# ICES SGRF REPORT 2011

ICES ADVISORY COMMITTEE

ICES CM 2011/ACOM:31

# Report of the Study Group on Recruitment Forecasting (SGRF)

17–20 October 2011

Copenhagen, Denmark

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

**International Council for the Exploration of the Sea**
**Conseil International pour l'Exploration de la Mer**

H. C. Andersens Boulevard 44–46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2011. Report of the Study Group on Recruitment Forecasting (SGRF), 17–20 October 2011, Copenhagen, Denmark. ICES CM 2011/ACOM:31. 40 pp.
https://doi.org/10.17895/ices.pub.19280792

# Contents

# 1 Executive Summary


The Study Group on Recruitment Forecasting (SGRF) met at ICES HQ (Copenhagen) from October 17–20, 2011, with five participants (covering three nationalities) and Dr Sam Subbey (Norway) as Chair.

The formal mandate for this SG meeting was established in 2010/2/ACOM32 under Action Plan No: 1.2, 1.10, and 2.5. The objectives of the SG are to decide on guidelines and standards with regards to (1) How to develop models for recruitment projections which incorporate both abundance indices and environmental drivers, and (2) Criteria for validating models and for choosing the "best" or a set of the best models.

The meeting adopted an approach involving review of the fisheries literature on model-based recruitment forecasting in order to obtain an overview of key modelling frameworks. Aspects of the modelling framework considered as best practice were highlighted based on ecological and statistical considerations.

This report deals with best practice for recruitment modelling and short-term forecasting. It is organized into a review section, a section dealing with guidelines to best practice in modelling and forecasting, and a section which identifies key case studies for to which the proposed guidelines are to be applied and tested in a subsequent meeting.

The report reviews:

- Biological and statistical terminologies in the fish recruitment modelling and forecasting literature.
- Current framework on recruitment modelling and forecast.

In defining best practice, the report

- Highlights the important (early life) biological and ecological processes which underpin the stock to recruitment process.
- Suggests a statistical framework (with an ecological underpinning) for the development of recruitment models. This framework includes
    - Guidelines for model development based on several considerations other than just correlations;
    - Diagnostic tools for determining appropriateness of models and choosing between competing class of recruitment models;
    - Evaluation of the significance of covariates (environmental and stock indices) on recruitment forecasting.

The SG recognizes that understanding the ecological underpinning to recruitment is a prerequisite in the development of appropriate models, which link fish recruitment to biotic and abiotic process drivers. A key component of a framework for development of recruitment is a set of diagnostic tools to evaluate model prediction performance, and especially to identify cases where model failure is due to information collapse in one or several data drivers (biotic and abiotic). This report exemplifies the use of a selected number of diagnostic tools through evaluation of two recruitment models currently used in stock assessment.

The contents of this report provides the basic ingredients/building blocks for defining best practice in recruitment modelling/forecasting and model performance evaluation. Given the limitations (in terms of attendance and time), this report only ad-

dresses recruitment forecasting in the short term (max. three years). It is to be viewed as a first year interim report of a three-year process in the development of best practice guidelines for recruitment modelling and forecasting.

Future developments will involve

- An elaborate list of steps and guidelines (recipe) for the development and evaluation of recruitment forecasts in the short and medium term;
- Application of the recipe to identified case studies.

## 2 Terms of Reference 2011

2010/2/ACOM32 The **Study Group on Recruitment Forecasting** (SGRF), chaired by Samuel Subbey, Norway, will be established and be held at ICES Headquarters, 17–20 October 2011 to:

a ) Develop a framework and methodology for modelling of recruitment for use in short- and medium-term projections in stock assessment, incorporating abundance indices and environmental drivers;

b ) Test this framework using designated case studies;

c ) Provide guidelines and standard criteria for recruitment model validation and for choosing the "best" or a set of "best" recruitment models.

SGRF will report by 31 October 2011 for the attention of ACOM.

## Supporting Information

| | |
|---|---|
| Priority: | Because the relationship between early survey indices and recruitment is fundamental to the scientific approach to fisheries management, the work of this group should be considered of high priority to ICES. |
| Scientific justification and relation to action plan: | Action Plan No: 1.2, 1.10, and 2.5.<br><br>Recruitment models are used to provide input into short- and medium-term stock projections made as part of an assessment. For a number of stocks, the ICES standard tool RCT3 is used. This method uses survey indices as input. For other stocks, e.g. NEA cod, recruitment models are based on both environmental drivers and survey indices. Correlations have a tendency to vanish and especially models including environmental drivers (biotic and abiotic) are important to validate frequently. However, it is debatable which methods should be used in model validation and which criteria should be used for selecting recruitment models for input in forecast models. This is a general problem and we propose a study group within the ICES framework that consists of international expertise on statistical and ecological modelling. Questions that needs to be addressed are:<br>1) How to develop models for recruitment projections which incorporate both abundance indices and environmental drivers<br>2) Criteria for validating models and for choosing the "best" or a set of the best models.<br>The SG should decide on some guidelines and standard methods regarding these questions. |
| Resource requirements: | No specific resource requirements beyond the need for members to prepare for and participate in the meeting. |
| Participants: | Participants would include scientists working with fisheries management and scientists with international expertise on statistical and ecological modelling. |
| Secretariat facilities: | No additional software/hardware is anticipated beyond that which is currently available. |
| Financial: | None specific. |
| Linkages to advisory committees: | Reliable stock forecasts are highly dependent on good models for recruitment forecasting |
| Linkages to other committees or groups: | The activities of the SG are designed to provide input of knowledge to various Assessment WGs. There is no potential overlap in activities because the latter do not have the resources to consider the nature of this new knowledge outside the scope of their current activities. |
| Linkages to other organizations: | ICES will seek widened participation for this group including contact with relevant academic and intergovernmental organizations (including FAO, OECD, and IIFET) for this meeting. |
| Secretariat marginal cost share: | ICES 100%. |

# 3 Review of Terminologies

## 3.1 Definition of recruitment

The term recruitment is used in a number of ways depending on the context. In biological and ecological contexts recruitment is used in conjunction with sexual maturation and the addition of individuals to the breeding population. Sometimes the term is loosely defined as individuals joining the adult population and thus not specifically linked to the process of sexual maturation. In this case the term often refers to the location of the non-juvenile population, or the part of the population on the feeding grounds which do not form part of the nursery areas. In the context of fisheries science the term is more often used for individuals that become part of the fishable population. In this context, the numbers of individuals within a year class that recruit each year should rise in a similar fashion to a selection ogive as the fish become vulnerable to the fishery. However, to simplify the situation the recruitment is usually defined by age and treated as knife edged i.e. recruitment occurs at one age only. The choice of age-at-recruitment differs between stocks and species but often it is the first year class which is fully vulnerable (100% selection) to fishing.

The assessment literature is not very rigorous with respect to its use of the term recruitment. Within the stock assessment reports ages between 0 and 3 are often used. In regard to ages 1 to 3 the numbers usually refer to the 1st January in a year. In the case of Age 0 this refers to either approximately 1st of June, connected to a specific survey time period or some undetermined fraction of a year. One further complication that should be mentioned is that in much (but not all) of the herring literature the ages are reported as winter rings. Therefore, 1-winter ring fish in the case of autumn spawned fish are 1 year and 3 or less months old on the 1st January and approximately 8–9 months old in the case of spring spawned fish.

Examples of different usages of recruitment age can be seen in northeast Arctic cod (age 3; ICES 2011a), North Sea cod (age 1; ICES 2011b), Norwegian Spring-spawning (NSS) herring (age 0; ICES 2011c) and North Sea herring (age 0/0-winter ring; ICES 2011d).

The numbers of recruits at age are derived in a number of ways and if VPA generated abundances are used in any stock and recruitment relationship or any other modelling exercise using recruitment then the user should be aware of their source. In the cases of NEA and North Sea cod the age 3 and age 1 (respectively) data are generated based on survey and catch data and are internally consistent with the year class as its abundance changes over its lifetime. In the case of NSS herring the 0-group abundance is back-calculated from older age classes that have a catch history. For North Sea autumn spawning herring the 0-group abundance is tuned to the MIK survey, however, the abundance is internally consistent with the change in abundance of the year class over time as seen in the catch statistics of older ages.

The overall purpose of predictions or forecasts of recruitment are to provide numbers of individuals in the future. In all species the further once projects in to the future the greater the influence of one’s perception of the recruitment process and levels of recruitment. In fact in species that live for e.g. 10 years, projections or forecasts of a population beyond 10 years means the stock and its structure is solely the consequence of the model that was used to generate the annual recruitment. The further recruitment is projected in to the future the greater the importance of predicting future environmental and physical drivers for survival of young to a recruiting age.

## 3.2 Simulations, Forecasts and Projections

In science, simulation is the imitation of a real process based on key characteristics representing this process. The key characteristics of this process are usually extracted and modelled using some learning sample of data of interest. Simulations are often used to generate scenarios. . Simulation and scenario planning are thus usually related to studying the response of input variables under the influence on output variables when these are specifically varied.

The terminologies 'forecast' and 'projections' have been loosely used in the recruitment literature. The SG therefore realized the need to clarify and distinguish between these terminologies.

In statistics (and other sciences), extrapolation is the process of constructing new datapoints outside the supporting data interval while interpolation constructs new datapoints between given (known) datapoints. Statistically two general types of extrapolation techniques do exist: forecasting and projection methods.

### 3.2.1 Forecasts

Forecasts are also called predictions or prognoses. Literally, forecasting, predicting or prognosticating address the same issue: fore-knowing or foreseeing unknown future events. Statistically, they address the process of making statements about the most likely outcome of future values of a process or time-series variable whose actual outcomes are unknown; in other words, attempts to estimate true future values with some reliability to be specified. This requires that the pattern of observed time-series data is identified and more or less formally described.

Thus, in sciences forecasting, predicting, and prognosticating refer to formal statistical methods employing time-series data, cross sectional data or longitudinal data.

### 3.2.2 Projection

Projection is the process of generating new datapoints outside the given supporting interval of observed or sampled data. A projection is closely related to simulation and scenario generation, but differs from forecasting in that the aim is not to make statements about the most likely outcome of future events as no forecast error is estimated that takes into account the distance of the forecasted value from the centre of the supporting data.

## 3.3 Definition of short, medium and long-term forecasts

It is often desirable in fisheries science to forecast future stock abundance. Such forecasts can be divided into "short", "medium" and "long-term". The classification however, is to considered in relationship to the life cycle of fish being modelled. In general, short-term assessment is by far the most reliable because uncertainties around future recruitment are avoided; explaining why this type of forecast is most used in stock assessment and management. In contrast, the more uncertain medium- and long-term forecasts are more often used in the context of management strategy evaluations.

With respect to stock assessment, the expression 'short-term forecast' usually refers to predictions associated with year classes that have already been spawned but yet to enter the fishery. For NEA cod and NEA haddock stocks for which recruitment to the fishery occurs at age 3, a short-term forecast is 1–3 year ahead (ICES CM 2011/ACOM:05). However, for species like Greenland halibut and redfish, short-term forecast may be as much as 6 years ahead. Barents Sea capelin is a short-lived species, which enters the fishery at age 2–3, and therefore a short-term forecast is only one year ahead. Further, because fishing at present only takes place on mature individuals just prior to spawning (mainly ages 3–5).

A 'medium-term forecast' is synonymous to prediction of the next generation, i.e. fish that will be spawned by the current generation. In the case of NEA cod, which have an average generation time of about seven years, this gives a medium-term forecast range of about 3–10 years. For Barents Sea capelin, which seldom lives for more than five years, the medium-term forecast should be in the range 2–4 or 2–5 years.

A "long-term forecast" is associated with future populations that will be spawned by generations that are themselves yet to be spawned. Usually such forecasts involve periods of ten years and more, but for short-lived species even a five year forecast may be considered to be "long term".

# 4 Modelling and model overview

Based on their formulation, two main classes of models have been identified in the literature. These included classes (1) Parametric and (2) Semiparametric and non-parametric models.

Parametric recruitment models are expressed in terms of specific analytical or semi-analytical parametric equations for deriving the recruitment relationship. Classical parametric recruitment models include the two-parameter Beverton–Holt (Beverton and Holt, 1957) and Ricker (Ricker, 1975) models, and various reformulations.

The Beverton–Holt model expresses a density-dependent relationship between the number of recruits per spawner as a (decreasing) function of the number of spawners.

The Beverton–Holt model is based on the assumptions that juvenile competition results in a mortality rate that is linearly dependent upon the number of fish alive in the cohort at any time and that, predators are always present. The Beverton–Holt model is appropriate "if there is a maximum abundance imposed by food availability or space, or if the predator can adjust its predatory activity immediately to changes in prey abundance" (Wootton, 1990).

The Ricker model defines a density-dependent relationship, where the number of recruits per spawner is expressed as a decreasing function of the number of spawners.

The Ricker stock–recruitment model assumes that the mortality rate of the eggs and juveniles is proportional to the initial cohort size. Thus a high mortality rate of eggs and juveniles can be supported by high initial number of eggs. Biological realities that might lead to this assumption being met include cannibalism of the juveniles by the adults (Ricker, 1975), density-dependent reductions in growth coupled with size-dependent predation (e.g. increase in the time it takes for the young fish to grow through a size range vulnerable to predation; Ricker (1975)), and a time-lag in the response of a predator to the abundance of the fish (Wootton, 1990).

For both models, the parameters involved are usually determined by non-linear least-squares regression, where the goodness-of-fit is assessed using summary statistics such as the $R^2$.

Variants of the Ricker and B–H models have been reported in the literature. These include the Power (Cushing, 1973), Saila–Lorda (Iles, 1994), Shepherd (Shepherd, 1982) and the Sigmoidal Beverton–Holt (Myers *et al.*, 1995). Included in this class of models are those for which the classical functional relationship between SSB and recruitment is augmented with secondary (climatic and ecological) data such as, temperature and prey interactions. For instance Gjøsæter and Bogstad (1998) included a term for juvenile herring in the B–H model for capelin recruitment.

The RCT3 model for stock–recruitment (Shepherd, 1997) is a regression model, which belongs to the class of parametric models. This model allows for the combination of multiple estimates of stock–recruitment derived from different index-series based on inverse variance weighted averages. The RCT3 model is used as the main stock recruitment model for NEA cod and NSSH.

Another subclass of parametric models encapsulates models referred to in the literature as time-series models. In a subclass of time-series models, the dependent parameter (recruitment) is regressed on one or several (and often time-lagged versions of) independent dataseries, and may include time-lagged values of the independent variable. These models are also referred to as Box–Jenkins (Box and Jenkins, 1976) models. The underlying model philosophy is to model recruitment as a combination of autoregressive (AR) and moving average (MA) effects, leading to an Auto-Regressive Moving Average (ARMA) model. In the literature, ARIMAX models have been developed for recruitment forecasting for NSSH (Gröger *et al.*, 2010) and cod (Gröger *et al.*, 2011).

A second subclass involves the use of state–space model formulations to link recruitment to population parameters. A discrete-time state–space model is defined by two equations namely, the observation (or measurement) equation and the system (or transition) equation. The system equations have the flexibility of introducing time varying parameters into the modelling framework. Discrete-time state-space models provide the same type of linear difference relationship between the inputs and the outputs as a linear ARX (Auto-Regressive models with eXogenous inputs, which are defined as variables that are determined outside the modelling process) model, but are rearranged such that there is only one delay in the expressions. As an example, a Bayesian state–space model for stock–recruitment has been reported for Fraser Piver pink salmon (Myers and Millar, 2001). The approach addressed two major problems encountered in traditional stock–recruitment analyses, that of errors-in-variables bias and time-series bias. Both process and observation errors were explicitly captured in the state–space model and quantified through posterior distributions of the parameters via the Bayesian paradigm.

Semi-parametric and nonparametric estimation methods for recruitment relationships avoid strong assumptions implied by parametric approaches, and are thus becoming increasingly popular. Classical nonparametric methods include: construction of the distribution of recruitment given stock biomass through nonparametric density estimators (Evans and Rice, 1988); using generalized additive models to estimate the relationship of recruitment with spawning biomass and an environmental variable, such as sea surface temperature (Jacobson and MacCall, 1995); fitting a locally weighted smoothing function with nonparametric regression and spline methods (Cook, 1998); and using neural networks to estimate the recruitment function (Chen and Ware, 1999). A practical drawback of these methods involves uncertainty quantification for the estimates of the recruitment function and of management reference points resulting from the estimated relationship. This is a direct consequence of the fact that classical nonparametric estimation techniques do not involve probabilistic modelling of the underlying (conditional) distribution. Hence, by avoiding potentially suspect parametric distributional forms, i.e. by avoiding likelihood specification, they are inevitably limited to point estimation. When developed, error bounds depend heavily on asymptotic results, which are unreliable because of the small sample sizes typically available for recruitment inference.

## 4.1 Examples of usage of fish recruitment forecast models that combine environmental and stock information

Predictions of recruitment based on a combination of stock and environmental variables have been implemented by few stock assessment working groups, although there appears to be a growing trend in the number of such models. Here we list some examples where such models have been implemented by stock assessment working groups. For the sake of completeness, we also list a small selection of relevant models developed under the same paradigm but currently not used by any stock assessment working group.

## 4.2 Operating recruitment models in stock assessment

### NEA cod and AFWG

Since 2008 several regression models, which includes stock and climate variables, have been combined into a "hybrid" model (i.e. the average of the individual model predictions) used by the Arctic Fisheries Working Group (AFWG) for prediction of NEA cod recruitment-at-age 3 (ICES CM 2011/ACOM: 05). The individual models used are described as below, where the tilde sign (~) is used to express dependence:

TITOV1: $R3_t \sim (DOxSat_{t-13})^2 + DOxSat_{t-13} + CodA2_{t-11} + Tw_{t-17}$,
TITOV2: $R3_t \sim (DOxSat_{t-13})^2 + ITa_{t-39} + CodA1_{t-23} + Tw_{t-17}$,
TITOV3: $R3_t \sim ITa_{t-39} + \log(CodC0_{t-28}) + Tw_{t-26}$,
JES1: $R3_t \sim Tw_{t-3} + Age1_{t-2} + \log(CapMatBio_{t-2})$.

The terms DOxSatt-13~ Exp(OxSatt-13) − OxSatt–38 and ITat-39 ~ It-39 +Tat–44. The term OxSat refers to the oxygen saturation at bottom layers of the Kola section stations 3–7, Ta the air temperature at the Murmansk station, Tw the water temperature: 3–7 stations of the Kola section (layer 0–200 m), and I is the ice coverage in the Barents Sea. The terms CodA1 and CodA2 represent the acoustic abundances indices for age 1 and 2 respectively, and Age1 is the bottom-trawl abundance (age 1 year index) of NEA cod from the joint winter Barents Sea acoustic survey. CodC0 is the age 0 index of NEA cod from the Barents Sea ecosystem survey in August–September. The numbers in parentheses are the time-lags in months (relative to 1 January) for the TITOV models and in years for the JES model. The ITa index coincides in time with the increase of horizontal gradients of water temperatures in the area of the Polar Front (Titov, 2001).

**Table 1. Overview of which models that went into the Hybrid model at the AFWG 2011 assessment (ICES AFWG:CM:05 2011). The Hybrid model was just an average of the individual regression model forecasts.**

| Model | One year forecast | Two year forecast | Three year forecast |
|---|---|---|---|
| Titov1 | X | | |
| Titov2 | | x | |
| Titov3 | X | x | x |
| JES1 | | x | |

## 4.3 Examples of non-operating models in stock assessment

### NEA cod

T. Bulgakova (ICES CM 2011/ACOM:05) developed a model that is a modification of Ricker's model for stock–recruitment defined by:

$R3_t \sim m_{t-3} \exp[-SSB_{t-3} + N_{t-3}]$

Where R3 is the number of age3 recruits for NEA cod, m is an index of population fecundity, SSB is the spawning–stock biomass and N is equal to the numbers of months with positive temperature anomalies (TA) on the Kola Section in the birth year for the year class. The subscript denotes the time-lag in years. For the years before 1998 TA was calculated relative to the monthly average for the period 1951–2000. For intervals after 1998, the TA was calculated relative to a linear trend in the temperature for the period 1998–present. The model was run using two time intervals (using cod year classes 1984–2000, and 1984–2004) for estimating the model coefficients.

Hjermann *et al*. (2007) developed a model with a one year prognosis. Based on a modification of this model, Dingsør *et al*. (WD 19) developed four models with 1–2 year projection capabilities. The four models (H1–H4) are defined by:

H1: $\log(R3_t) \sim Temp_{t-3} + \log(Age0_{t-3}) + BMcod3\text{–}6 / ABMcapelin_{t-2,t-1}$
H2: $\log(R3_t) \sim Temp_{t-2} + I(surv) + Age1_{t-2} + BMcod3\text{–}6 / ABMcapelin_{t-2,t-1}$
H3: $\log(R3_t) \sim Temp_{t-1} + Age2_{t-1} + BMcod3\text{–}6 / ABMcapelin_{t-1}$
H4: $\log(R3_t) \sim Temp_{t-1} + Age3_{t-0}$

Temp is the Kola yearly temperature (0–200 m), Age0 is the 0-group index of cod, Age1, Age2 and Age3 are the winter survey bottom-trawl index for cod age 1, 2 and 3, respectively, BMcod3–6 is the biomass of cod between age 3 and 6, and ABM is the maturing biomass of capelin. The subscript denotes to time-lag in years.

### Barents Sea capelin

The capelin spawns close to the shore on the north Norwegian and northwestern Russian coast during March–April. After hatching the larvae rise from bottom and drift northeastward. The age of maturation vary from age 2–5. The capelin is an important food source for all the large fish species, as well as for seals, whales and seabirds. In years with high abundance of herring in the Barents Sea, the survival of capelin larvae tend to be low due to predation from herring. The capelin stock biomass may fluctuate widely between. Annual catches of capelin have varied from zero to 3 million tonnes.

Stiansen *et al*. (2005) developed a model:

$Rec1_t \sim TempSkinBS_{t-1} + 0group_{t-1} + capmatbio_{t-1}$

where the subscript denotes the time-lag in years. Rec1 is the number of recruits (acoustic survey estimates back-calculated to 1 August), TempSkinBS the skin temperature from the NCEP reanalysed database average from January to March and over the Barents Sea subarea between 30–45°E and 71–75°N one year earlier, 0group the capelin 0-group trawl survey index one year earlier (in August) and capmatbio the capelin maturing biomass in tonnes (acoustic survey estimates of fish above 14 cm length) one year earlier. Data available for the model reach back to 1984 for the response variable. The model gives a one-year prognosis of the capelin recruitment. The surface temperature in the Southern part of the Barents Sea during winter was chosen as climatic parameter. The chosen area is occasionally partly covered in ice in this period, and thereby influencing the estimate of skin temperature. This climatic term is therefore a proxy for both heat conditions (temperature) and available area (ice cover). The 0-group term is the link back to the parent population. It is not obvious how the maturing biomass is coupled to the 1-group, because this is not the parent population that gives the 1-group we are looking at. Most capelin die after spawning (Gjøsæter, 1998) so there should not be a direct link between the maturing term on one side and the 0-group and 1-group term on the other side. However, this combination of parameters and time-lag gave the best fit. If one should speculate about how they are coupled, two mechanisms are likely. First, feeding conditions would be equal for both the maturing and the 0-group populations, which gives the survival from 0-group to 1-group. Second, the maturing populations may act as a buffer for predation, i.e. cod, as a major predator (Mehl, 1989), prefer larger individuals to small ones (0-group/1-group), but will eat anything if large individuals are not accessible (years with low mature population).

Titov *et al*. (ICES CM 2011/ACOM:05) developed a model for one-year forecast of Barents Sea capelin recruitment, using the capelin 0-group abundance indices of capelin 0-group in the year of spawning (Cap0t-12) and the capelin mature biomass the year before spawning (Mbiot-24) as stock variables. As climate drivers mean monthly anomalies of ice coverage of the Barents Sea (I), mean monthly anomalies of air temperature at the Murmansk station (Ta) and mean monthly anomalies of saturation by oxygen of near-bottom-water layers at 3–7 stations of the Kola Section (OxSat) were used. The subscript denotes number of months that the variables are lagged compared to the predictor.

$$Cap1_{t} \sim (OxSat_{t-12})^2 + ITa_{t-20} + Cap0_{t-12} + MBio_{t-24}$$

### NSAS herring

Motivated to understand better the recent years of reproductive failures of commercially valuable North Sea herring large-scale climate changes in the North Atlantic Ocean and their potential effects on stock regeneration were studied by Gröger *et al*. (2010). Applying time-series analyses, it was possible to reconstruct the full-time-series of recruitment solely from climate cycles, indexed by the North Atlantic Oscillation (NAO) as an index of atmospheric variability (lagged by five years) and the Atlantic Multidecadal Oscillation (AMO) as an index of sea surface temperature (lagged by three years). Based on these two climatic drivers a prognostic model was developed to provide forecasts of herring stock–recruitment three years in advance which, in contrast to competing models (Ricker, Beverton–Holt, Ricker extended by climate), did best in terms of the diagnostic measures used (AICC, performance, forecast power) and explained most of the recruitment variability.

### NSS herring

Stiansen *et al*. (2005) developed the model:

$$Rec3_{t} \sim TempSkinNS_{t-3} + 0group_{t-3}$$

where the subscript denotes the time-lag in years. Rec3 is the number of three recruits of Norwegian spring-spawning herring from ICES Northern Pelagic and Blue Whiting Working Group (WGNPBW) 2004 SEASTAR assessment (ICES CM 2004/ACFM:24), TempSkinNS the NCEP skin (sea surface) temperature in degree C in the Norwegian Sea subarea between 64–70°N and 6°W–8°E averaged from January to March three years earlier and 0group the 0-group logarithmic index of herring larvae from the 0-group survey in August three years earlier. The data used for the model are from the period 1983–2002 for the response variable. The model gives a three-year prognosis of the herring recruitment. The winter surface temperature around the spawning sites was picked as the climatic term, while the 0-group was picked as the link back to the parent population. The latter is very closely linked to the number of three-s, and also optimizes the time-lag used in the prediction. Even if it is possible to go even further back in time, there is a large change in the distribution of young herring around 1983, when the stock started to recover (stock overview is given in Chapter 6.1).

### Georges Bank cod

Climatic influences on Georges Bank cod recruitment were investigated by Gröger and Fogarty (2011) using the North Atlantic Oscillation (NAO) as an index of atmospheric variability and the Atlantic Multidecadal Oscillation (AMO) as an index of sea surface temperature. A quantitative approach based on a simple Cushing-type stock–recruitment model was developed and extended to include climate influences using the technique of generalized transfer functions (ARIMAX modelling). This allowed the autoregressive nature of the interacting exogenous and endogenous processes to be taken into account. Based on two information criteria, the resulting best transfer function contains winter NAO with a lag of three years, annual AMO with a lag of one year (both as exogenous climate factors), loge(spawning–stock biomass) as a structural model component, plus two autoregressive parameters. The model is characterized by the smallest information criteria, 92% of explained recruitment variation (vs. 55% from the simple Cushing-type model), excellent forecasting behaviour, and all model assumptions being fulfilled. It is proposed that the model's recruitment hindcasts (*ex post* forecasts) and forecasts be incorporated into stock and risk assessments as well as management strategy evaluations, either as a climate-induced recruitment index for projections or as real forecasts to establish sustainable cod fisheries on Georges Bank conditioned by climate as a forcing factor.

# 5 Developing a framework for best practice

## 5.1 Understanding the mechanism (linkages)

### 5.1.1 Early life-history and recruitment processes

The numbers of recruits is considered as the number of individuals that either join the adult population or become vulnerable to the fishery (see definitions above). Starting from the numbers of breeding individuals (represented as Spawning–Stock Biomass, SSB) or probably more importantly the numbers of mature females (female only SSB) a number of eggs are spawned, of which a small percentage survive through to adulthood. In general the major commercially exploited species have a high fecundity with either benthic/demersal eggs (e.g. herring and sandeel) or pelagic eggs (e.g. cod, haddock, plaice, etc). The numbers of eggs spawned is a function of a number of different factors including the number of mature females, the extent of skipped spawning, condition and fat reserves of the females along with the annual investment in fecundity. The eggs develop over time, the development rate generally being species-specific and temperature modulated (see Pepin, 1991; Geffen and Nash, 2011). Egg mortalities are a due to a number of diverse factors, including malformations, disease and predation. The latter is probably stage dependent thus longer development times may result in large absolute losses through the egg phase.

The termination of the egg phase occurs at hatch resulting in the larval phase. The initial larval phase is usually supported by endogenous nutrition through the use of yolk-sac reserves. The end of the yolk-sac stage results in the larva being fully dependent on exogenous feeding and the presence of adequate quantities of suitable prey. Predation is also still a major influence on survival through the larval phase. Growth and hence duration of the larval phase depends on both available prey quality and levels along with the effect of temperature on metabolic rates. Once again stage duration can have a fairly large impact on the absolute losses of individuals through the larval phase.

Autumn spawned fish e.g. North Sea herring generally undergo the winter period still in the pelagic larval phase. Often, there is very limited prey availability and feeding conditions e.g. light levels are relatively poor. Most spring-spawning fish metamorphose in the early summer, arriving on nursery grounds and undergoing settlement. The onset of the juvenile phase is often associated with a shift in habitat and thus a change in diet and behaviour. This is the most extreme in the case of flatfish where metamorphosis is a radical change in body shape and there is a major shift from a three-dimensional pelagic habitat to what is essentially a two dimensional benthic existence (Geffen *et al.*, 2007). Nursery ground existence can be accompanied by elevated densities and predation from a range of larger predators. Growth is again influenced by productivity and environmental conditions of the nursery ground i.e. both prey levels and physical factors e.g. temperature. The overwintering period is accompanied by a scarcity of prey and adverse environmental conditions and may result in elevated mortality rates, especially for smaller individuals within a cohort.

The remaining period of time through to recruitment varies depending on the species and which age or development period is utilized for recruitment in the stock assessment. In the case of e.g. cod then losses during the subsequent two juvenile years can occur through cannibalism (density-dependent effects and spatial dynamics between the adult and juvenile population).

Much of the above is summarized in Houde (2002, 2008) and Nash and Geffen (2011) with relevant references given therein.

There are a number of ways to visualize the early life-history dynamics of fish. The one method shown here is through the use of Paulik diagrams (Paulik, 1973; Ulltang, 1996; Nash, 1998). This report concentrates on the prediction of recruitment and the classical stock to recruitment relationship is illustrated in quadrant 4 (top right hand part of the four panel graph; Figure 1).

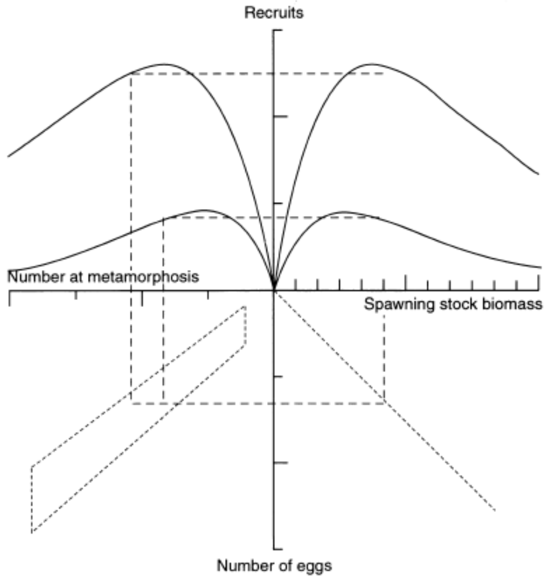



**Figure 1. Conceptual model of plaice life history, redrawn from Nash (1998). Solid lines and dashed box indicate the range of abundance in each quadrant. Quadrant 4 is the classic stock and recruitment plot. Quadrants 1 to 3 follow the development of the spawning–stock through to recruitment using a number of transitions (SSB to eggs, eggs to larval metamorphosis and metamorphosis to recruitment). Each of the three quadrants therefore corresponds to processes occurring internally to the fish (Quadrant 1, egg production), the pelagic phase (Quadrant 2) and the nursery ground (Quadrant 3). The summation of Quadrants 1 to 3 is Quadrant 4 i.e. Stock to recruitment. The present conceptual model accepts non-stationarity in the system and the ranges accept that the productivity of the various phases may change between years, not all changing in tandem.**

## 5.2 Developing the statistical/Mathematical Model

### 5.2.1 Principal model types

Before selecting a recruitment model, the purpose of the model to be identified must be clarified beforehand, i.e. the intention whether it should serve explaining a mechanism which specifies a more or less clear cause–response relationship or whether it should be purely used for prognostic reasons.

An endemic problem to be resolved in this context is the model type selection; i.e. the clear formulation of the biological working hypotheses and research foci, their translation into statistical hypotheses ($H_0$, $H_1$) and the derivation of the principal model type from these.

#### 5.2.1.1 Causality and mechanistic models

Causal models are designed to explain causal mechanisms. If the model to be established aims at explaining causal mechanisms it is important to note that imposing causality always implies imposing a causal direction which restricts the number of suitable methods to those which establish asymmetric relationships; causal models in a true sense usually address unidirectional functional relationships of type $y = f(x)$ or $output = f(input)$ where $y$ or *output* represent the endogenous variable (response variable) and $x$ or input the exogenous one (factor). Sometimes input does not only represent one variable, but a set of variables (vector of variables). These type of models are then called multiple or multifactor models. If also output represents a vector of variables then we have it to do with a multivariate case.

In contrast to asymmetric or causal models, symmetry is a feature of non-directed (a-functional) relationships that are usually identified by procedures originating in the family of correspondence analyses being usually based on association techniques (association measures). These include techniques such as simple and multiple as well as canonical correlation. It should be noted that correlation techniques are in principle linear methods why they cannot be uncritically applied to non-linear cases, i.e. without any specific adjustment.

The clear unidirectional cause may sometimes be confused by feedback effects. Feedback models represent a class of functions that are at minimum bidirectional where y does not only depend on x, but at the same time $x$ on $y$; hence $y = f(x)$ and $x = f(y)$. For feedback models it is rather complicated to correctly specify them and estimate the associated model parameters. Moreover, these usually require many datapoints, hence long time-series. Feedback models play a specific role in control theory and cybernetics. Some of these models are called interdependent models or simultaneous equations.

Another important issue is the type and timing of effects. The effects of factors on recruitment, for instance, can be direct or indirect, i.e. mediated through other components of the system such as the food chain, the adults, or hydrographical and meteorological features. Moreover, they can affect recruitment, for instance, immediately or in a delayed manner. The latter case can often be observed when so called latent factors that are global scale play an important role. These then need some time to penetrate or cascade through all intermediate components of the system before they can affect the endogenous variable, for instance the recruitment. Consequently, it is an important business identifying and taking into account lagged influences of exogenous factors.

Another important issue in this context is the complexity of the model which usually depends on the number of exogenous variables included and the number of parameters to be estimated. The problem here is that exogenous variables (and thus parameters) are not uncorrelated among each other. Otherwise it would be sufficient to establish one univariate model per each exogenous variable, with always the same endogenous variable as response. Correlation means overlapping (doubled, redundant) information and in this sense represents and specifies a joint interaction effect. Incorporating highly correlated exogenous variables would thus over-pronounce the joint effect resulting in an inflation of the variance estimates linked to the model parameters. In an extreme case this can lead to an inability of estimating the parameters (rank loss in the variance-covariance matrix). Given this, it is good practice to select a more parsimonious model with a smaller number of variables and/or parameters, even if the fit of this is a little worse compared with that of a more complex one. Specific performance criteria (information criteria, adjusted coefficients of determination, etc.) may help to resolve the issue of variable selection.

Endemic problems to be resolved in this context are variable selection, multi-collinearity, redundancy, structural constancy or misspecification.

#### 5.2.1.2 Prognostic models

Prognostic models are designed to forecast the future by first identifying the internal characteristics of one or more processes independently of each other rather than by identifying a functional (asymmetric) cause–response relationship between them. In contrast to causal models they do not necessarily make use of exogenous variables, but are solely be based on the information gathered from only one single endogenous variable of interest. These types of models are typically represented by TSA models. The simplest TSA model is a time-trend model based on simple regression. All other TSA models are linear or non-linear extensions of this, starting with so called lag-endogenous models.

However, even if these models are not intended to describe a functional mechanism involving exogenous variables, they also assume a specific minimum variant of causality in the sense that past values, for instance, of recruitment influence present and future values of recruitment but not *vice versa*. This means, that the causality assumption which is normally related to external effects is reduced to the fact that the presence is a reaction to the past; a process is thus considered to be chronologically predetermined prior to the current period, implying that the past contains all relevant information for the future. Consequently, the endogenous variable (for instance, recruitment) is considered to be a response to the environment accumulating all relevant information from past outside effects. This type of causality is called Granger causality. Granger causality implicitly supports the idea of including delayed or lagged effects, even of external factors.

Endemic problems to be resolved in this context are stationarity as well as the correct specification of lags related to, for instance, AR and/or MA components in case of ARIMA models.

#### 5.2.1.3 Combined models

In this context, the focus of combined models is also on forecasting the future. Combined models such as time based regressions models or transfer functions (including intervention functions) such as ARIMAX models are typically prognostic models that are not only based on picking up the features of one endogenous variable representing the process to be studied, but at the same time integrate additional external information represented by exogenous variables (factors). It is thus a model type that combines prognostic and mechanistic features.

The functional relationship between endogenous and exogenous variables as well as potentially delayed influences is usually found by performing one-by-one cross-correlations between both groups of variables being made stationary and pre-whitened.

The most sophisticated combination of endogenous with exogenous variables is represented by so called rational transfer functions. Rational transfer functions are complex transfer functions composed of numerator and denominator polynomials associated with the exogenous variables; they contain much more information regarding the integrated exogenous processes than regular mechanistic models as they do not only take into account the delay-structure of the exogenous effects, but the lag-structure of the integrated exogenous processes themselves.

Endemic problems to be resolved in this context are pre-whitening and cross-correlation (beside all other issues related to prognostic models).

### 5.2.2 Evaluating model choice

Diagnostic tools are required to evaluate model fit to observations, fidelity to underlying assumptions, and to assess model predictions of recruitment.

Standard hypothesis testing methods (F-tests, likelihood ratio test, score test, etc) can be used to compare complicated model (highly parameterized) to less complicated models (McCullagh and Nelder, 1989). Such methods are however, only directly applicable to nested models. Information-theoretic methods, e.g. Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are applicable to non-tested models.

#### AIC(c)

If we define k as the number of parameters in the statistical model and L as the maximized value of the likelihood function for the estimated model, then the AIC is defined as:

$AIC=2k–2\ln(L)$ or alternatively, $AIC=\chi2 + 2k$.

The latter form is often convenient, because most model-fitting programs produce $\chi^2$ as a statistic for the fit.

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. Hence AIC not only rewards goodness-of-fit, but also includes a penalty that is an increasing function of the number of estimated parameters, and thus limits overfitting. The classical AIC, though, is only valid asymptotically. Thus if the number of datapoints is small (such as in recruitment data), then the corrected AIC, AICC must be used. The AICc is defined as:

$AICC =AIC + 2k(k+1)/(n-k-1)$,

Where n is the number of datapoints.

Burnham and Anderson (2002) strongly recommend using AICc, rather than AIC, if n is small or k is large. Because AICC converges to AIC as n gets large, AICC generally should be employed regardless, (Burnham and Andersen, 2002).

#### BIC

A convenient formulation of the Bayesian Information Criterion (BIC) is given by

$BIC= \chi^2 + k \ln(n)$.

A comparison of AIC/AICC and BIC is given by Burnham and Anderson (2002, Section 6.4). The authors argue that AIC/AICC has theoretical advantages over BIC. The authors present a few simulation studies that suggest AICC tends to have practical/performance advantages over BIC. In particular, AIC is asymptotically optimal in selecting the model with the least mean squared error, under the assumption that the exact "true" model is not in the candidate set (as is virtually always the case in practice); BIC is not asymptotically optimal.

#### Cross-validation

Cross-validation involves using a ‘training dataset’ (a subset of the total dataset) to estimate the parameters of the model and using the resulting model to predict the remaining data (the 'test dataset'). The ability of the model to predict the test set is used to select the explanatory variables to include in the model. If too many explanatory variables are used, one obtains good model fit for the training data but a model with poor predictive ability because the good fit also includes fitting noise, rather than only signal. If too few explanatory variables are used, the model performance is poor for both test and training datasets. K-fold validation is a version of cross-validation that makes more use of the data than simple cross-validation. In K-fold CV, the data are divided into k-equal parts and the model is run k-times, each time rotating through each of the k-subsets as the test dataset and using the remaining data as the training dataset. A potential drawback with cross-validation: It does not necessarily parallel the likelihood inference that is used to estimate the model parameters. This is because a test criterion is required, and simple least squares are often used (Trevor Hastie *et al.*, 2001). However, the likelihood function may differ from the least-squares criterion (perhaps not, for most of recruitment models in use?)

### 5.2.3 Graphical model performance diagnostics

In general, there are four principal assumptions which justify the use of linear regression models for purposes of prediction:

1) Linearity of the relationship between dependent and independent variables;
2) Independence of the errors (no serial correlation);
3) Homoscedasticity (with respect to time and predictions);
4) Normality of the error distribution.

If any of these assumptions is violated (i.e. if there is non-linearity, serial correlation, heteroscedasticity, and/or non-normality), then the forecasts, confidence intervals, and economic insights yielded by a regression model may be (at best) inefficient or (at worst) seriously biased or misleading.

The SG discussed simple diagnostic tools to help detect model misspecification and also to detect when model performance becomes unreliable because the underlying assumptions no more hold valid (break down), such as when there is change in the data generating statistics.

A selected number of methods are briefly discussed below. Where convenient, example plots are presented, using recruitment models from Arctic Fisheries Working Group (AFWG).

a) Plot standardized residuals vs. fitted (expected) values – Assessing model misspecification

It is usually assumed that in a bivariate or multiple linear regression analysis, the distribution of residuals (observed data, Yobs - predicted data, Ypred), is, in the population, normal at every level of predicted Ypred and constant in variance across levels of Ypred.

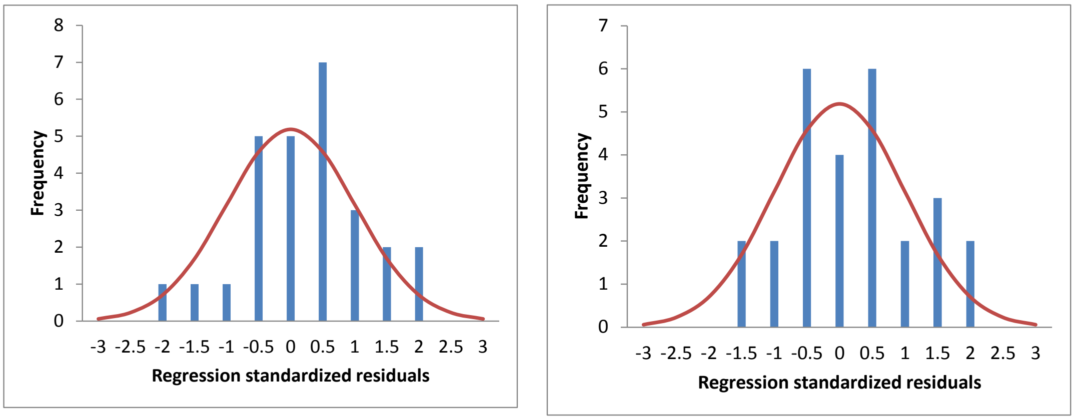



**Figure 2. Example plot of the residuals vs. predicted Y for JES1 model (left) and Titov1 model (right). The patterns indicate no problems with the assumption that the residuals are normally distributed at each level of Y and constant in variance across levels of Y.**

b ) Plot square-root of absolute values of standardized residuals vs. standardized fitted (expected) values – Assess whether variance changes as a function of predicted value (should not)

Homoscedasticity defines a constant variance of the errors with respect to (i) time or (ii) predictions (or vs. any independent variable). The opposite (non constant variance) is referred to as Heteroscedasticity. The presence of heteroscedasticity may have the effect of giving too much weight to e.g. a small subset of the data (namely the subset with the largest error variance) when estimating model coefficients. Plots of residuals vs. time and residuals vs. predicted value should give evidence of heteroscedasticity, i.e. residuals that are getting larger (i.e. more spread-out) either as a function of time or as a function of the predicted value.

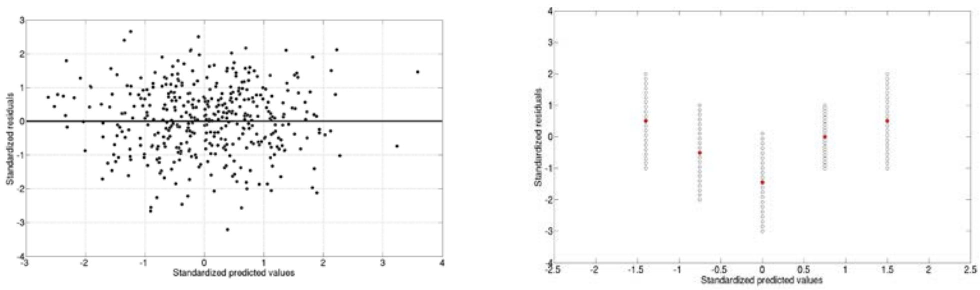



**Figure 3. Example residuals plots— Left: The residuals are skewed towards the top of the plot, which indicates that the residuals are not normally distributed Right: The residuals are mostly positive with low and high predicted values of the dependent variable and mostly negative with medium predictions. A curve with one bend links the mean values (see red dots) of the residuals, suggesting that an appropriate modelling approach for the data is a non-linear (rather than a linear) model.**

c ) Plot observed vs. predicted (expected) values: Assess qualitatively whether explanatory variables are indeed able to reduce variance in the data

Non-linearity is usually most evident in a plot of the observed vs. predicted values or a plot of residuals vs. predicted values, which are a part of standard regression output. The points should be symmetrically distributed around a diagonal line in the former plot or a horizontal line in the latter plot. A "bowed" pattern indicates that the model makes systematic errors whenever it is making unusually large or small predictions.

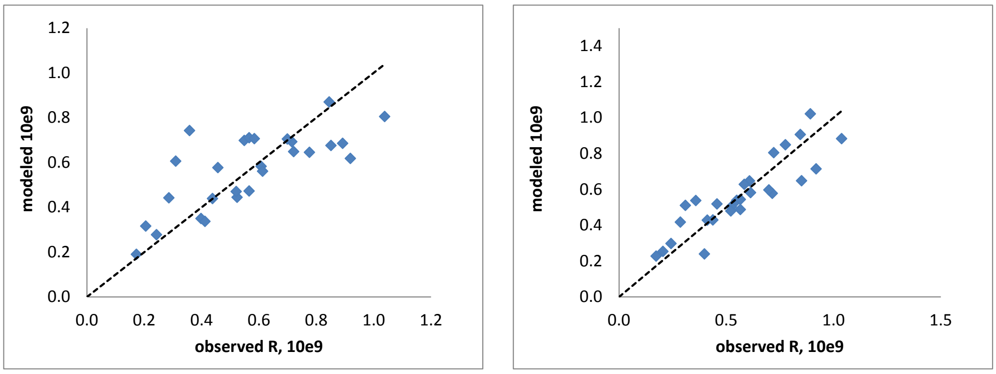



**Figure 4. Plot of observed (horizontal axis) vs. predicted (vertical axis) values for JES1 model (left) and Titov1 model (right).**

d ) Recursive (time varying) parameters and residuals

Non-linear dependence of the level of a series on previous datapoints is of interest because it can indicate the advantage of using predictions derived from non-linear models, over those from linear models, as for example in non-linear autoregressive exogenous models.

Among other types of non-linear time-series models, there are models to represent the changes of variance along time (heteroscedasticity).

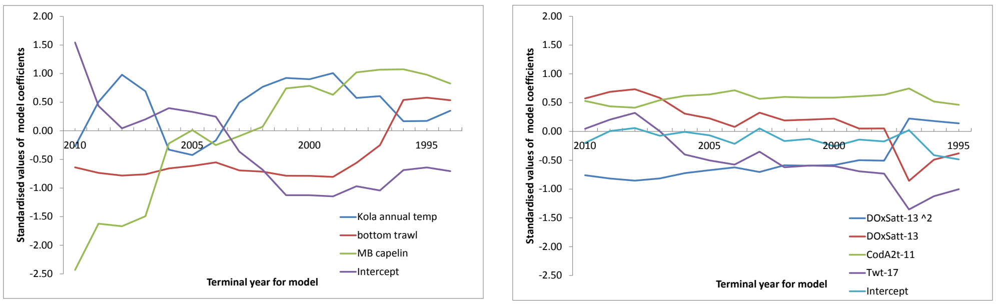



**Figure 5. Plot of model parameter trajectory with time, for the models JES1 (left) and Titov1 (right). Each model is fitted to data from 1984 until the terminal year.**

e ) Q-Q (Quantile-Quantile) plots: Determine whether the residuals are consistent with assumed error model

A Q-Q plot is a plot of the quantiles of two distributions against each other, or a plot based on estimates of the quantiles. The pattern of points in the plot is used to compare the two distributions. The points plotted in a Q-Q plot are always non-decreasing when viewed from left to right. If the two distributions being compared are identical, the Q-Q plot follows the 45° line y = x. If the two distributions agree after linearly transforming the values in one of the distributions, then the Q-Q plot follows some line, but not necessarily the line y = x. If the general trend of the Q-Q plot is flatter than the line y = x, the distribution plotted on the horizontal axis is more dispersed than the distribution plotted on the vertical axis. Conversely, if the general trend of the Q-Q plot is steeper than the line y = x, the distribution plotted on the vertical axis is more dispersed than the distribution plotted on the horizontal axis. Q-Q plots are often arced, or "S" shaped, indicating that one of the distributions is more skewed than the other, or that one of the distributions has heavier tails than the other.

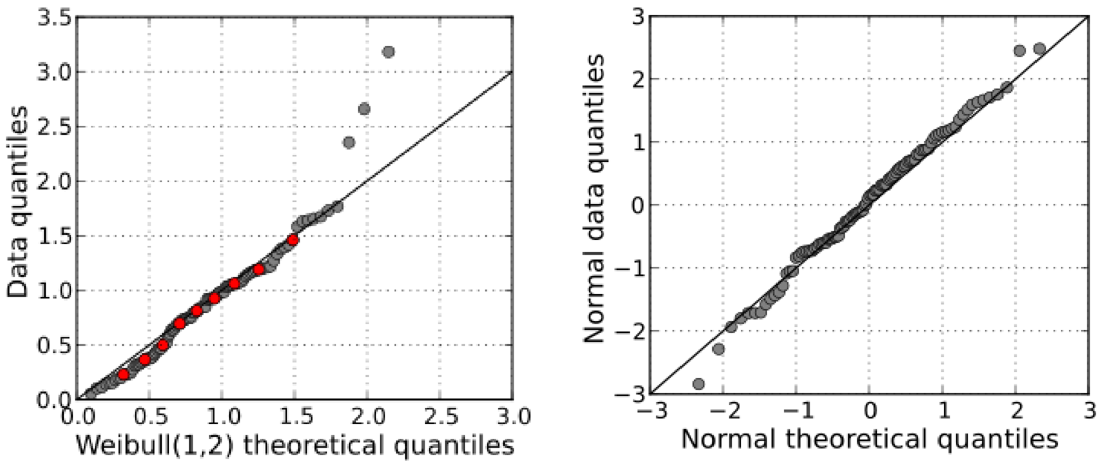



**Figure 6. Left: A Q-Q plot of a sample of data vs. a Weibull distribution. The deciles of the distributions are shown in red. Three outliers are evident at the high end of the range. Otherwise, the data fit the Weibull(1,2) model well. Right: A normal Q-Q plot comparing randomly generated, independent standard normal data on the vertical axis to a standard normal population on the horizontal axis. The linearity of the points suggests that the data are normally distributed.**

# 6 Case studies

## 6.1 Description of the stocks in the case studies

### Northeast Arctic cod (NEA cod)

The main spawning is along the Norwegian coast, mainly north of 67°N, during March–April. Most of the larvae drift into the Barents Sea, where the cod spend the rest of its life (Figure 7), except for the spawning migration. Age of recruitment to the fisheries has been defined as age 3. The cod can reach an age of at least 20 years and a size greater than 130 cm. Economically it is the most important fish stock in the area, with typical annual catches between 400 and 800 thousand tonnes. The cod is an opportunistic feeder, eating most available species of suitable size; however it seems that it prefers capelin where it is available. In years when capelin abundance is low or the cod density is high, cannibalism may cause a substantial mortality on juveniles. Seals and whales are important predators of cod.

### North Sea cod (NS cod)

The North Sea cod stock is widely distributed over the North Sea. In general the younger age classes (1 and 2) are generally found in the southern part and adults tend to be concentrated in discrete groups. The North Sea cod stock is really a metapopulation where each of the components fluctuates in abundance and contribution to the overall stock. This information is summarized in the stock annex of the ICES WG report (ICES 2009). The metapopulation structure is reflected in the spatial separation of the spawning areas and there are probably significantly different trends in the SSB for each component. The majority of the spawning occurs between January and April. Maturity occurs between 4 and 5 years old with fishing mortality high from age 2 onward. At present the SSB is relatively low. The juveniles (O groups) currently are predated by grey gurnards but several other predators contribute to the predation mortality, namely whiting and seabirds.

### Norwegian spring-spawning herring (NSS herring)

During the 1950s and 1960s overwintering concentrations were observed in the Norwegian Sea. However, since the 1970s (collapse and recovery of the NSS herring stock) adult herring generally overwinter along the North Norwegian west coast (Figure 7). In late winter the herring migrate to the spawning grounds along the Norwegian coast (mainly between 62 and 67°N in recent periods). After spawning the herring starts on a feeding migration into the Norwegian Sea, mainly feeding on copepods and euphausiids. The eggs are benthic and after hatching the larvae move up into the water column where they are advected along the Norwegian coast. Some larvae enter the coastal fjords and others are dispersed across the Barents Sea, where they spend their first three years. The Barents Sea juveniles feed mostly on zooplankton, but can also feed on capelin and cod larvae. Important predators on herring are seals, whales and to some extent cod, saithe and seabirds. The herring is one of the largest fish stocks in the Norwegian Sea/Barents Sea and is economically very important. Annual catches have varied from nearly zero to 2 million tonnes.

### North Sea Autumn Spawning herring (NSAS herring)

The North Sea autumn spawning herring is composed of a number of components that currently are considered as Orkney/Shetland, Buchan, Banks and Downs. Each of these components has a distinct spawning location, spawning generally starting in August/September in the north (Orkney/Shetland) and finishing in winter in the south (Downs). The larvae generally drift eastward from the spawning grounds toward the Skagerrak and German Bight. In this species the overwintering stage is as larvae. The juveniles spend the early part of their life in relative shallow coastal waters, moving further offshore as they age. Recruitment to the adults stock and sexual maturity is generally between 2 and 3-winter rings (3+ years old). Much of the background information and historical record for this stock is summarized in Dickey-Collas *et al*. (2010). It should be noted that in NSAS herring the ages are given in winter rings and because they are autumn spawners there is no winter ring for the first winter. Thus to match up the age with the correct year class and hence SSB it is necessary to go back the age (in winter rings) plus 1 to get the correct SSB i.e. a 3wr fish in 2000 is from the 1996 year class and SSB.

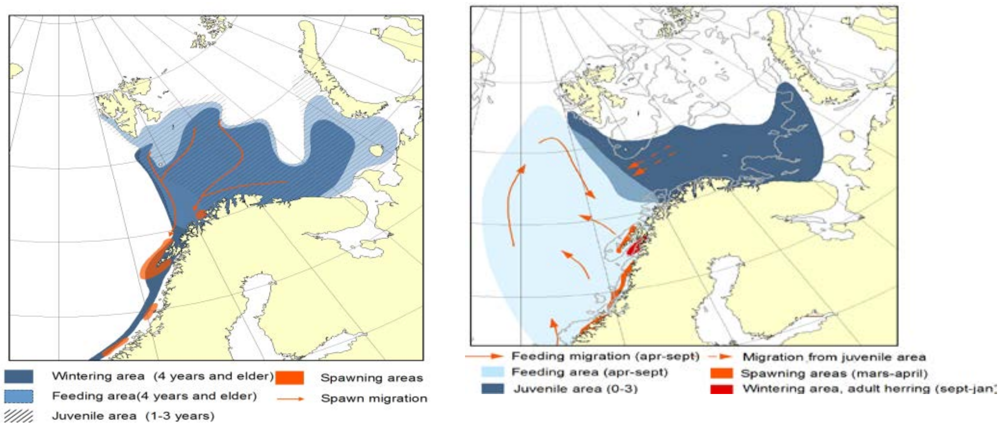



**Figure 7. Spawning, hatchery and feeding area of NEA cod and NSS herring.**

## 6.2 Abundances of early life-history stages during the recruitment process of the stocks in the case studies

### 6.2.1 Cod

#### North Sea (NS) cod

Numbers-at-age and SSB were obtained from ICES (2007d). Egg productions estimated as laid out in Kell *et al*. (in prep) using the mean-weights-at-age in the stock as reported by ICES (2007d) and assuming males and females had the same weight-at-age and there was a 50:50 male to female sex ratio. The fecundity relationship was obtained from the EU project RASER (O.S. Kjesbu, A. Thorsen, IMR, Bergen and P.J. Witthames, Cefas, Lowestoft, pers. comm.).

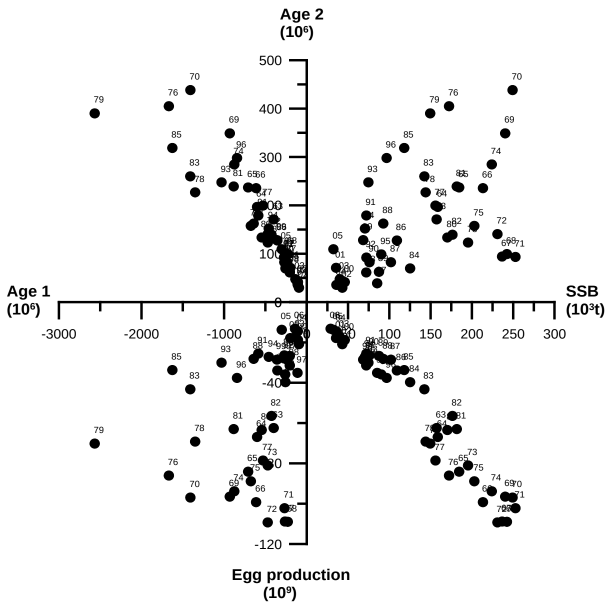



**Figure 8. Paulik diagram for North Sea cod. Estimated egg production (Kell *et al*., pers. comm.) and age 1 and age 2 abundance data and the estimated SSB from the VPA.**

## Northeast Arctic (NEA) cod

Numbers-at-age and survey indices were taken from ICES (2007c). Egg productions were estimated here based on the formulae and procedures laid out in Marshall *et al*. (2004, 2006). The most recent relationships e.g. length–weight were obtained from Marshall CT (University of Aberdeen, pers. comm.). Planktonic early life-history stage abundances were obtained from Russian surveys (N. Yaragina, PINRO, pers. comm.).

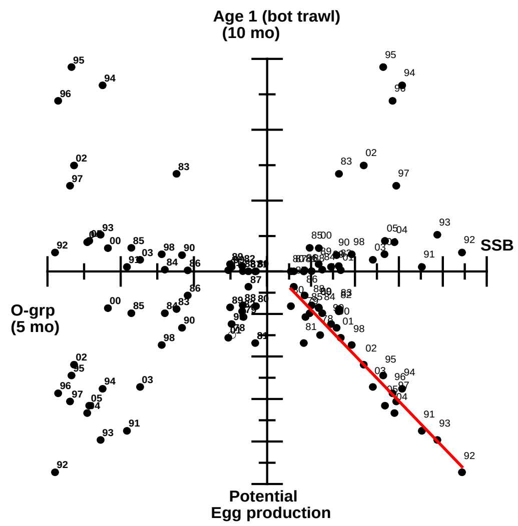
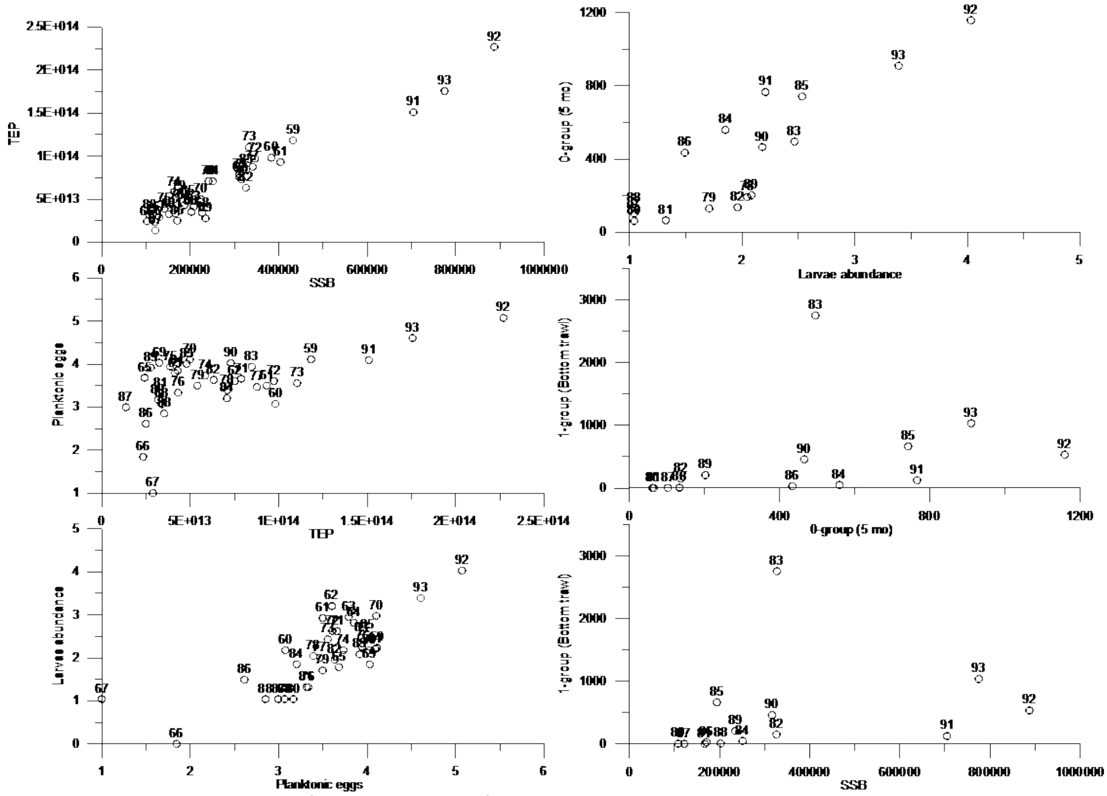



**Figure 9. Paulik diagram for northeast Arctic cod. Survey estimates of O- and 1-group juveniles are used (note that surveys of O- and 1-group juveniles started in the early 1980s).**




*Planktonic egg and larvae data from N. Yaragina (PINRO, Murmansk, Russia)*

**Figure 10. Multi-panel life-history transition graph for northeast Arctic cod. Data on early life-history phases from Russian surveys.**

## 6.2.2 Herring

### North Sea autumn spawning (NSAS) herring

The methods for estimating egg production are laid out in Nash and Dickey-Collas (2005). The original Paulik diagram was updated using data given in ICES (2007a).

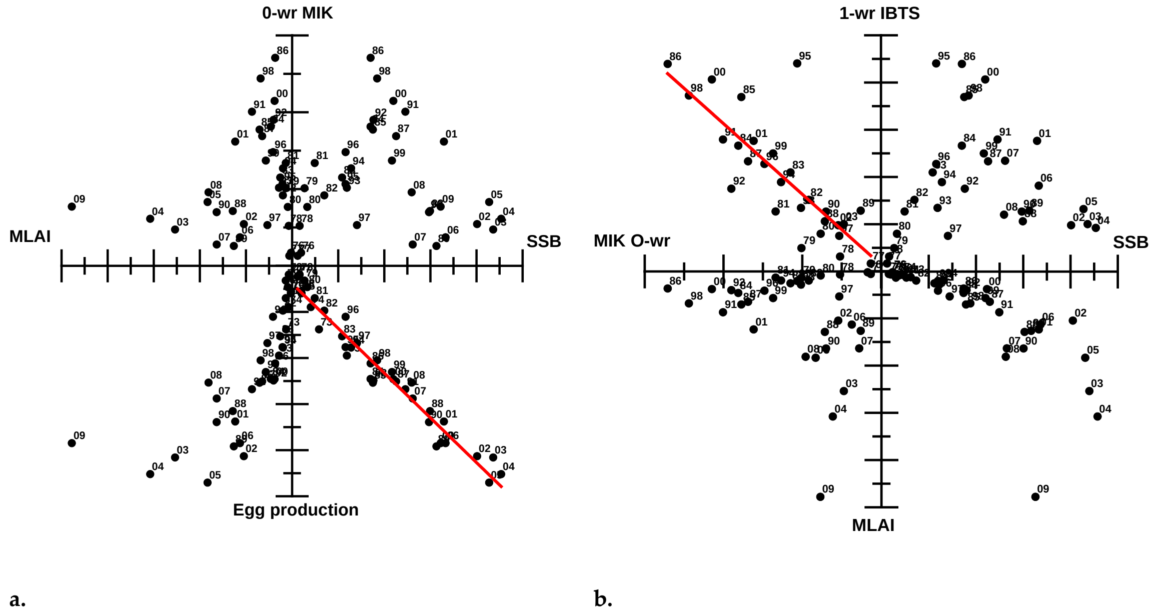



**Figure 11. Paulik diagrams for North Sea autumn spawning herring. a. utilizing SSB to egg (fecundity transition) with recruitment-at-age 0-wr b. utilizing SSB to larvae transition and recruitment at 1-wr. See text for methods.**

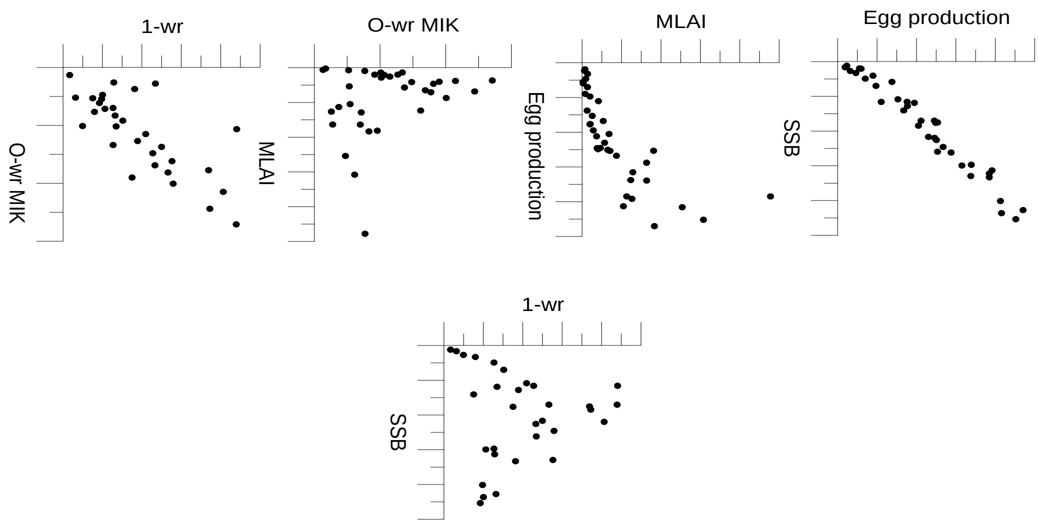
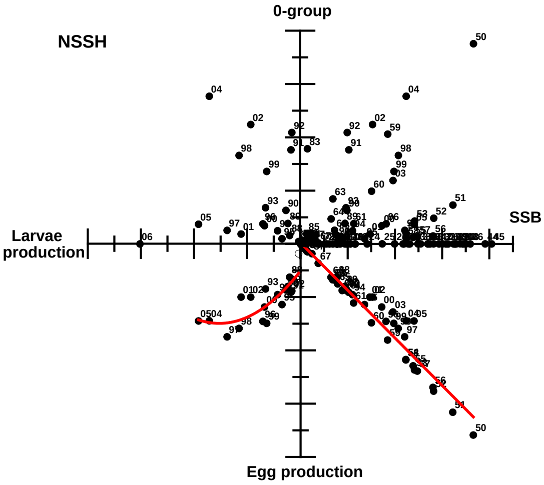


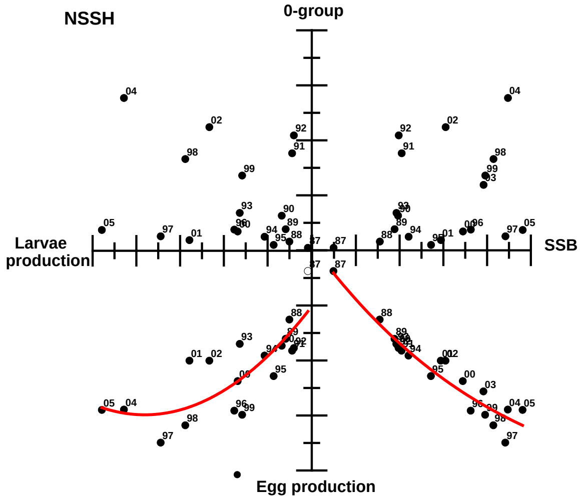

**Figure 12. Life-history transitions for North Sea autumn spawning herring. Each graph represents a life-history stage from egg production, early pelagic phase, first-feeding/overwintering, remainder of first year of life and then the summation of all life-history stages in recruitment with SSB to the end of the second winter period.**

## Norwegian Spring-spawning (NSS) herring

Data on larvae abundance (1987–2006) and VPA numbers-at-age and SSB (1950–2006) were taken from ICES (2007b). Egg production was estimated from the numbers and mean weights-at-age given in ICES (2007b), assuming males and females had the same weight-at-age and there was a 50:50 sex ration, using the formula presented in Óskarsson *et al*. (2002).




a.




b.

**Figure 13. Paulik diagrams for Norwegian Spring-spawning herring. a. with all available data (stock to recruit panel is for 1907–2006). b. Data restricted to the period 1987 to 2006 as this is the period when larvae abundance has been estimated.**

# 7 Preliminary conclusions and further directions

This preliminary report has identified the key components for defining a framework for best practice in recruitment modelling and forecasting.

In particular, the SG has identified that when applied to large statistical populations, prognostic estimates from TSA can be very accurate, although they might be solely based on one given time-series. As there is no single right forecasting method to use, method selection (Trend models or time regression, ARIMA models, transfer functions, Kalman-filtering, spectral analysis, etc.) should be based on objectives and conditions (type, scale and length of dataseries, number and type of variables, etc.).

As the knowledge of the future is incomplete, quantifying risk and uncertainty are important components to forecasting. It must be considered as good practice to indicate the degree of uncertainty associated with a given forecast. This can be done in terms of specifying the forecast errors or forecast intervals. In contrast to confidence intervals a forecast error or interval always contains a component that quantifies the distance of the forecasted value from the centre of the supporting data. Thus the greater the distance, the wider the intervals and the more uncertain is the forecast. Reliable forecasts can only be made for a few time units ahead (normally for only one time unit), depending on the time unit itself (years, quarters, months, etc.), the length of the TS as well as the lag structure of the TS model specified and the exogenous variables if included (ARMA, transfer function).

It is however imperative that any models or projections of recruitment must satisfy the basic principles of ecology and population dynamics, irrespective of the statistical or mathematical approach. This must be the ultimate litmus test for applicability of a recruitment model in stock assessment.

There are a number of basic tenets, in this respect, that need to be adhered to. For instance, in relation to the bounds of recruitment (1) recruitment cannot be less than zero and (2) the number of recruits can never be greater than the total egg production. The consequence of the latter statement is that the numbers of breeding adults can have an influence on the level of recruitment and the scope for very large numbers of recruits will be impacted to a greater extent at lower population sizes. There is probably also a minimum population size whereby successful recruitment can occur i.e. Allee or depensatory effects will occur. The absolute number of pre-recruits is probably determined by the carrying capacity of the habitats occupied prior to joining the adult phase, similar constraints will apply to the adult population thus density-dependent factors need to be considered at higher population sizes. Both the biotic and abiotic environment can influence the survival of early life-history stages and the controls or drivers of survival are most probably not single factors. The principal drivers may not always be the same or in the same combination each year, the ecosystem fluctuates.

The results presented in this report are preliminary, and only address issues related to a framework for short-term forecasts. Subsequent meetings will aim at finalizing the framework and applying it to the case studies identified in this report. Further work will also seek to extend the analysis and case studies to medium and long-term recruitment forecasts.

# 8 Recommendations

The SG suggests meeting in Barcelona in October 2012. The group also suggests an extension of the meeting duration from four days (at present) to at least twice as long (eight days), to allow for adequate time to address the ToRs.

# 9 References


Beverton, R. J. H. and S. J. Holt. 1957. On the dynamics of exploited _sh populations, Fisheries Investigations (Series 2), volume 19. United Kingdom Ministry of Agriculture and Fisheries, 533 pp.

Box, George; Jenkins, Gwilym. 1976. Time-series analysis: forecasting and control, rev. ed., Oakland, California: Holden-Day.

Brillinger, D. R. 1975. Time-series: Data analysis and theory. New York: Holt, Rinehart. & Winston.

Brezinski, C. and Zaglia M. R. 1991. Extrapolation methods, theory and practice. Studies in computational mathematics, Vol.2, North-Holland, Amsterdam, 1991, 464p.

Burnham, K. P., and Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag. ISBN 0-387-95364-7.

Burnham, K. P., and Anderson, D.R. 2004. "Multimodel inference: understanding AIC and BIC in Model Selection", Sociological Methods and Research, 33: 261–304.

Chen, D.G., and Ware, D.M. 1999. A neural network model for forecasting fish stock–recruitment. Can. J. Fish. Aquat. Sci. 56: 2385–2396.

Claeskens, G, and N.L. Hjort. 2008. Model Selection and Model Averaging, Cambridge.

Cook, R.M. 1998. A sustainability criterion for the exploitation of North Sea cod. ICES J. Mar. Sci. 55: 1061–1070.

Cushing, D.H. 1973. Dependence of recruitment on parent stock. J. Fish. Res. Bd Can. 30, 1965.

Dickey-Collas, M., Nash, R.D.M., Brunel, T., van Damme, C.J.G., Marshall, C.T., Payne, M.R., Corten, A., Geffen, A.J., Peck, M.A., Hatfield, E.M.C., Hintzen, N.T., Enberg, K., Kell, L.T. and Simmonds, E.J. 2010. Lessons learned from stock collapse and recovery of North Sea herring: a review. ICES Journal of Marine Science 67: 1875–1886.

Evans, G. T., and Rice, J.C. 1988. Predicting recruitment from stock size without the mediation of a functional relation. ICES J. Cons.,44: 111–122.

Fahrmeir, L., Hamerle, A., and Tutz, G. 1996. Multivariate statistische Verfahren. Walter de Gruyter, Berlin, Germany.

Fang, Yixin. 2011. "Asymptotic equivalence between cross-validations and Akaike Information Criteria in mixed-effects models", Journal of Data Science, 9:15–21.

Geffen, A.J. and Nash, R.D.M. 2011. Egg development rates for use in egg production methods (EPMs) and beyond. Fish. Res.

Geffen, A.J., van der Veer, H.W. and Nash, R.D.M. 2007. The cost of metamorphosis in flatfishes. Journal of Sea Research 58: 35–45.

Gjøsæter, H. and B. Bogstad. 1998. Effects of the presence of herring (*Clupea harengus*) on the stock–recruitment relationship of Barents Sea capelin (*Mallotus villosus*). Fisheries Research 38:57–71.

Gröger, J. P., Kruse, G. H., and Rohlf, N. 2010. Slave to the rhythm: how large-scale climate cycles trigger herring (*Clupea harengus*) regeneration in the North Sea. – ICES Journal of Marine Science, 67: 454–465.

Gröger, J. P., and Fogarty, M. J. 2011. Broad-scale climate influences on cod (*Gadus morhua*) recruitment on Georges Bank. – ICES Journal of Marine Science, 68: 592–602.

Helfenstein, U. 1991. The Use of Transfer Function Models, Intervention Analysis and Related Time-series Methods in Epidemiology, International Journal of Epidemiology Vol. 20, No. 3, pp. 808–815.

Houde, E.D. 2002. Mortality. Chapter 3. pp. 64–87 In: L.A. Fuiman and R.G. Werner (eds). Fishery Science. The unique contributions of early life stages. Blackwell Science.

Houde, E.D. 2008. Emerging from Hjort's Shadow. J. Northw. Atl. Fish. Sci. 41, 53–70.

Hurvich, C. M., and Tsai, C.-L. 1989. "Regression and time-series model selection in small samples", Biometrika, 76: 297–307.

ICES. 2010. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak - Combined Spring and Autumn (WGNSSK). ICES CM 2009/ACOM:10.

ICES. 2011a. Report of the Arctic Fisheries Working Group (AFWG). ICES CM 2011/ACOM:05.

ICES. 2011b. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK). ICES CM 2011/ACOM:13.

ICES. 2011c. Report of the Working Group on Widely Distributed Stocks (WGWIDE). ICES CM 2011/ACOM:15.

ICES. 2011d. Report of the Herring Assessment Working Group for the area South of 62 deg N (HAWG). ICES CM 2011/ACOM:06.

Iles, T.C. 1994. A review of stock–recruitment relationships with reference to flatfish populations. Neth. J. Sea Res. 32(3/4), 399–420.1976.

Ingvaldsen, R., Loeng, H, Ottersen, G, and Ådlandsvik, B. 2002. The Barents Sea climate during the 1990s. ICES Marine Science Symposia, 219:160–168.

Jacobson, L. D., and MacCall, A.D. 1995. Stock–recruitment models for Pacific sardine (*Sardinops sagax*). Can. J. Fish. Aquat. Sci. 52:566–577.

Jiao, Y., Schneider, D., Chen, Y., and Wroblewski, J. 2004. An analysis of error structure in modelling the stock–recruitment data of gadoid stocks using generalized linear models. Canadian Journal of Fisheries and Aquatic Sciences,61: 134–146.

Kassandra Fronczyk, Athanasios Kottas, Stephan Munch. Flexible modelling for stock–.recruitment relationships using Bayesian nonparametric mixtures

Lütkepohl, H. 1993. Introduction to multiple time-series analysis. Springer-Verlag, Berlin, Germany.

Martel, L., Caron-Malenfant, E., Vezina, S. and A. P. Belanger. 2007. Labour Force Projections for Canada, 2006-2031. Canadian Economic Observer, 20(6), 1–13.

McCullagh, P. and J.A. Nelder. 1989. Generalized Linear Models. Chapman and Hall: London. (mathematical statistics of generalized linear model).

Meyer, R., and Millar, R. 2001. State–Space Models for Stock–Recruit Time-series. The Sixth World Meeting of the International Society for Bayesian Analysis (2001).

Michielsens, C.G.J., and McAllister, M.K. 2004. A Bayesian hierarchical analysis of stock–recruitment data: quantifying structural and parameter uncertainties. Canadian Journal of Fisheries and Aquatic Sciences, 61: 1032–1047.

Morgan, M.J., A. Perez-Rodriguez, F. Saborido-Rey. 2011. Does increased information about reproductive potential result in better prediction of recruitment? Can. J. Fish. Aquat. Sci. 68:1361–1368.

Myers, R.A. and Barrowman, N.J. 1995. Time-series bias in the estimation of density-dependent mortality in stock–recruitment models. Can. J. Fish. Aquat. Sci. 52(1), 223–232.

Nash, R.D.M. 1998. Exploring the population dynamics of Irish Sea plaice, *Pleuronectes platessa* L., through the use of Paulik diagrams. J. Sea Res. 40, 1–18.

Nash, R.D.M. and Geffen, A.J. 2011. Mortality through the early life-history of fish: What can we learn from European plaice (*Pleuronectes platessa* L.)? J. Mar. Syst.

Nelder J. A. and R. W. M. Wedderburn. 1972. Generalized Linear Models Journal of the Royal Statistical Society. Series A (General) Vol. 135, No. 3 (1972), pp. 370–384.

Nelissen, J. H. M. 1991. Household and education projections by means of a microsimulation model. Economic Modelling, 8(4), 410–511.

Paulik, G.J. 1973. Studies of the possible form of the stock–recruitment curve. Rapp. P-V Réun. Cons. Int. Explor. Mer 164, 302–315.

Pepin, P. 1991. Effect of temperature and size on development, mortality, and survival rates of the pelagic early life history stages of marine fish. Can. J. Fish. Aquat. Sci. 48, 503–518.

Pindyck R. S., Rubinfeld, D. L. 1991. Econometric Models and Forecasts. McGraw Hill, New York, 596 pp.

Post, E. 2004. Time lags in terrestrial and marine environments. In Marine Ecosystems and Climate Variability, pp. 165–167. Ed. by C. Stenseth, G. Ottersen, J. W. Hurrell, and A. Belgrano. Oxford University Press, Oxford, UK.

Quinn II, T. J. and R. B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press.

Ricker, W. 1975. Computation and interpretation of biological statistics of fish populations. Technical Report Bulletin 191, Bulletin of the Fisheries Research Board of Canada.

SAS/ETS Software. 1991. Applications Guide 1, Version 6, 1st Edition: Time-series Modelling and Forecasting, Financial Reporting and Loan Analysis, Cary, NC: SAS Inc., 380 pp.

Schlittgen, R. 2001. Angewandte Zeitreihenanalyse. Oldenbourg Verlag, München, Germany.

Schlittgen, R., and Streitberg, B. H. J. 2001. Zeitreihenanalyse. Oldenbourg Verlag, München, Germany.

Shepherd, J.G. 1997. Prediction of year-class strength by calibration regression analysis of multiple recruit index series. ICES Journal of Marine Science, 54: 741–752. 1997.

Shepherd, J.G. 1982. A versatile new stock–recruitment relationship for fisheries and the construction of sustainable yield curves. J. Cons. Int. Explor. Mer 40, 67–75.

Stiansen, J.E., A. Aglen, B. Bogstad, H. Loeng, S. Mehl, O. Nakken, G. Ottersen and E. Svendsen. 2005. Including climate into the assessment of future fish recruitment, using multiple regression models. ICES CM 2005/O:25.

Stock, J. H., M. W. Watson. 2007. Introduction to Econometrics, 2nd ed., Pearson.

Tereshchenko, V.V. 1996. Seasonal and year-to-year variations of temperature and salinity along the Kola meridian transect. ICES C.M. 1996/C:11.

Titov, O. V. 2001. System approach to the analysis of natural long-term changes of the Barents Sea ecosystem. Murmansk. PINRO. P. 119 (in Russian).

Trevor Hastie, Robert Tibshirani and Jerome Friedman, "Elements of Statistical Learning: Data Mining, Inference and Prediction" Springer-Verlag, New York.

Ulltang, Ø. 1996. Stock assessment and biological knowledge: can prediction uncertainty be reduced? ICES J. Mar. Sci. 53, 659–675.

Van Imhoff, E. and Post, W. 1998. Microsimulation methods for population projection [Special Issue New Methodological Approaches in the Social Sciences]. Population: An English Selection, 10(1), 97–138.

Vujic, S. 2009. Econometric Studies to the Economic and Social Factors of Crime, Rozenberg Publishers.

Walters, C.J. 1990. A partial bias correction factor for stock–recruitment parameter estimation in the presence of auto-correlated environmental effects. Canadian Journal of Fisheries and Aquatic Sciences, 47: 516–519.

Wang, Xiaolan L., Francis W. Zwiers, Val R. Swail. 2004. North Atlantic Ocean Wave Climate Change Scenarios for the Twenty-First Century. J. Climate, 17, 2368–2383. doi: 10.1175/1520-0442(2004)017<2368:NAOWCC>2.0.CO.

Wolfe, J., and Crookall, D. 1998. Developing a scientific knowledge of simulation/gaming. Simulation & Gaming: an International Journal of Theory, Design and Research 29 (1): 7–19.

Wootton, R. J. 1990. Ecology of Teleost Fishes. Fish and Fisheries Series 1, Chapman & Hall.

Woth, K. 2005. North Sea storm surge statistics based on projections in a warmer climate: How important are the driving GCM and the chosen emission scenario?, Geophys. Res. Lett., 32, L22708, doi:10.1029/ 2005GL023762.

Yang, Y. 2005. "Can the strengths of AIC and BIC be shared?" Biometrika, 92: 937–950.

## Annex 1: Participants list

| Name | Address | Phone/Fax | E-mail |
|---|---|---|---|
| Joachim Gröger | Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries Institute for Sea Fisheries<br>Palmaille 9<br>22767 Hamburg<br>Germany | Phone +49 4038905266<br>Fax +49 4038905263 | joachim.groeger@vti.bund.de |
| Yuri Kovalev | Knipovich Polar Research Institute of Marine Fisheries and Oceanography(PINRO)<br>6 Knipovitch Street<br>183763 Murmansk<br>Russian Federation | Phone +7 8152 472 469<br>Fax +7 8152 473 331 | kovalev@pinro.ru |
| Richard Nash | Institute of Marine Research<br>PO Box 1870<br>Nordne<br>5817 Bergen<br>Norway | | Richard.Nash@imr.no |
| Henrik Sparholt | International Council for the Exploration of the Sea<br>H. C. Andersens Boulevard 44–46<br>DK-1553 Copenhagen V<br>Denmark | | henriks@ices.dk |
| Jan Erik Stiansen | Institute of Marine Research<br>PO Box 1870<br>Nordnes<br>5817 Bergen<br>Norway | Phone +47 55 238 626<br>Fax +47 55 238 687 | jan.erik.stiansen@imr.no |
| Samuel Subbey<br>Chair | Institute of Marine Research<br>PO Box 1870<br>Nordne<br>5817 Bergen<br>Norway | Phone +47 5523 5383<br>Fax +47 5523 8687 | samuel.subbey@imr.no |

## Annex 2: Draft Terms of Reference 2012

2011/10/ACOM31 The **Study Group on Recruitment Forecasting** (SGRF), chaired by Samuel Subbey, Norway, will be established and be held in Barcelona, XX–XX October 2012 (eight days) to:

a ) Develop a framework and methodology for modelling of recruitment for use in short- and medium-term projections in stock assessment, incorporating abundance indices and environmental driver;

b ) Test this framework using designated case studies;

c ) Provide comprehensive guidelines for model choice, validation and diagnostics for recruitment forecasting in the short and medium term.

SGRF will report by 31 October 2012 for the attention of ACOM.