005), Rice and Rochet (2005), Rochet and Rice (2005), and other papers in the conference proceedings edited by Daan et al. (2005). Published studies of the performance of selected indicators in fished situations are also available (Piet and Jennings 2004). Rice and Rochet (2005) argue that the number of indicators chosen should be minimal to prevent conflicting signals and arguments. Most of these writers were probably unaware of the spatial indices recently developed under FISBOAT, see table 2. There can often be reasons to expect that the geographic distribution of fish stocks will change in response to fishing pressures, or to variations in oceanographic conditions or climate. Spatial indicators therefore provide another way of looking at a fish stock. Usefully, those shown in table 2 are unaffected by zero catches which can distort comparisons of geographic distributions over time when using spatial indicators that do not allow for zeros.

The biological indicators listed in table 1 relate to most key biological processes including growth, condition, maturity, reproduction, abundance, and mortality. There are of course hundreds more indicators to choose from (Methratta and Link 2006). The first step in selecting indicators is the identification of fishing impacts most likely to compromise attainment of the operational objectives. These are then prioritised according to severity and likelihood of impact, and state indicators relevant to impacted components of the ecosystem selected, depending on resources (Jennings 2005), and seasonal timing of the survey relative to the biological processes. Different interest groups in the management body are likely to favour different indicators. Simulation studies using the ALADYM age-length-based model, table 3, or the FLR system (Kell *et al.* 2007), both developed under the FISBOAT project, may provide an objective evaluation of which strategy is likely to bear most fruit (De Oliveira *et al.* in press). Several simulation studies concerned with the performance of ecological state indicators have been reported recently (Fulton *et al.* 2005; Hall *et al.* 2006; Methratta and Link 2006; Travers *et al.* 2006).

A further consideration when selecting biological indicators for fish stocks and ecosystems is whether they can be adequately sampled by the available surveys. Species that are poorly caught by the survey gear will occur infrequently in survey catches and are likely to display a high variance from station to station.

Consequently, the data make a poorly informative indicator series. There is more discussion of compatibility between indicator and survey in section 3.3.4 below.

A high degree of functional independence among indicator series implies that they are measuring different processes with minimal overlap and redundancy. As examples, indicators of growth and condition are likely to show a high degree of functional dependence because of a mutual dependence on adequate food, while growth and, say, age composition are likely to be more nearly independent because the latter depends on the additional factor of mortality. The amount of dependence acceptable among indicator series itself depends on the extra work and expense created by each, and on whether they inform about relevant, different aspects of a process.

Sampling independence is just as desirable as functional independence but is likely to be harder to achieve when, as is usually the case, different indicators are measured from the same nets (or other sampling devices) on the same surveys. To understand this, imagine sending out a different survey team and vessel for each indicator, supposing that were an affordable option: the indicators would then be sampled independently (or more nearly so) and variances among series would be higher because results would include the between-teams variance and give a truer picture of the reliability of, and the relationships among the different series.

3.3.3 Select indicators of fishing pressure

Jennings (2005) points out that ecological state indicators are inadequate by themselves for managing a fishery; another important class of indicators measures the fishing pressure being applied to an ecosystem. Less research attention has been given to this class of indicator but recent papers by Piet et al. (2007) and Hiddink et al. (2006) provide different perspectives on the pressures of trawling effort. Ideally, the links between fishing pressure indicators and ecological state indicators will be well understood so that fishing can be managed in relation to properties of the ecosystem (Jennings 2005) but, if the links are not well understood, some sort of adaptive management system would be needed as discussed further below. Pressure indicators can be studied with similar statistical methods as state indicators, for example, those listed in table 3. Canonical correlation (Rencher 1995; Everitt 2005) is an obvious choice of statistical method to consider when trying to identify the most important links between suites of pressure and state indicators but an example of a research study was not known to the FISBOAT project team.

For short lived species caught early in their life, density-dependent phenomena may condition growth and natural mortality rates and, consequently, any indices based on them. These processes can have as much effect as increasing fishing effort, as is known to happen for red mullet and several species of bivalve molluscs in the Mediterranean, and, almost certainly, elsewhere. Caddy (2004) states that “Indicators of growth rate and mortality should ideally be quantified through time (i.e. under different densities) and in space (by fishing ground) to evaluate variations in density-dependent processes and habitat quality”. Moreover, considering that in the Mediterranean area, mainly due to the small mesh size of the utilized trawl net and the high fishing pressure, age 0 predominates by many times in survey and commercial catches, as well as in the stock. Consequently, changes in recruitment strength may have a strong influence on mean size.

3.3.4 Select survey type and design

Many existing fish surveys serve many purposes (Ehrich *et al.* 2007) but were primarily designed to support or 'tune' fish stock assessments based on models of the numbers of fish landed by commercial fleets. A different situation exists in Mediterranean areas, where monitoring of landings was sparse at least until 2002 when the EC Data Collection Regulation started to be applied, and fish surveys served also as assessment tools (Abella et al., 1999; Anonymous, 2002). Ecosystem based management of fisheries using survey-based models and indicator series might benefit from different emphases in the designs of these various surveys. However, immediate, radical changes are unlikely because of the general wish in fisheries science communities to retain temporal continuity in survey results. The easiest modification to consider without damaging continuity is the deployment of additional fishing or other devices such as a standardised 2-metre beam trawl towed for 5 minutes at each station fished with the main gear (Callaway *et al.* 2002a; Callaway *et al.* 2002b). Benthic grabs (Rees *et al.* 2006; Rees *et al.* in press), acoustic equipment (Greenstreet *et al.* 1997; Mackinson *et al.* 2004; Mackinson and van der Kooij 2006), and plankton collectors (Beaugrand 2005) could also be considered, as could systematic observations of sea birds and marine mammals from the decks of survey vessels.

Given that one or more standard, unmodified surveys provide the only source of information for managing a fish stock, the limitations of those surveys for that purpose must be carefully considered beforehand. Trenkel (2007) provides examples to illustrate that:

- The survey should encompass the distribution of the whole stock, particularly if it is a mobile species such that the proportion inside the survey area varies from year to year, e.g. with abundance.
- The size and age classes sampled should give unbiased impressions of the length and age frequency distributions of the stock. Problems could arise because the distribution of fishing stations within the survey area gives a biased impression of geographic variability, or because the size selectivity of the gear is not reasonably constant from small to large fish.
- Survey catchability for the species should be reasonably constant geographically and temporally.

Numerous other technical issues (Anonymous 2004b; 2005b) should also be considered before placing heavy reliance on the results of a survey . As a first guide, alarm bells should ring for any species that is not consistently caught where it is expected, when migrations – either horizontal or vertical – perhaps combined with variations in the timing of the survey are likely to cause substantial variations in abundance indices, or when catchability is likely to be related to abundance of the stock due to contraction around favoured locations when numbers decline. Note that these reasons imply that the precision of survey-based indicators for a stock may decline significantly with its declining abundance. Another point to consider concerns the statistical power of a survey to detect future trends. Precision must be adequate to detect an undesirable trend in indicator values. A method for this is summarised in table 3, method 5.3.2.

In the Mediterranean Sea, fishing fleets, allocated along the narrow continental shelves, generally exert their fishing pressure near the ports and hence, a stock may show different levels of abundance and demographic structure over relatively short distances along these narrow stripes. In consequence, the spatial coverage of the surveys can include quite heterogeneous situations. Analyses should only be done with data proceeding from the tows performed within the areas exploited at the same rate. If this is not done, the analysis of time series becomes problematic, for instance for trends in Z. For instance, consider a region where no fisheries-dependent data are available. Information from a trawl survey may include data from 2 contiguous areas, e.g. a northern portion that is exploited at a quite high rate and a southern portion that is lightly exploited. The results will consist of an average value with no practical use for management purposes. In fact, results and derived advice will tend to penalize the fleet that exerts a lower fishing pressure in the Southern portion even though it is likely to be able to support heavier fishing pressure in the future.

Certain indicators are very dependent on the seasonal timing of a survey. Indicators of reproductive capacity, e.g. SSN, GSI, LaM50, AaM50 (table 1), are mostly best measured just before the breeding season of a species when gonad development is most advanced. Supplementary histological examinations of fish can permit more flexibility in the timing but they add considerably to the workload. Similarly, indicators of growth and condition, e.g. C, L25 (table 1), are likely to display seasonal highs and lows, especially for small, young fish that respond especially rapidly to good feeding conditions. Spatial indicators (table 2) may also vary with season depending on migrations, and perhaps with mortality. If the season of a survey cannot be altered, the choice of indicators should be restricted to those that are compatible with the time of year.

Those finding themselves with the challenge of designing a survey specifically to manage a fishery without relying on other data would need to keep the foregoing points in mind but, in addition, should ask what other design features would enhance precision of the chosen indicators. Fixed station designs are vulnerable to contraction of the stock to locations in between fishing stations. This would be a case of bias varying with abundance. The migration-related bias of a fixed station designs may itself also vary from year to year depending on hydrographic or seasonal conditions. Randomised designs, if they are feasible practically, should be free of bias but may show high variance and thus be inefficient, particularly for the less common species. Stratified or adaptive designs may be efficient for priority species but can make the survey very inefficient for others. In short, there are no easy answers about design of a survey intended to provide information for many indicators.

3.3.5 Select reference period and reference values/trend directions for indicators

Reference points are values chosen on best available information to help managers to decide whether the level of an indicator signifies that a stock or an ecosystem are in a good state, or a bad state needing corrective action. The points may be chosen to signal for example, “no impact”, a precautionary “need for corrective action”, or a limiting “need for extreme action” such as closure of the fishery. Deciding these values in advance of any problems and possibly without extensive experience of monitoring the ecosystem may prove difficult and subjective. Simulating the fishery under known conditions representing low and high fishing pressures combined with different levels of observation error could assist; the ALADYM model (table 3, methods 5.2.1 and 5.2.2) and the FLR system (Kell *et al.* 2007) developed under FISBOAT are two possible methods for doing this. Reference points and directions are certainly matters to be negotiated with all interests represented on the fishery management body. Jennings (2005) discusses reference points in more detail, pointing out that hitting targets may be better policy than avoiding limits, and that reference directions (e.g. towards improvement) may be easier to find and agree than absolute values or slopes.

Finding reference points or directions is often greatly assisted by knowledge of a period when fish stocks or the ecosystem were in an acceptable condition, probably at a time of low fishing effort, or before the fishery matured and yields per unit of effort started to decline substantially. Values or trends of each indicator may then be interpreted relative to this so-called 'reference period' taking into account elementary facts about the biology and behaviour of the species (table 3, method 5.5.1). In some cases, research results already permit convincing explanations of what is happening in the ecosystem, as for some groups of size-based ecosystem indicators (Shin *et al.* 2005). Since such understanding greatly facilitates intelligent, adaptive managerial actions that are most likely to be agreeable by different interest groups on the stock management body, research directed at improving understanding of the basic biological and spatial processes appears to offer much promise for successful stock management based on indicators.

3.3.6. Select statistical methods relating to individual indicators

Indicator series derived from fish surveys are likely to be affected by considerable sampling and measurement variance so a statistical approach to assessing compliance with pre-set reference points or reference directions is more or less essential. Several methods have been documented to assist comparisons:

- Recent trends are likely to be of most interest; they can be assessed from second derivatives of the smoothed series (table 3, method 5.3.1).
- Among industrial quality assurance schemes, the Cusum method offers a sensitive method for checking whether fishery and environmental quality indicators are behaving as expected (table 3, method 5.3.3). A reference period is essential.
- Nonparametric statistical methods can be used to assess prevailing levels and overall trends without using models and with the minimum of assumptions, making them relatively objective and easy to explain (table 3, method 5.3.4).

3.3.7 Select statistical means of combining indicator results

Management of a fishery will probably need many indicators. Two problems result. One is how to understand collectively the many, possibly different signals about the stock and the ecosystem; the other is to recognise that results of some indicators are linked with results for others, either through a functional dependence, or through a sampling dependence arising because the material upon which indicator values were measured was obtained on the same surveys or with the same hauls.

One approach to both these problems is to form new composite indices from groups of individual indices. Principal components analysis (PCA) weights the different indices so that the weighted averages can be used as composites that are statistically independent. The first two or three principal components – orthogonal axes through the data cloud – usually explain most of the variability and can be very helpful for understanding the signal from groups of correlated indicators. See table 3, method 5.4.1 in which indicators are composited into a multivariate distance from a central reference point so allowing a new perspective on evolution of the stock. Multi-factorial analysis allows PCA applied to each of several fishing stations to be combined. This can be useful for following changes in the prevailing spatial location of a species as individuals get older. See table 3, method 5.4.2. Min/Max autocorrelation factors (MAFs) are used to combine a set of indices thought to be representative of a fish stock into components (factors) that present maximal continuity in time. The trends extracted do not allow reference periods to be defined because of their continuity. On the other hand, the continuity allows the status of the stock to be monitored in time (Woillez *et al* 2007).

Another approach to the problems of multiple indicators is simpler and amounts to a preparation of a systematic diagnosis about the stock from the different individual signals. The cause-effects table has already been mentioned. It attempts to stimulate biological interpretations of joint results. See table 3, method 5.5.1. A simple visual tabulation of results for groups of variables using red, orange, and green colouring to indicate perceived harm, insignificance, or benefit, respectively, to the stock and ecosystem is called a 'traffic light table'. It can reveal at a glance whether a stock is in difficulty or not. An application to Cusum results is summarised in table 3, method 5.5.2. A third diagnostic method again uses PCA but, in this approach, to summarise independent groups of indicators so as to assist understanding of the signals underlying them. See table 3, method 5.5.3.

3.3.8 Select survey-based assessment method

In addition to providing indicator series for managing a fishery, surveys can also support variants of various stock assessment models. This has already been commented upon in connection with provision of survey-based assessments in parallel with assessments based on landings statistics. The survey-based methods estimate CPUE (rather than absolute stock numbers), total mortality *Z*, and annual recruitments in a relative sense. These standard outputs are likely to be valued by fisheries biologists and others used to interpreting fish stock models and wishing to use them as complementary to indicator-based methods. They might also serve as indicators in their own right. The options are to use LENSUR for length-structured assessments, BREM for modelling biomass, and TSA, SURBA, or YCC for age-structured assessments. Vigilance should be maintained to prevent the same survey results for CPUE and *Z* being used twice, once to support a survey-based assessment, and once again to provide indicator series. This practice would lead to spurious relationships between signals coming from the two methods, and to double inclusion of any sampling and measurement errors (Cotter *et al.* 2004).

3.3.9 Agree management responses to good and bad assessment results.

Few fish surveys provide absolute estimates of numbers or weight of fish in a stock; they only provide relative estimates from year to year or place to place. This is because the catchability coefficient relating CPUE and population abundance locally is seldom known even approximately for trawl surveys[2]. The disadvantage of relative estimates of quantity is that they cannot be used directly to recommend total

---

[2] Exceptionally, some acoustic surveys are considered to give absolute estimates Gjøsaeter, H., Bogstad, B. and Tjelmeland, S. (2002) Assessment methodology for Barents Sea capelin, Mallotus villosus (Müller). *ICES Journal of Marine Science* **59**, 1086-1095..

allowable catches (TACs) that correspond to proposed absolute levels of fishing mortality, *F*, as carried out annually to manage many stocks under the present European Common Fisheries Policy (CFP). Note, however, that the situation is different in the Mediterranean where fisheries are regulated mainly by control of effort and technical measures.

TACs based on *F* became important for managing many international fisheries partly because they are easy to divide up by political agreement among nations, and partly because *F* is a readily estimated parameter in VPA and other stock assessment methods based on fishery-dependent data (Pope 1979). However, whether or not such TACs actually achieve control of the absolute numbers of fish in a stock, even if accurately implemented according to the recommendations of scientists, is now debatable. Only the total mortality, *Z*, is directly estimable - from the decline in numbers in each year class over time. *F* must be estimated indirectly by subtracting natural mortality, *M*, from *Z*, and, since *M* is one of the hardest fishery parameters to estimate (Vetter 1988; Hewitt and Hoenig 2005), significant bias in estimates of *F*, stock numbers, and thus TACs would not be surprising (Rivard 1989), even if bias in the market sampling data and stock assessment models (Cotter *et al.* 2004) is somehow successfully avoided. Furthermore, TACs can fail to maintain a stock even if implemented perfectly (Kell *et al.* 2005). The TAC-*F* system is now widely acknowledged to have a variety of disadvantages (Beverton 1998; Demaré 2006) and reform of the European Common Fisheries Policy (CFP) is being actively discussed (Commission 2001b). It follows therefore that superiority of the supposedly absolute TAC management system over relative, survey-based methods should not be presumed.

Given acknowledgement that management of the absolute size of stocks is not feasible, some sort of organised trial and error procedure, generally known as adaptive management (Walters 1986) becomes necessary. Jennings (2005) proposed that management should focus on fishing activities that are most likely to cause unsustainable impacts in the ecosystem; the relationships of current values of priority indicators to reference levels or trajectories provides guidance on the actions to take. Survey-based assessments may also contribute. Having decided whether there is a problem or not, the management body then has to decide how to adjust controls on fishing effort or harvesting. ‘Harvest control rules’ (HCR) is the name applied to the limits set on fishing by judgement and agreement among fishers and managers and without presuming a link with absolute stock size. A variety of HCRs exists with various properties and rules for adjustment (see Bogaards report in FISBOAT). These can be tested using simulation under the FLR system (Kell *et al.* 2007) or the ALADYM simulation model, both developed under FISBOAT, and are likely to form an important part of the management strategy adopted. Technical measures, e.g. mesh regulations, closed areas, by-catch rules, etc., are likely to form another important part. They may be particularly suited for controlling specific damage on the ecosystem, e.g. to reefs, or nursery areas, as well as being used to control harvesting of the target species. All controls would have to be reviewed and re-negotiated on a regular basis but not necessarily every year if monitoring results cannot provide clear trends, as distinct from possible noise, within that period. Increasing the frequency of surveys from annual to multi-annual, possibly with reduced numbers of stations being fished on each occasion to minimise extra costs, could be a useful strategy for picking up and acting upon signals available from priority indicators earlier.

4. Experiences with case studies

FISBOAT case studies covered the 8 diverse European stocks and surveys whose details are presented in table 5. They were selected to represent a varied range of European fish stocks that have been assessed by ICES for many years. The national survey data are collected under the international coordination of ICES and/or the European Commission (EC regulation No 1639/2001). Most of the indicators and methods listed in tables 1 to 4 were applied in each case, when practicable. The results are reported fully in Final Report Document 'Case Studies'. Here we summarise selected experiences from practical application of the indicators and methods to these case studies.

4.1 Barents Sea cod

Survey Zs compared well with ICES survey-independent estimates of fishing mortality. Cusums appeared to provide more signals than trend analyses but this may be an artifact of the high variability of recruitment, growth and productivity shown by this stock over the time-series. The effectiveness of spatial indicators was reduced by variability in the area of survey coverage from year to year. The "traffic light" summary of Cusum results seemed promising. MFA indicated a strong, age-dependent spatial structure for the stock.

4.2 North Sea cod

The derivatives method indicated reducing total mortality in recent years while LaM50 showed a consistent decline over the period. The Cusum method provided an alert for abnormally decreasing abundance from 1997. Trends over time were found for spatial location indices and some other spatial indices using various nonparametric methods but an obvious interpretation was not available. Comparable results were obtained with multivariate methods. Constructed multivariate indicators were found useful for identifying the most influential factors in groups of correlated indicators.

4.3 North Sea herring

Awaiting report.

4.4 Western (Biscay) hake

The power method showed that the power to detect trends was low due to high inter-annual variability and missing values for biological and spatial indicators. The derivatives method detected significant trends for all the length indices and Z during the recent 5 years. The Cusum method found various spatial changes for different age groups, and an increase in Z. Composite multivariate indices did not provide results with obvious interpretations. Small discrepancies between the survey and ICES assessments were observed in the timing and range of variation of recruitment.

4.5 Biscay anchovy

The daily egg production survey confirmed the importance of recruitment indicators. The Cusum and derivatives method successfully revealed changes in abundance related indicators. Overall abundance was related to the area occupied by the stock and inversely related to anisotropy. Composite multivariate indices were found to be difficult to interpret. Analysis of the acoustic survey also found that spatial distribution and abundance indicators were well correlated, that abundance is heavily dependent on successful recruitment. These relationships corroborate alerts about the stock triggered by ICES.

4.6 Central Mediterranean red mullet

Trends were observed in Z, length-based indicators and spatial indicators but these were not highly statistically significant, possibly because of the shortness of the time-series. The Cusum method triggered an alert when the survey abundance index reached its lowest level in 1997. Composite multivariate indicators were found useful for interpreting results from correlated indicator series. In general, the findings were difficult to interpret, although combination of the approaches based on the Cusum method and trend analysis indicates that the population dynamics of red mullet are affected by impacts that influence demography and production probably with cyclical phases, although the most recent condition displays signs of an increased exploitation pattern.

4.7 Baltic cod

The Cusum method successfully signalled a significant reduction of age 5 Baltic cod in 2000, and a reduction in the spatial indicator, positive area in 2002. Biological indices were less successful at signalling changes, except for L75 which signalled a decrease in numbers of large fish in the last year of observation. Trend analysis in the cause-effects table suggested that fishing mortality was having the most effect on the stock. The age structure of the stock was predominantly young, implying a strong dependence on recruitment. LaM50 was found to be unreliable because the seasonal timing of the survey was not optimal for estimating maturity. However, MFA and spatial indices revealed different spatial distributions for young and old individuals, probably related to their stage of maturity. These analyses supported existing ICES assessments that Baltic cod are too heavily fished.

4.8 Eastern Mediterranean hake

Aegean Sea

No significant trends were observed for any of the biological parameters considered using the linear and derivatives methods. However, there was an indication from the CUSUM analysisof a general increase of abundance towards 2003 compared to before 1998 when it was poor. Although there was no signal for increasing abundance of recruits, there were negative signals in the abundance of immature fish from 1994 to 1997 and in the abundance of mature fish in 1995 and 1996, which consequently did not give any signal (CUSUM). Furthermore, the different lengths did not give any signals after 1997 (CUSUM). These results show a stable situation at least after 1997 for the biological indices, however, trends may be observed if more recent years are considered. Concerning the spatial indices increasing trends in the positive area of the younger age groups (A0, A1, A2) were observed for the whole period, whereas some recent trends were found in the distribution characteristics of different ages (mostly A4).

Assessments for hake in the Aegean Sea are not available. Given the facts that hake were considered overexploited at the beginning of the studied period, that after 1994 the fishing effort was reduced gradually until 2005, and that a larger cod-end mesh size was imposed recently, it is suggested that enforcement of the existing measures be continued because they may have assisted the observed increased abundance. The protection of recruits by expansion of closed seasons in the main nursery grounds is also recommended.

Ionian Sea

There was an increasing trend in abundance during the studied period (linear). This result in combination with the decreasing trend of L25 and Lbar obtained for the recent years (derivatives method) could be showing an increasing trend in recruitment. However, this scenario was not confirmed by the trend of the recruitment index (not significant), although there was an indication of increased recruitment in the last year (CUSUM). The simultaneous decreasing trend of L75 (derivatives) could show a slower growth rate. However, this scenario was not supported by the increasing trend of abundance and the decreasing trend of L25. Furthermore, the CUSUM analysis did not give any alarm for the different lengths. Additionally, the CUSUM analysis did not give any signal for the abundance of the different ages with the exception of recruits and the mdbio showed to be driven only by the positive alarms of the recruitment index (1995, 2003). The results are not clear and perhaps this is due to the short data series (1994-2003). The consideration of more recent years in the analysis is suggested in order to obtain a clearer image.

Assessments for hake in the Ionian Sea are not available. Given the facts that hake was considered overexploited in the beginning of the studied period, that after 1994 fishing effort was reduced gradually until 2005, and that a larger cod-end mesh size was imposed recently, it is suggested to continue the enforcement of the existing measures, which may have resulted in an amelioration of recruitment. The consideration of more recent years in the analysis is suggested before imposing new measures.

4.9 Comments

The results of using the FISBOAT indicators and methods clearly varied with the stock and the survey. The pattern of variability and its time scale along the indicator series strongly influenced the detection of change in the time series. Short series and missing values created additional problems. There were also difficulties in choosing reference periods as baselines for detecting change when stocks were heavily impacted by fishing throughout the survey series. For the same reason, some stocks showed no major changes in biological or spatial variables and, in those circumstances, it is not unreasonable that FISBOAT methods failed to detect changes.

It is important that indicators and methods be chosen to reveal the prevailing state of a stock, as well as changes to it. Abundance, weight, spawning stock biomass, and age structural indices often serve this purpose satisfactorily, as they already do in the current assessment system. Total mortality, on the other hand, tended to be constant, or else was too noisy to give a clear signal about the stock. Although length-based indices had much to offer conceptually, their signal and coherence with other indices was difficult to interpret in several cases.

Low stock levels imply that fine tuning of managerial measures to control fishing is not required. Instead, it is necessary, at least initially, to aim to achieve improving trajectories for all important indicators. Negotiations on the desirable level of long-term fishing effort can begin when the security of the stock is clearer.

Table 1. FISBOAT project: Biological indicators

Indicator	Symbol	Description and properties	Section in Manual
Intrinsic population growth rate	r	Slope of log total abundance (all ages) over time. Decreases with fishing particularly if recruitment is also affected. Other factors could also cause a decline. Estimated from time-series of abundance indices.	3.2
Total mortality	Z	Coefficient of total mortality averaged over a given age range. Increases with fishing, or net migration out of the survey area. Estimated from log year class numbers from year to year.	3.3
Numbers-at-length, numbers-at-age	NaL, NaA	Length- or age-frequency distributions. Lack of large or old fish may indicate over-fishing, low productivity economically, and vulnerability to high fishing pressures.	3.4
Spawning stock in number	SSN	Number of mature fish per tow. Low SSN implies a stock vulnerable to interference with reproductive processes, and high fishing pressures. Requires accurate maturity staging and surveys in the season leading up to spawning. SSN is reduced by fishing.	3.5
Length statistics	Lbar, L25, L50, L75	Mean or percentiles of fish lengths found in survey catches. They indicate growth, recruitment, and numbers of older, spawning fish. The different percentiles respond differently to fishing, recruitment pulses, and loss of spawning stock.	3.6
Catch weight	W	Total weight, or weight per unit effort of one (or more) species. Relates to stock biomass and size composition and is affected by seasonal growth and reproduction. Less influenced by varying recruitment annually than length indices.	3.7
Condition	C	Average body weight for a given body size. Reflects nutritional status and reproductive fitness. Estimated by regression from length frequencies but surveys must take place in the appropriate season. Condition varies with gender.	3.8
Gonadosomatic index	GSI	The ratio of gonad weight to body weight. Affected by nutritional status, maturity stage, and reproductive fitness. An advantage over SSN is that maturity stages are not used.	3.9
Length and age at maturity	LaM50, AaM50	Length or age at which 50% of the individuals in a fish stock have reached reproductive maturity. Can decrease slowly with fishing but can also vary widely with latitude. Consistent maturity staging is required.	3.10

Table 2. FISBOAT project: Spatial indicators

Indicator	Symbol	Description and properties	Section in Manual
Centre of gravity	CG	Mean location of the individuals of a population. A shift may reflect effects of fishing. CG is sensitive to high densities of fish.	4.2
Inertia	I	Variance of the location of the individuals of a population. Indicates dispersal but is sensitive to high densities of fish.	4.3
Anisotropy, Isotropy	An, Is	Anisotropy measures the elongation of the spatial distribution of the population. Isotropy is the inverse. Can be affected by the appearance or disappearance of patches of fish.	4.4
Global index of collocation	GIC	Measures the geographic distinctness or overlap of two populations of fish.	4.5
Number of spatial patches	NoP	Measures the geographic patchiness of fish populations. NoP depends on the threshold distance separating two patches and is sensitive to the locations of high densities of fish.	4.6
Positive area	PA	Measures the area where fish of a species occur. PA is greatly increased when fish occur at low densities over a large area.	4.7
Spreading area	SA	Measure of the area occupied by the stock, based on how the abundance is spreading in space. SA equals PA when the population is evenly spread with a constant density.	4.8
Equivalent area	EA	Represents the area that would be covered by the population if all individuals occupied the same area. Independent of the absolute abundance and sensitive to the highest density values.	4.9
Microstructure index	MI	Measures the relative importance of structural components having a scale smaller than the sample lag. 0 corresponds to a very regular, well-structured density surface, and 1 corresponds to a highly irregular, poorly structured, density surface, or to measurement noise.	4.10

Table 3. FISBOAT project: Methods for integrating and interpreting indicator series

Task/Category	Method	Description and properties	Section in Manual
Simulating population dynamics	ALADYM	An age-length based simulation model for predicting the effects of different fishing pressures on a single population of fish.	5.2.1
	Estimating indicators and reference points	A Monte Carlo approach using ALADYM	5.2.2
Indicator time-series methods	Derivatives method for determining recent trends	A method to estimate the direction of recent changes – up or down – using the first and second derivatives of the smoothed time-series.	5.3.1
	Assessing the power to detect future trends	A model-based method for estimating the power to detect future, linear trends.	5.3.2
	Statistical process control (SPC) schemes	Application of industrial quality control schemes, e.g. Cusum, to monitor fishery and environmental qualities derived from time series of indicators.	5.3.3
	Nonparametric statistical methods for assessing trends	Statistical method for assessing trends in fishery and environmental indicators without models and with a minimum of assumptions.	5.3.4
Construction of multivariate indicators	Principal components analysis (PCA) and biological indicators	Method to represent the evolution of a stock, characterised by many biological indicators, as a multivariate distance from a reference centre of gravity.	5.4.1
	Multi-factorial analysis (MFA) and spatial indicators	MFA extends the PCA method to cases where the same variables (spatial indicators) are measured for the same individuals (e.g. stations) at different times. Can be applied to summerise the spatial organisation of a species life cycle through different age classes.	5.4.2
	Min/Max autocorrelation factors (MAFs) and time continuity	MAFs are linear combinations of indicator series whose autocorrelation decreases from the first to last (much as PC factors explain decreasing proportions of the variance). The first MAF extracts the variation which is most continuous in time and can be used with spatial indices to follow distribution of a fish population over time.	5.4.3

Table 3 continuation. FISBOAT project: Methods for integrating and interpreting indicator series

Diagnosing stock status from indicator series	Combining trend signals using a cause-effects table	A simple, visual system for interpreting different types of biological and fishery information provided by indicator time series, and by other sources if available.	5.5.1
	A 'traffic light' procedure based on Cusum out-of-control tables	Illustration of the construction of a simple, visual traffic light table from the results of Cusum procedures applied to various indicator series.	5.5.2
	A multi-variate statistical procedure	Demonstration of how PCA can be informatively applied to the results of Cusum procedures applied to the various indicator series.	5.5.3

Table 4. FISBOAT project: Fishery-independent assessment models

Method	Abbrev-iation	Description and properties	Section in Assessment methods manual
Biomass random effects model	BREM	Relative biomass is modelled as a function of last year's biomass, a net growth coefficient incorporating growth and mortality, and annual recruitments. The growth coefficient and recruitments are treated as random walks on the log scale.	1.1
Length structured model	LENSUR	Generates an artificial population in numbers by length class and time step, as specified by a set of parameters. Model observations are derived from the operating model in an observation model, and parameters are estimated by minimising the deviation of the model observations from real observations.	1.2
Survey-based, age structured model	SURBA	Abundance at each age and year of a cohort is given by the recruiting abundance of the relevant cohort modified by the cumulative effect of (separable) mortality during its lifetime. Parameters are estimated by minimising the weighted sum-of-squares of observed and estimated abundance indices. All abundance estimates are relative.	1.3
Time-series analysis	TSA	A state-space, random walk framework for modelling abundance-at-age indices from a single survey. Fitting of parameters by Kalman filter.	1.4
Year-class curve regression	YCC	Estimates annual recruitments and total mortality, Z, by regressing log abundance indices on age by year class. Also estimates relative catchabilities and residual variances for different surveys.	1.5

Table 5. FISBOAT project: Details of case studies.

Stock	Behaviour	Life span	Survey Type	Survey timing	Reference period	Age range, years	Data on Length	Data on Maturity	Manageme nt problem
Barents Sea Cod	Demersal	Long	Botttom trawl & acoustics	1989-2004, qtr. 1	1996-2004	1-10+	+	+	
North Sea Cod	Demersal	Long	Bottom trawl	1985-2005, qtr 1	1999-2003	1-6+	+	+	+
North Sea Herring	Pelagic	Long	Acoustics	Long		+	+	+	
Western Hake	Demersal	Long	Bottom trawl	1987-2004 not 91,93,96	1987-1997	0-5+	+	+	+
Biscay Anchovy	Pelagic	Short	Acoustics & eggs (DEPM)	1989-2005	1990-2001	?	+	+	
Central Mediterranean Red mullet	Demersal	Short	Bottom trawl	1994-2003	1999-2003	1-5	+	+	
Baltic Sea Cod	Demersal	Long	Bottom trawl	1994-2004, qtr. 1	1994-1999 excl. 1997	1-5	+	+	+
Eastern Mediterranean Hake	Demersal	Long	Bottom trawl	Medium		-	+	+	

References

Abella, A., Belluscio, A., Bertrand, J., Carbonara, P., Giordano, D., Sbrana, M., and Zamboni A. (1999) - Use of MEDITS trawl survey data and commercial fleet information for the assessment of some Mediterranean demersal resources. *Aquat. Living Resources*., **12** (3): 155-166.

Anonymous (1986) Progress report 1985/1986 by ICES Statistician. *International Council for the Exploration of the Sea*, Copenhagen. **ICES CM 1986/D:19**, 11 pp.

Anonymous (2002) - *Stock Assessment in the Mediterranean*. SAMED project. Final Report EU Project n° 99/047.

Anonymous (2004a) Report of the workshop on survey design and analysis (WKSAD). *International Council for the Exploration of the Sea*. **ICES CM 2004/Fishing Technology Committee**, 73pp. www.ices.dk/reports/FTC/2004/WKSAD04.pdf

Anonymous (2004b) Report of the workshop on survey design and data analysis (WKSAD). *International Council for the Exploration of the Sea*, Copehagen. **ICES CM 2004/B:07, Ref. D.G.**, 261pp.

Anonymous (2005a) Report of the workshop on survey design and analysis (WKSAD). *International Council for the Exploration of the Sea*. **ICES CM 2005/B:07 Ref. D, G**, 170 pp. www.ices.dk/reports/FTC/2005/WKSAD05.pdf

Anonymous (2005b) Report of the workshop on survey design and data analysis (WKSAD). *International Council for the Exploration of the Sea*, Sete, France. **ICES CM 2005/B:07**. file:///C:/ICES%20Survey%20Design%20%26%20Analysis/WKSAD05.pdf

Anonymous (2006) Report of the study group on survey trawl standardisation (SGSTS). *International Council for the Exploration of the Sea*. **ICES CM 2006/FTC:05**, 67 pp. www.ices.dk/reports/FTC/2005/SGSTS05.pdf

Armstrong, M.J., Payne, A.I.L. and Cotter, A.J.R. (2007) Contributions of the fishing industry to research through partnerships. In: *Advances in fisheries science 50 years on from Beverton and Holt.* ed Payne, A.I.L., Cotter, A.J.R. and Potter, T., Blackwell Publishing.

Beaugrand, G. (2005) Monitoring pelagic ecosystems using plankton indicators. *ICES Journal of Marine Science* **62**, 333-338.

Beverton, R.J.H. (1998) Fish, fact and fantasy. *Reviews in Fish Biology and Fisheries* **8**, 229-249.

Caddy, J.F. (2004). Current usage of fisheries indicators and reference points, and their potential application to management of fisheries for marine invertebrates. *Canadian Journal of Fisheries and Aquatic Sciences*, **61**, 1307-1324

Callaway, R., Alsvåg, J., de Boois, I., Cotter, J., Ford, A., Hinz, H., *et al.* (2002a) Diversity and community structure of epibenthic invertebrates and fish in the North Sea. *ICES Journal of Marine Science* **59**, 1199-1214.

Callaway, R., Jennings, S., Lancaster, J. and Cotter, J. (2002b) Mesh-size matters in epibenthic surveys. *Journal of the Marine Biological Association of the United Kingdom* **82**, 1-8.

Commission, E. (2001a) Elements of a strategy for the integration of environmental protection requirements into the Common Fisheries Policy. Brussels. **COM(2001) 143 final**, 23pp.

Commission, E. (2001b) The future of the common fisheries policy. *European Commission*, Luxembourg. 47pp.

Commission, E. (2001c) Report on the economic and social situation of coastal regions. *European Commission*, Luxembourg. 31-72

Cotter, A.J.R., Burt, L., Paxton, C.G.M., Fernandez, C., Buckland, S.T. and Pan, J.-X. (2004) Are stock assessment methods too complicated? *Fish and Fisheries* **5**, 235-254.

Cotter, A.J.R., Petitgas, P., Mesnil, B., Trenkel, V.M., Rochet, M.-J., Woillez, M., *et al.* (2007) Manual of indicators and methods for assessing fish stocks using only fishery-independent, survey based information. *International Council for the Exploration of the Sea*. **ICES CM 2007/O:27**, 135pp.

Cotter, A.J.R. and Pilling, G.M. (2007) Landings, logbooks and observer surveys: improving the protocols for sampling commercial fisheries. *Fish and Fisheries* **8**, 123-152.

Daan, N., Christensen, V. and Cury, P.M. (2005). *Quantitative ecosystem indicators for fisheries management.*, ICES Journal of marine Science.

Daskalov, G.M. (2007) Overfishing affects more than fish populations: a case of trophic cascades and regime shifts in the Black Sea. In: *Advances in fisheries science 50 years on from Beverton and Holt.* ed Payne, A.I.L., Cotter, A.J.R. and Potter, T., Blackwell Publishing.

Daskalov, G.M., Grishin, A., Rodionov, S. and Mihneva, V. (2007) Trophic cascades triggered by overfishing reveal possible mechanisms of ecosystem regime shifts. *Proceedings of the National Academy of Sciences* **104**, 10518-10523.

De Oliveira, J.A.A., Kell, L.T., Punt, A.E., Roel, B.A. and Butterworth, D.S. (in press) Managing without best predictions: the Management Strategy Evaluation framework. In: *Advances in fisheries science 50 years on from Beverton and Holt.* ed Payne, A.I.L., Cotter, J. and Potter, T., Blackwell Publishing.

Demaré, W. (2006) Fisheries-based management and advice in Europe. In: *The knowledge base for fisheries management.* ed Motos, L. and Wilson, D.C., Elsevier**:** 295-305.

Ehrich, S., Adlerstein, S., Brockmann, U., Floeter, J., Garthe, S., Hinz, H.*, et al.* (2007) 20 years of the German small-scale bottom trawl survey (GSBTS): A review. *Senckenbergiana Maritima* **37**, 13.

Everitt, B.S. (2005) *An R and S-plus companion to multivariate analysis.*, Springer, London

FAO (2003) The ecosystem approach to fisheries. *FAO, Fisheries Department*, Rome. **No. 4, Suppl. 2**, 112pp.

Fulton, E.A., Smith, A.D.M. and Punt, A.E. (2005) Which ecological indicators can robustly detect effects of fishing? *ICES Journal of Marine Science* **62**, 540-551.

Garcia, M. and Cochrane, K.L. (2005) Ecosystem approach to fisheries: a review of implementation guidelines. *ICES Journal of Marine Science* **62**, 311-318.

Gjøsaeter, H., Bogstad, B. and Tjelmeland, S. (2002) Assessment methodology for Barents Sea capelin, Mallotus villosus (Müller). *ICES Journal of Marine Science* **59**, 1086-1095.

Greenstreet, S.P.R., Tuck, I.D., Grewar, G.N., Armstrong, E., Reid, D.G. and Wright, P.J. (1997) An assessment of the acoustic survey technique, RoxAnn, as a means of mapping seabed habitat. *ICES Journal of Marine Science* **54**, 939-959.

Hall, S.J., Collie, J.S., Duplisea, D.E., Jennings, S., Bravington, M. and Link, J. (2006) A length-based multispecies model for evaluating community responses to fishing. *CANADIAN JOURNAL OF FISHERIES AND AQUATIC SCIENCES* **63**, 1344.

Hewitt, D.A. and Hoenig, J.M. (2005) Comparison of two approaches for estimating natural mortality based on longevity. *Fisheries Bulletin* **103**, 433-437.

Hiddink, J.G., Jennings, S. and Kaiser, M.J. (2006) Indicators of the ecological impact of bottom-trawl disturbance on seabed communities. *Ecosystems* **9**, 1190.

Jennings, S. (2005) Indicators to support an ecosystem approach to fisheries. *Fish and Fisheries* **6**, 212-232.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.M., Garcia, D., Hillary, R.*, et al.* (2007) FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science* **64**, 640-646.

Kell, L.T., Pilling, G.M., Kirkwood, G.P., Pastoors, M., Mesnil, B., Korsbrekke, K.*, et al.* (2005) An evaluation of the implicit management procedure used for some ICES roundfish stocks. *ICES Journal of Marine Science* **62**, 750-759.

Mackinson, S. and van der Kooij, J. (2006) Perceptions of fish distribution, abundance and behaviour: Observations revealed by alternative survey strategies made by scientific and fishing vessels. *Fisheries Research* **81**, 306.

Mackinson, S., Freeman, S., Flatt, R. and Meadows, B. (2004) Improved acoustic surveys that save time and money: Integrating fisheries and ground-discrimination acoustic technologies. *Journal of Experimental marine biology and ecology* **305**, 129.

Mesnil, B., Cotter, J., Fryer, R.J., Needle, C., Skagen, D., Spedicato, M.T., *et al.* (2007) Review of fishery-independent assessment methods. *International Council for the Exploration of the Sea*. **ICES CM2007/O:04**,

Methratta, E.T. and Link, J.S. (2006) Evaluation of quantitative indicators for marine fish communities. *Ecological Indicators* **6**, 575.

Mitson, R.B. (1995) Underwater noise of research vessels - review and recommendations. *ICES Cooperative Research Report*. **209**, 61 pp.

Motos, L. and Wilson, D.C. (2006) *The knowledge base for fisheries management.*, Elsevier, Amsterdam

Piet, G. and Jennings, S. (2004) Response of potential fish community indicators to fishing. *ICES Journal of Marine Science* **62**, 214-225.

Piet, G., Quirijns, F.J., Robinson, L. and Greenstreet, S.P.R. (2007) Potential pressure indicators for fishing, and their data requirements. *ICES Journal of Marine Science* **64**, 110-121.

Pope, J.G. (1979) Population dynamics and management: current status and future trends. *Investigacion Pesquera* **43**, 199-221.

Rees, H.L., Boyd, S.E., Schratzberger, M. and Murray, L.A. (2006) Role of benthic indicators in regulating human activities at sea. *Environmental Science and Policy* **9**, 496-508.

Rees, H.L., Ellis, J.R., Hiscock, K., Boyd, S.E. and Schratzberger, M. (in press) Benthic communities, ecosystems and fisheries. In: *Advances in fisheries science 50 years on from Beverton and Holt.* ed Payne, A.I.L., Cotter, A.J.R. and Potter, T., Blackwell Publishing.

Rencher, A.C. (1995) *Methods of multivariate analysis.*, J. Wiley & Sons, Inc., New York

Rice, J.C. and Rochet, M.-J. (2005) A framework for selecting a suite of indicators for fisheries management. *ICES Journal of Marine Science* **62**, 516-527.

Rice, J.C., Trujillo, V., Jennings, S., Hylland, K., Hagstrom, O., Astudillo, A., *et al.* (2005) Guidance on the application of the ecosystem approach to management of human activities in the European marine environment. *ICES Cooperative Research Report* **273**, 22pp.

Rivard, D. (1989) Overview of the systematic, structural, and sampling errors in cohort analysis. *American Fisheries Sociey Symposium* **6**, 49-65.

Rochet, M.-J. and Rice, J.C. (2005) Do explicit criteria help in selecting indicators for ecosystem-based fisheries management? *ICES Journal of Marine Science* **62**, 528-539.

Rose, G.A. and Kulka, D.W. (1999) Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. *Canadian Journal of Fisheries & Aquatic Sciences* **56**, 118-127.

Shin, Y.-J., Rochet, M.-J., Jennings, S., Field, J.G. and Gislason, H. (2005) Using size-based indicators to evaluate the ecosystem effects of fishing. *ICES Journal of Marine Science* **62**, 384-396.

Travers, M., Shin, Y.J., Shannon, L. and Cury, P. (2006) Simulating and testing the sensitivity of ecosystem-based indicators to fishing in the southern Benguela ecosystem. *CANADIAN JOURNAL OF FISHERIES AND AQUATIC SCIENCES* **63**, 943.

Trenkel, V.M. (2007) Potential limitations of survey data as the unique source of information for stock assessment and management. In: *Manual of indicators and methods for assessing fish stocks using only fishery-independent, survey-based information. ICES CM 2007/O:27*. ed, FISBOAT final report**:** section 2.2, pp.14-17.

Vetter, E.F. (1988) Estimation of natural mortality in fish stocks: a review. *Fishery Bulletin* **86**, 25-43.

Walters, C. (1986) *Adaptive management of renewable resources.*, Macmillan Publishing Co., New York

Ward, T., Tarte, D., Hegerl, E. and Short, K. (2002) Ecosystem-based management of marine fisheries - policy proposals and operational guidance for ecosystem-based management of marine capture fisheries. *World Wide Fund for Nature*. 80pp.

Woillez, M., Rivoirard, J., Petitgas, P. and Deerenberg, C. (2007) Selecting and combining survey based indices of fish stocks using their correlation in time to make diagnostic of their status. *International Council for the Exploration of the Sea*. **ICES CM2007/O:07.**