Contribution to the Special Issue: 'Commemorating 100 years since Hjort's 1914 treatise on fluctuations in the great fisheries of northern Europe'

Quo Vadimus

Modelling and forecasting stock – recruitment: current and future perspectives

Sam Subbey[1,2]*, Jennifer A. Devine[1], Ute Schaarschmidt[1,2], and Richard D.M. Nash[1]

[1]Institute of Marine Research, PO Box 1870, Bergen 5817, Norway
[2]Institute of Informatics, University of Bergen, PO Box 7800, Bergen 5020, Norway

*Corresponding author: tel: +47 46 83 68 23; fax: +47 55 23 85 31; e-mail: samuels@imr.no

This paper presents a brief review of the present state of knowledge in stock–recruitment forecasting, including process and current methodological challenges to predicting stock–recruitment. The discussion covers the apparent inability of models to accurately forecast recruitment even when environmental covariates are included as explanatory variables. The review shows that despite the incremental success in the past hundred years, substantial challenges remain if the process of modelling and forecasting stock–recruitment is to become relevant to fisheries science and management in the next 100 years.

Keywords: environment, forecasting, modelling, review, stock–recruitment.

Introduction

Recruitment is the result of many factors that affect survival from the egg-stage (including parental effects) through to when individuals recruit to the fishery or stock, i.e. a stepwise process through parent, egg, larva, juvenile, and back to adults, where abundance at one stage is a function of abundance at a previous stage (Paulik, 1973; Rothschild, 1986; Ulltang, 1996). The state of a population in any given year is a function of the stock (e.g. reproduction and growth) and recruitment, which is itself a function of past events (e.g. state of the stock, environmental conditions). Recruitment is therefore explicitly linked not only to the amount of spawning-stock biomass (SSB), which is often used as a proxy for reproductive potential or egg production, but also parental size, growth history, and environment of the individual.

When providing strategic, rather than operational advice, it is clear that greater knowledge is required about the variability and trends in fish recruitment. This is because operational advice, in the strictest sense, is concerned with the ground-level approaches of putting management plans into action. Therefore, it tends to be short-term focused, and requires information over a short period, e.g. whether stock levels have decreased, increased, or remained level since last assessment. Strategic advice, on the other hand, tends to be for the long haul (e.g. stock rebuilding plans, maintenance of stock levels) and covers the "what" and "why" of fisheries management decisions. Such decisions will require knowledge of the variability and trends in recruitment, to evaluate the possible effects of management decisions on stock trend variability and to assess the efficacy of decisions in meeting long-term goals.

The issue of recruitment forecasting (for strategic and operational advice) has been a focus for fisheries research over the last hundred years. Unfortunately, it is often now dismissed as an unreachable goal because of the various mechanisms interacting to influence dynamics throughout the life of an individual. Many of these mechanistic links are either poorly known or the link with recruitment holds for only a short period (Myers, 1998), during which the strength can be intermittently strong/weak. Attempts to understand and forecast outcomes, e.g. survival and recruitment, are now in progress through the use of complex and sophisticated individual-based modelling techniques (Peck and Hufnagl, 2012). However, these models still rely on a basic recognition and understanding of the key drivers and mechanisms which influence survival through all prerecruit life stages.

Here we present a brief review of the present state of knowledge in stock to recruitment modelling and forecasting, including methodological challenges to predicting recruitment and potentially its

linkages with the parent stock. We discuss several reasons for the current apparent inability to forecast recruitment accurately, especially when environmental covariates are included as explanatory variables. The main questions we ask are (i) what are the influences on recruitment? (ii) do we need to understand the factors influencing each life stage to predict recruitment?, and (iii) do we have the necessary tools to model recruitment? Finally, we examine preconditions necessary for stock–recruitment modelling and forecasting to be recognized as relevant to fisheries science and management in the next 100 years.

Biological mechanism and recruitment drivers

The construct of stock to recruitment is a method of predicting recruitment levels from a known stock size. For various reasons, the SSB is taken as a proxy for the total egg production or stock reproductive potential (see Trippel, 1999). By using stock size and recruitment in a single model, it was implicit that the parent stock size had an influence on the numbers of young surviving to join the parent stock or fishery (Ricker, 1954; Beverton and Holt, 1957). At low stock sizes, recruitment is primarily driven by density-independent factors and thus recruitment increases monotonically with stock size. However, at large stock sizes, other factors, e.g. density-dependent effects (to varying degrees) are more influential on the survivorship of young. This method of estimating recruitment levels into the future is primarily for the purpose of modelling and is not really designed for understanding early life-history dynamics. The necessity to derive such relationships is driven essentially by stock assessment and management needs, such as accurately quantifying the exploitable segment of the population and sustainable stewardship of fisheries resources. It also provides a method of closing the loop in fish population modelling whereby generational input of young fish (new year classes) can be estimated.

A shift in the linkage between SSB and egg production can occur through interannual variations in individual fecundity (McBride *et al.*, 2013), non-participation in spawning by part of the mature population, i.e. skipped spawning (Rideout and Tomkiewicz, 2011; Skjæraasen *et al.,* 2012), or through viability of offspring due to the structure of the mature population (Marshall *et al.,* 2010) Any of these can lead to the errors-in-variables problem when using SSB as the metric to predict productivity (Rothschild and Fogarty, 1989; Morgan *et al.,* 2011). A stock–recruitment relationship (SRR) that fails to consider other factors or their interactions does not incorporate a large amount of biological realism, which leads to an inability to accurately quantify the true effect of factors influencing recruitment.

Stock–recruit models often fail to adequately show the link between spawners and recruits or the link weakens when new or longer datasets are added (Myers, 1998). Factors not directly a consequence of the spawning stock can impact cohorts within the same stock differently and this lack of consistency is often viewed as a breakdown of the relationship. Hutchinson (2008) surmised that the reason for SRR failure is due to either incorrect assumptions regarding the spawning stock or failure to account for the spatio-temporal scales of factors describing the relationship. Furthermore, changes in population demography (Wright and Trippel, 2009; Fitzhugh *et al.,* 2012) may result in a dissociation between SRR and mechanistic relationships, e.g. Atlantic cod SRR with temperature, the North Atlantic Oscillation (NAO), or dynamics of the North Atlantic Subpolar Gyre (see Ottersen *et al.,* 2010). Many SRRs have the assumption of stationarity and so do not effectively deal with shifts in productivity of the ecosystem or the stock itself (Nash *et al.,*

2009). Moreover, system complexity and non-linear interactions between factors, which act to prevent the observed variable from behaving in an expected way, or mask, mitigate, or intensify specific information types, affect the perceived performance of SRRs.

A number of authors have shown that incorporating community dynamics (e.g. prey to-predator loops, Bakun and Weeks, 2006; cannibalism, Hjermann *et al.,* 2007) was more important than solely using environmental factors for accurate predictions of recruitment. The effect of the NAO in the Barents Sea typically reinforces oceanic responses, causing overall higher water temperatures, but it may be that it is not solely temperature that is responsible for strong year classes; the NAO influences the influx of Atlantic Water, which imports not only warmer water but also food (zooplankton) for juvenile cod (Ottersen *et al.*, 2010). Furthermore, one type of interacting term may not be strongly related to recruitment dynamics until another term is added to the model. This was the case for the Northern Gulf of St. Lawrence Atlantic cod (*Gadus morhua*) population, where herring (*Clupea harengus* L.) predation was not strongly correlated with cod survival until the effect of temperature was included (Duplisea and Robert, 2008). Fish live in multispecies communities and the effect of interactions between species, or between fish of different sizes within the same species, should not be overlooked. A reversal in the predator–prey dynamic between Atlantic cod and pelagic fish (e.g. herring, mackerel (*Scomber scombrus*)) appeared to drive cod recruitment in the 1970s and 1980s in the Southern Gulf of St. Lawrence, where, when pelagic fish reached high abundance, they preyed heavily upon the eggs and larvae of cod (Swain and Sinclair, 2000). Another type of trophic dynamic exists in the Barents Sea; good year classes of herring result in high mortality of their prey, capelin (*Malotus villosus*) larvae, which acts to deplete the capelin stock (at a lag), leaving little food for adult cod, which then may, depending on cod stock size, turn cannibalistic, resulting in poor cod recruitment (Hamre, 2003). Cannibalism has also recently been suggested for the North Sea autumn spawning herring population as a consequence of a demographic shift in the abundance of substock abundances and a potential overlap of adults with a part of the larval drift pattern (Corten, 2013).

Early life-history dynamics

Traditional stock–recruitment models use the endpoint (i.e. the number of individuals recruiting to the fishery) of a complex relationship that integrates several processes operating over multiple spatio-temporal scales (Rothschild, 2000), but whether this encompasses enough detail is questionable. Stock–recruitment is not only a function of the stock, but is intrinsically linked to past events that influence growth and mortality, not only of the individual recruiting but also of the parents. Recruitment must decline if there is insufficient spawning biomass, but recruitment will also decline with reduced body size, as fecundity, egg size, and spawning extent is inextricably linked to the fish's growth history and condition (Kjesbu *et al.*, 1996). Any factor changing the demography of the population will impact reproductive potential, and hence recruitment of the stock. This multi-scale complexity then argues for a life-stage approach to stock–recruitment modelling (Rothschild, 2000).

An integrated life cycle approach is one that investigates the importance of different mechanisms acting on various life stages and identifies the critical stages and mechanisms for recruitment (e.g. Nash, 1998, and references therein). By partitioning recruitment relationships into different life stages, scale can be explicitly investigated by examining the numerical variability in the relationship, as well as the interaction between life stages and the environment, and

the non-linear dynamics regulating each stage (Rothschild, 2000). This approach has been successfully used to investigate recruitment for several stocks, e.g. North Sea herring (Nash and Dickey-Collas, 2005); Northeast Arctic cod (Mukhina *et al.*, 2003); and Barents Sea stocks (Dingsør *et al.*, 2007). Paulik (1973) highlights this in his seminal work on predictive SRR whereby the various life stages are taken into account when progressing from stock to eventual recruitment.

Paulik diagrams, in their present form (as initially presented by Nash, 1998), provide a graphical format for indicating where bottle-necks or shifts in survival are occurring in the life cycle. An example of this is shown in North Sea autumn spawning herring (Payne *et al.*, 2009), where a major shift in survivorship occurred in the first winter of life. From around 2001 to the present, overwintering survival is considerably less than in previous years. In Paulik's original paper (Paulik, 1973), he assumed there would be recognizable relationships between life-history stanzas that could be modelled. The presentations by Nash (1998), Nash and Dickey-Collas (2005), Payne *et al.* (2009), and Nash and Geffen (2012) indicate that the situation is more complex and relationships between life-history stanzas can vary quite considerably.

The impact on environmental factors may be crucial for the early life stages. The link between environment and early life stage survival has been thoroughly investigated through match–mismatch hypothesis (Cushing, 1982, 1990), critical period hypothesis (Hjort, 1914), ocean stability hypothesis (Lasker, 1981), optimal environmental window (Cury and Roy, 1989), and ocean triad (Bakun, 1998). However, larval abundance alone may not be an adequate predictor of recruitment (Stige *et al.*, 2013). Furthermore, the environmental link may be indirect; the gadoid outburst in the North Sea occurred when stock levels were not extremely high, but when environmental conditions were good for the prey of early life-history stages of gadoids, namely the prey abundance, seasonal timing, and mean size (Cushing, 1984; Beaugrand *et al.*, 2003). Shifts in productivity or survival through the early life history during different thermal regimes was also indicated by Nash (1998) and Nash and Geffen (2012) in Irish Sea plaice (*Pleuronectes platessa*).

In early life history dynamics, there is a recognition that both density-dependent and independent factors occur, but the relative importance of the two shift through ontogeny (Beverton and Iles, 1992a, b). Both density-dependent and -independent processes are assumed to be occurring within a certain recruitment interval, i.e. the period between spawning and recruitment, and all mortality outside of this interval (i.e. post-recruitment) is considered density independent. This assumption is an oversimplification because the length of the recruitment interval is often not based on an understanding of mechanisms or underlying processes, which is an important consideration when attempting to determine the impact of specific mortality influences (e.g. fisheries) on multiple stages (Brooks and Powers, 2007).

By using multiple stages, both the level and type of mortality on each stage can be allowed to vary before recruitment. The question of whether this matters for the timing of compensatory processes, i.e. whether the timing (early or late) in the stage, has a strong effect on the numbers of recruits. As Brooks and Powers (2007) pointed out, additional mortality from exogenous sources will have a larger impact if occurring later in the stage, after compensatory responses have occurred, regardless of whether the Beverton–Holt or Ricker model is used to investigate recruitment. Various authors have illustrated the differences in mortality rates through the early life-history stages (e.g. Nash and Geffen, 2012) and highlighted the

fact that the absolute level of mortality is a function of both the instantaneous rate and the stage-duration.

Modelling stock – recruitment

In the fisheries literature, recruitment models can be classified as being parametric, semi-parametric, or non-parametric. Parametric recruitment models are analytical or semi-analytical functional expressions for recruitment. Three subclasses of parametric approaches to modelling stock–recruitment are identifiable from the fisheries literature, which we classify as classical, Box–Jenkins type, and state-space models. The first subclass includes the two parameters Beverton and Holt (1957) and Ricker (1975); the general recruitment function by Deriso (1980), Cushing (1973), Iles (1994), and Shepherd (1982); and the Sigmoidal Beverton–Holt (Myers and Barrowman, 1995) models. A good summary of the most widely used parametric recruitment models can be found in Needle (2001). Included in this subclass is cases where one of the classical functional relationships (Beverton–Holt or Ricker) is augmented with secondary data (climatic and ecological) such as temperature and prey interactions (see, e.g. Gjøsæter and Bogstad, 1998; Planque and Frédou, 1999; Olsen *et al.*, 2011). The second subclass of parametric recruitment functions is those in which the dependent variable (here recruitment) is regressed on one or several (often time-lagged) independent dataseries, including time-lagged values of the independent variable. This subclass of time-series models (also referred to as Box-Jenkins models, Box *et al.*, 1994) consider recruitment as a combination of autoregressive and moving average effects, leading to an autoregressive moving average model (see, e.g. Gröger *et al.*, 2010; Gröger and Fogarty, 2011). The third subclass involves the use of state-space model formulations to link recruitment to population parameters. In general, discrete-time state-space models are defined by two equations, namely, the observation (or measurement) equation and the system (or transition) equation. This model class, based on linear difference relationships between the input and output variables, provides and offers the additional flexibility of including parameters that are exogenous to the model. An example is the Bayesian state-space stock–recruitment model for Fraser River pink salmon (*Oncorhynchus nerka*) (Meyer and Millar, 1999). The state-space approach addresses two major problems encountered in traditional stock–recruitment analyses that of errors-in-variables bias and time-series bias. In Meyer and Millar (1999), 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 non-parametric methods derive the recruitment relationship based on less stringent assumptions than those implied by the use of parametric approaches. This class of models spans a wide range including modelling the distribution of recruitment as a function of biomass by non-parametric density estimators (Evans and Rice, 1988), locally weighted smoothing functions with non-parametric regression, LOESS smoothers and spline methods (Cook, 1998; Bravington *et al.*, 2000; Cadigan, 2013; Munch *et al.*, 2005) and neural networks (Chen and Ware, 1999). Extensions of this class of models to include environmental variables has also been reported in the literature (see, e.g. Jacobson and MacCall, 1995).

This manuscript focuses on establishing an SRR outside the stock assessment model, and involves fitting functional relationships to stock–recruit data. The fitting process results in deterministic (expected recruitment) and stochastic (characterized by residuals resulting from the fit) components of recruitment. It must be

mentioned, however, that a common practice is to estimate the recruitment relationship within the model used to assess the stock (see, e.g. Maunder and Deriso, 2003; Lee *et al.*, 2012). When recruitment relationships are integrated into assessment models, the annual recruitment is usually decomposed into two components: an average annual recruitment (synonymous to the relationship obtained by an analytical fit to stock–recruit data) and a stochastic component (e.g. residuals from a functional fit). The stochastic component is usually defined by an assumed probability distribution function, and the parameters for the recruitment relationship (the annual average recruitment and parameters of the probability distribution) are usually integrated into the total objective function for the assessment model. The optimized probability distributions are then used to characterize the uncertainties in the annual recruitment, as well as in recruitment projections. The integration of SRRs can be extended to the case where environmental correlation with recruitment or a stock–recruitment model is being considered (Maunder and Starr, 2001; Maunder and Watters, 2003). See Maunder and Deriso (2003) for instance, for several estimation methods involving catch-at-age assessment models with integrated recruitment estimation components.

Functional SRRs

The classical approach to stock recruitment assumes the existence of a functional relationship, $F(\cdot)$, between spawners, S_t, and recruits, R_t, over some indexed time (usually years) of $t = 1,\ldots,n$. This relationship can be expressed in a compact form as

$$R_t \equiv F(S_t, \Theta), \tag{1}$$

where Θ is a vector of parameters, usually of direct relevance to the fishery management policy (Hilborn and Walters, 1992; Quinn and Deriso, 1999; Chen, 2004).

Deriso (1980) introduced the general three-parameter SRR in (2), which was further developed by Schnute (1985). The resulting Deriso–Schnute model,

$$R_t = \alpha S_t (1 - \gamma \beta S_t)^{1/\gamma}, \tag{2}$$

is such that

$$R_t = \begin{cases} S_t e^{\alpha - \beta S_t}, & \lim \ \gamma \to 0 \\ \dfrac{\alpha S_t}{1 + \beta S_t}, & \gamma = -1. \end{cases} \tag{3}$$

The models resulting from setting $\lim \gamma \to 0$ and $\gamma = -1$ define, respectively, the Ricker (1975) and the Beverton and Holt (1957) models, which are the most commonly used stock–recruitment models in fisheries science. The parameter α is the density-independent parameter, and β measures the level of density-dependence. The shape parameter γ is, however, not limited to $\gamma \leq 0$. For instance, $\gamma = 1$ leads to the Schaefer model (Schnute, 1985; Schaefer, 1991) for recruitment. Also, note that we can derive the following equation (dropping the time subscript) from (2):

$$\frac{d(R/S)}{dS} = -\frac{\beta}{1 - \beta \gamma S} \left(\frac{R}{S} \right).$$

Hence the compensatory property, $d(R/S)/dS < 0$ is only guaranteed for $\beta(1 - \beta \gamma S) > 0$ for all S. For a comprehensive discussion, see Quinn and Deriso (1999).

The Beverton–Holt and Ricker models both describe a decrease of the per capita recruitment with increasing stock size. Mortality is assumed to be density independent at low stock sizes and there is scope for some underlying level of density-dependent mortality at any stock size. The choice of the stock–recruitment model is based upon the severity of the density-dependent mortality (the existence of "overcompensation") believed to influence population dynamics most during the period before recruitment. 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. It also assumes the ever presence of predators. 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 is based on the assumption that the mortality rate of the eggs and juveniles is proportional to the initial cohort size. The ecological conditions that result in a Ricker-type recruitment include (i) cannibalism of the juveniles by the adults, (ii) disease transmission, (iii) damage by adults of one another's spawning sites, (iv) density-dependent reductions in growth coupled with size-dependent predation, and (v) a time-lag in the response of a predator or parasite to the abundance of the fish (Ricker, 1975; Wootton, 1990).

At low stock sizes, depensatory effects may occur. The decrease of per capita recruitment is commonly referred to as the Allee effect (Allee *et al.*, 1949). In terms of modelling, the Allee effect is a zero-recruitment term, which represents an offset of the recruitment function from the origin (see Frank and Brickman, 2000). Models incorporating depensatory effects may be three-parameter generalizations of the Ricker (see, e.g. Chen *et al.*, 2002), the Beverton–Holt model (see, e.g. Liermann and Hilborn, 1997), or the Saila–Loda model (Saila *et al.*, 1988).

In the literature, the classical equations (Beverton–Holt and Ricker) have also been reformulated to provide direct information about the productivity and resilience of a fish stock, and hence to management. The reformulation incorporates a parameter, referred to in the literature as *steepness*, and defined as the proportion of virgin recruits produced by 20% of the virgin spawning stock, (see, e.g. Mace *et al.*, 1988; Brooks and Powers, 2007). First, let h, R_0, E_0, ψ_0 represent the steepness, virgin recruitment, virgin egg production, and virgin level of egg production per recruit ($\psi_0 = (E_0/R_0)$), respectively. We denote by $R_i(t)$ the number of fish surviving to time t of stage i ($0 \leq t \leq q\Delta t_i$), where Δt_i ($i = 1,2,\ldots,n$) is a unit duration of the recruitment process, assumed to occur over n discrete stages. Then the number of fish at time 0 of stage 1 is the number of eggs produced, i.e. $E \equiv R_1(0)$.

The reparametrized Beverton–Holt and Ricker functions are then defined by

$$R = \begin{cases} \dfrac{4hR_0 E}{\psi_0 R_0 (1 - h) + (5h - 1)E}, & \text{Beverton-Holt} \\ \dfrac{E}{\psi_0} (\beta \psi_0)^{(1-E)/E_0}, & \text{Ricker.} \end{cases} \tag{4}$$

A high steepness value is indicative of a resilient population, which is robust to harvesting, including a high probability of rebuilding when fishing pressure is relaxed. It must be mentioned, however, that despite its attractiveness, the steepness parameter is difficult

to estimate because there is often little information in the data on this parameter (Lee *et al.*, 2012).

Incorporation of environmental effects

Demographic uncertainty is the result of inherent temporal variability in population processes due to environmental stochasticity. The goal of including environmental covariates in stock–recruitment modelling is to reduce the degree of unexplained stock–recruitment variation caused (at entry to the fishery) by demographic variation. When building recruitment models, a challenge is the question of whether or how environmental influences should be included in the model definition. The most common models for incorporation of environmental variables $Z_1, \ldots, Z_n$ in SRRs are modified Ricker or Beverton–Holt functions (Hilborn and Walters, 1992; Chen and Irvine, 2001) of the form

$$R_t = S_t e^{\alpha - \beta S_t + c_1 Z_1 + c_2 Z_2 + \cdots + c_n Z_n}, \tag{5}$$

$$R_t = \frac{\alpha S_t}{1 + \beta S_t} e^{c_1 Z_1 + \cdots + c_n Z_n}. \tag{6}$$

Of these two models, the modified Ricker model is the type of relationship more commonly investigated. This is perhaps because similar to a lognormal error; this allows for a linear and additive formulation of the model, thus affording computational ease. This type of model has, for instance, been used by Planque and Frédou (1999) and Drinkwater (2005) to describe recruitment of Irish Sea cod. The effects of predators and competitors of Northwest Atlantic cod have been studied using a linear regression of $\log(R/S)$ against SSB of cod, herring, and mackerel (Swain and Sinclair, 2000). A hybrid Beverton–Holt–Ricker model was used to incorporate zooplankton abundance and a log-linear effect of temperature (Olsen *et al.*, 2011).

Myers (1998) pointed out that environment–recruitment relationships are seldom used in stock assessment. For instance, Shepherd *et al.* (1984) listed 42 stocks in the North Atlantic and Northeast Pacific, for which the environmental impact on recruitment had been studied, but only one correlation had become part of stock assessment. Furthermore, Myers (1998) found that correlations, in general, performed poorly on retesting with longer time series, with exception of relationships at population's geographical limits. The problem of data-dependence on choice of tests for correlations and their results has also been highlighted by Iles and Beverton (1998).

More generally, a major drawback with models of the forms (5) and (6) is that they are usually linear approximations of non-linear environmental effects. This approximation may be inadequate or unreliable. For example, Drinkwater (2005) and Planque and Frédou (1999) found evidence for a non-linear relationship between temperature and recruitment. For cold-water species, warmer sea temperature would strengthen recruitment and for warm-water species, the opposite. However, the actual response was dependent on where in the latitudinal range of the species the observations were made. A quadratic relationship between sea surface temperature (SST) and the logarithm of the reproductive success was reported by Stocker *et al.* (1985). Such a dome-shaped functional relationship cannot be described by a log-linear model such as (5) or (6). The logarithm of a linear function can be a strictly monotonic decreasing or increasing function, but does not allow for both a positive and a negative slope of the environment–recruitment curve. This fact is of special significance, since the

rejection of a linear relationship does not necessarily imply the total absence of one, but that the functional form may be non-linear. On the other hand, suppose the influence of an environmental parameter (e.g. temperature), T, on recruitment is defined as the exponential of an n-dimensional polynomial with coefficients $a_0, a_1, \ldots, a_n$, and given by (7)

$$R_t = f(S_t) \exp\left(a_0 + \sum_{j=1}^{n} a_j T^j\right), \tag{7}$$

where $f(S_t)$ is a function of S_t. Then the effect of T on R_t will be not unique, but depend on the values of the coefficients of the polynomial. For the simplest case of $n = 2$ (quadratic polynomial) recruitment can be lower both at low and high temperatures (for fixed values of S_t and T), depending on the values of a_1 and a_2.

It is worth mentioning that functional forms other than log-linear have been suggested by Iles and Beverton (1998). Semi-parametric and non-parametric models incorporating environmental parameters have also been reported in the literature. For instance, Chen and Irvine (2001) used semi-parametric models of the form (8), where $g(E)$ denotes a linear or quadratic function of one or several environmental factors E and $f(S_t)$ is a linear function of S_t.

$$R_t = S_t e^{\alpha + f(S_t) + g(E)}. \tag{8}$$

Fitting stock–recruitment functions
Error assumptions

Although some authors (see, e.g. Jiao *et al.*, 2004) have modelled the SRR assuming normal, lognormal, gamma, and Poisson model errors, the predominant assumption is that the errors in the spawner–recruitment relationship are usually lognormally distributed (Hilborn and Walters, 1992). A stochastic version of the classical Beverton–Holt and Ricker functions, given lognormal errors, can be formulated as

$$R_t \equiv F(S_t, \Theta) e^{\epsilon_t}, \tag{9}$$

where $\epsilon \sim N(0, \sigma^2 I)$. The assumption of lognormal errors in the SR analysis can be attributed to principal contributions by Allen (1973), Walters and Hilborn (1976), Peterman (1978), Hennemuth *et al.* (1980), and Peterman (1981). The basis for the lognormal assumption, however, derives from empirical studies (see, e.g. Allen, 1973; Peterman, 1981) or driven by the need for computational convenience and ease (see, e.g. Hilborn, 1985), or a combination of both factors (Walters and Hilborn, 1976). However, a common practice in fitting the SRR is to first introduce a linearization of the model equation (e.g. by taking the logarithm of R_t, as in Chen, 2004), followed by parameter estimation using a least-squares approach, where the lognormal error assumption provides computational ease.

The literature, however, shows that non-parametric and Box–Jenkins models allow for consideration of other types of error structures other than lognormal (Meyer and Millar, 1999). Generalized linear models (GLMs), which are a flexible generalization of ordinary linear regression, are applicable to several error distributions other than the normal. Jiao *et al.* (2004), for instance, used a GLM to investigate the effect of normal, lognormal, gamma, and Poisson model error assumptions on the derived SRR. Other non-parametric methods in the literature include density estimation

(Evans and Rice, 1988), generalized additive models (Jacobson and MacCall, 1995), locally weighted spline smoothing with non-parametric regression (Cook, 1998), and neural networks (Chen and Ware, 1999). It must be noted that since most non-parametric methods are heavily dependent on asymptotic assumptions, their reliability, when applied to small sample sizes typically available for recruitment series, may be unknown.

Bias and bias correction

In general, the difference between an estimator's average and the true parameter value is called the bias. In stock assessment models, recruitment variability is usually assumed to be lognormally distributed (Maunder and Deriso, 2003). Methot and Taylor (2011) showed that the lognormal assumption calls for a bias correction in the modelled mean annual recruitment. We define the true (but unobserved) and modelled mean annual recruitment values, respectively, by R and R^*, and assume that R^* is lognormally distributed, i.e. $\log(R^*)$ is normally distributed with mean R and standard deviation σ_R. Then the true mean recruitment is considered as a random variable, which is related to the modelled recruitment through

$$R^* = R \exp\left(\left[u\sigma - \frac{1}{2}\sigma^2\right]\right), \qquad (10)$$

where u is a normally distributed random variable with zero mean and unit variance. The term $1/2\sigma^2$ is subtracted to make R^* unbiased, i.e. we have

$$\mathbb{E}(R^*) = R, \quad \mathrm{Var}(\log(R^*)) = \sigma^2. \qquad (11)$$

As stated in Methot and Taylor (2011), the adjusted value $R^* < R$, and represents the median recruitment value.

Formally, an estimator θ^* of a real-valued parameter θ is median-unbiased if

$$\text{Probability } [\theta^*(X) < \theta|\theta] = \text{Probability } [\theta^*(X) > \theta|\theta],$$

for each θ (Birnbaum, 1964), i.e. under- and overestimates are equally likely. On the other hand, θ^* is mean-unbiased if for each θ, the expectation of θ^* is defined such that

$$\text{Expectation } [\theta^*(X)|\theta] = \theta.$$

Both mean-unbiased and median-unbiased estimates of stock–recruitment have been reported in the fisheries literature. Unfortunately, the literature offers no guidelines for which to prefer and why, although the estimators have different attributes and probabilistic expectations. When the estimator of recruitment is to be used as input to different (not all linear) formulae (or models); however, then the median-unbiased estimator is to be preferred (see, e.g. Birnbaum, 1964). On the other hand, if one assumes that the recruitment data errors are normally distributed, then the best unbiased estimator will also be normally distributed. Under such assumption, the median-unbiased and mean-unbiased estimators are mathematically equivalent in any linear regression or multivariate analysis. Median-unbiased estimators are however preferable in situations where the normally distributed error assumption cannot be justified (i.e. asymmetric distributions). This is because the median is then insensitive to skewness and kurtosis (see, e.g. Andrews, 1993; Nelson and Plosser, 1982). Because recruitment data are usually heavy-tailed

(represent asymmetric distribution), a median-unbiased recruitment estimate may offer a precautionary and preferable alternative for short-term predictions or if estimates are to be used for computing reference points. Median-unbiased recruitment estimates are also to be considered when these values are input parameters to other submodels or equations.

For the general problem of fitting the SR function, the literature shows that a density-dependent SR relationship is generally assumed to exist (Sakuramoto, 2005). It has also been recognized that because estimates of numbers of spawners and recruits are uncertain, the recruit–spawner output is incomplete unless accompanied by estimates of uncertainty (Ludwig and Walters, 1981a). Some emphasis has therefore been placed on how to correct for bias in estimated parameters Θ, due to data uncertainty (see, e.g. Ludwig and Walters, 1981a; Walters and Ludwig, 1981; Walters, 1985; Chen, 2004). Two potential effects of uncertainty on recruitment have been reported, namely (i) the estimated parameters may be inconsistent due to violation of the independent error assumption, which underlies the regression equation and (ii) the over-estimation of the information content of the observation data (Ludwig and Walters, 1981b).

For the classical Beverton–Holt and Ricker models, if the errors are lognormally distributed, then the optimal parameter values obtained by multiple linear regression, $\hat{\Theta}$, and the associated variance, $\hat{\sigma}_\Theta^2$, can be easily estimated. The multiple regression approach is, in principle, finding a model that quantifies the expected (average) recruitment for a given stock size. However, this average has been shown to be (in most cases) larger than the modal recruitment (Ricker, 1975; Hilborn, 1985). Consequently, bias is introduced in estimates of Θ, with subsequent bias and uncertainty in management decisions, which are dependent on the parameter estimates. Ludwig and Walters (1981a) developed an elaborate estimation procedure, which is robust both to the limited number of observations that are usually available and to measurement errors. The computational framework relies on an independent assessment of the relative strengths of environmental and measurement variances in providing maximum likelihood estimates of Θ, a measure of uncertainty in Θ and a time series of errors. More recently, Chen (2004) argued that the assumption of lognormal errors in the Ricker and Beverton–Holt models introduces a positive bias into recruitment predictions, which is qualitatively dependent on the structure of the model residual variance, the historical SR data, and the specific spawner biomass used. The bias correction introduced by Chen (2004) is reported to be asymptotically unbiased with a finite-sample bias that is practically zero.

Forecasting stock–recruitment

Recruitment forecasting often involves the use of environmental indices because, often, prerecruit data are either unavailable or highly uncertain. Examples include recruitment forecasting of short-lived species such as anchovy in the Bay of Biscay (*Engraulis encrasicolus* L.), where upwelling conditions caused by north-easterly winds of medium and low intensity in spring–summer in the Bay were linked to good levels of recruitment (De Oliveira *et al.*, 2005), linking SST and recruitment to predict the impact of increasing sea temperatures on future recruitment (Roel *et al.*, 2004), and the use of SST to predict California sardine (*Sardinops sagax*) recruitment (Jacobson and MacCall, 1995). The attempt to predict short-term changes in recruitment of North Sea cod using SST has also been reported by Planque *et al.* (2003). Multiple regression recruitment forecast models have also been developed for

Northeast Arctic cod, Norwegian spawning herring, and Barents Sea capelin. These models combine the water temperature of 3−7 stations of the Kola section transect (layers 0−200 m) (see Bochkov, 1982), the bottom-trawl abundance indices of Northeast Arctic cod (ages 0 and 1), and the biomass of mature capelin (see details in Stiansen *et al.*, 2005). According to Stiansen *et al.*, (2005), the presented models accounted for 65−85% of the variance in the predicted recruitment. For South African anchovy, "fluctuations in copepod biomass and production on the spawning grounds, the incidence of oocyte atresia in adult females, the incidence of southerly winds at Cape Point and the distance offshore of the 16°C isotherm at Cape Columbine were associated with fluctuations in anchovy recruitment" (Cochrane and Hutchings, 1995).

The dominant motivation for recruitment forecasting appears for management planning to ensure optimal harvesting strategies (see, e.g. Sun *et al.*, 2009). This is particularly true for cases where management advice is provided before the level of recruitment can be assessed and/or when management advice involves multi-annual decisions such as total allowable catch (TAC). An example is provided by the management of South African anchovy (*Engraulis encrasicolus*), where an initial TAC is set at the start of the fishing season in January, before the level of the year's recruits of 0-year-old fish can be quantified, and the TAC is readjusted in May or June after the actual recruitment has been estimated (Cochrane and Starfield, 1992). Simulation results show that for this particular stock, the average annual catch could be increased theoretically by up to 48% given very precise (CV = 0.1) short-term predictions of recruitment at the start of the fishing season.

Since management decisions are defined in terms of time horizons, forecasting SR has been classified as being *short*, *medium*, or *long term*. The classification, however, is non-unique, and has to be considered in relationship with the life cycle of the specific stock of interest. Northeast Arctic cod and Northeast Arctic haddock (*Melanogrammus aeglefinus*) stocks, for which recruitment to the fishery occurs at age-3, a short-term forecast is 1−3 years ahead, while for species like Greenland halibut (*Reinhardtius hippoglossoides*) and redfish (*Sebastes mentella*), short-term forecast may be as much as 6 years ahead (see ICES, 2011a). Both short- and medium-term projections are usually based on fitted stock−recruitment functions and recent trends in recruitment. The literature also reports recruitment forecasts based on historical estimates of recruitment, such as the projected short/medium-term recruitment being a geometric mean of past estimates of recruitment, or the geometric mean of past recruitment values that fall below the arithmetic mean of the historic stock−recruitment data (see, e.g. ICES, 2003). For South African anchovies, a short-term forecast is defined in terms of months (see Cochrane and Starfield, 1992). In general, however, short-term forecasts have been considered most reliable, which explains their popular use in stock assessment and in the context of management advice (see, e.g. Bogstad *et al.*, 2000). In contrast, the more uncertain medium- and long-term forecasts are more often used in the context of management strategy evaluations (MSEs; see, e.g. Kimoto *et al.*, 2007).

The literature reports mixed results when evaluating the practical implementation, effectiveness, and impact of forecasted fish recruitment on management decision. For California sardine, SST has been successfully used in predicting stock−recruitment (in the stock assessment) and therefore in the harvest control rules for this particular stock (Jacobson and MacCall, 1995; Deriso *et al.*, 1996). Conversely, using North Sea cod as an example, Planque *et al.* (2003) concluded that currently, there is limited managerial

benefit of forecasting recruitment based on environmental information. This conclusion, however, does not discount the possible use of environmental information to forecast recruitment in the future. It merely emphasizes the fact that for this particular stock, there is limited understanding of the coupling between the environment and recruitment, which prevents reliable recruitment forecasts from environment-based recruitment models. The literature also reports of failure when harvest rules have been derived using environment-based recruitment forecasts. For instance in 1999, the TAC for the Bay of Biscay anchovy was drastically reduced when the projected SSB was found to be well below the precautionary SSB level. The projected SSB was based on predicted recruitment from a model that had previously shown a significant relationship between the upwelling index and recruitment of age-1 fish (see Borja *et al.*, 1998). On re-examination, the forecasted recruitment turned out to have been grossly underestimated. As a consequence, this particular environmental index is no longer considered a reliable indicator of the absolute levels of recruitment (De Oliveira and Butterworth, 2005). Based on simulation results, Cochrane and Hutchings (1995) report the possibility of obtaining a valuable increase in mean annual yield of South African anchovy fishery if below-average recruitment could be forecasted 6 months in advance of obtaining reliable measurements of recruitment, at the start of the fisheries.

Basson (1999) undertook a simulation study to investigate the value of including environmental data for management procedures, and arrived at the conclusion that there is no advantage in incorporating environmental drivers of recruitment in the short term for conservation or yields. Furthermore, there were situations where the incorporation of environmental factors made the management decisions worse. While these conclusions may appear to apply to some of the examples previously discussed (e.g. for the Bay of Biscay anchovy), the fact remains that each species (its life history and dynamics) and fishery needs to be carefully considered before being a candidate for the inclusion of environmental factors in recruitment forecasting.

Outstanding methodological challenges and caveats
Defining recruitment
One of the largest problems in contemporary discussions concerning SRRs is that the recruitment is neither to the fishery nor at the point of sexual maturity. While the point in the life history is not critical, the range of ages used mean that there is no consistency in the meaning of recruitment. In some instances, e.g. North Sea Autumn Spawning herring at age-0-winter rings, the year-class strength is apparent (ICES, 2013b), whereas in Northeast Arctic cod, processes such as cannibalism mean that the year-class strength may not be apparent until age-3 (ICES, 2013a).

Recruitment, in the strict fisheries definition (becoming vulnerable to the fishery or captured), is generally length or size dependent, rather than strictly age dependent. Likewise, recruitment to the adult, i.e. sexually mature, portion of the population is also primarily driven by length; however, age will also have an influence. Recruitment is assumed to be knife edged (in the case of sexual maturity, can be either deemed knife edged or, more commonly, may vary annually and is determined from surveys) and a convenient age class is chosen, often for simplicity, whereas it is well known that there is a selectivity curve for both capture and maturation.

To make matters more complicated, sometimes a year class that is not fully defined in regard to the life history, such as age 0, has been

chosen in the past for age at recruitment to the fishery, e.g. Irish Sea cod (ICES, 2011b). To add to the confusion, spawning for some species occurs on either side of 1 January, i.e. 1 year class effectively spans two different years, e.g. north–east Atlantic herring (ICES, 2013b). In the case of North Sea autumn spawning herring, the 0-winterring (or 0-group) is clearly defined as 1 January, but the larvae are enumerated during a survey in February when they are nominally 3–4 months old (ICES, 2013b). This definition of year class is clearly not appropriate for many winter and spring spawning herring populations.

Modelling stock–recruitment

The existence of a spawners–recruits relationship (and its particular form) is often accepted without question in the fisheries literature. However, empirical evidence that such a relationship may occasionally fail to exist in general has been reported by Sakuramoto (2005) and, for Pacific Salmonids, by Peterman (1978). Myers and Barrowman (1995) undertook a meta-analysis of 364 spawner–recruitment time series and noted that, on average, the highest recruitment occurred at high spawner abundance and the lowest spawner abundance produced the lowest recruitment. These conclusions arising from the meta-analysis contradict, for instance, the (overcompensation) assumptions underlying the Ricker-type recruitment relationship.

Incorporation of environmental effects

A general potential challenge to methodologies for incorporating environmental factors into SRRs is the fact that such dependencies are usually non-linear and asymmetrical. Further, because the time series of data is usually short and noisy, estimation of the underlying coupling between recruitment and environmental parameters is difficult (see, e.g. Lungarella *et al.*, 2007). Attempts to accommodate co-incidental trends in time (Iles and Beverton, 1998) may also lead to development of models with little predictive ability (see Hilborn and Walters, 1992; Chen and Irvine, 2001).

The search for environmental influences on recruitment has resulted in many studies which incorporate physical factors into the SRR, e.g. NAO in North Sea autumn spawning herring (Gröger *et al.*, 2010), temperature with cod (O'Brien *et al.*, 2000), which was taken up in a wider study by Drinkwater (2005). Changes in recruitment have often been found to be indirect responses to changes in third factors (Drinkwater, 2005). For instance, temperature may cause shifts in the composition of the zooplankton community, and thus indirectly strengthen or weaken recruitment (Beaugrand *et al.*, 2003). Furthermore, factors such as upwelling, salinity, and sea level might correlate with temperature and have thus been hypothesized to be the actual causes of changes in recruitment (Prager and MacCall, 1993). Because the impact of temperature may often be an indirect response to changes in many other parameters, temperature has been used as a proxy measurement of other environmental factors (see Jacobson and MacCall, 1995; Myers, 1998, and references therein). Temperature has also been shown to influence the spawning stock, the effects of which will eventually be seen in recruitment. For instance, Ottersen *et al.* (2006) investigated change in the relationship between sea temperature and recruitment of Atlantic cod over time and found that a decrease of length and age of average spawners occurred at the same time as a strengthening of the impact of temperature on recruitment. Direct temperature influence on spawning times and sites has also been reported by Morgan *et al.* (2013). Thus temperature may have indirect effects

on recruitment, either through its links with the spawning stock or as a proxy for other (unmeasured) environmental variables.

In lieu of models of form (5)–(6) and (8), a model that incorporates the effect of temperature could, in general, have the form described by Equations (12)–(13):

$$R = f(S(E)), \tag{12}$$

$$E_t = g(A_{t-\tau_1}^{(1)}, A_{t-\tau_2}^{(2)}, \ldots, A_{t-\tau_n}^{(n)}), \tag{13}$$

where there is a causal link between S_t and a finite n number of environmental factors, $A^{(j)} \leq ft(j = 1, \ldots, n)$, through E_t. The causal link is indicated by the time-lag, $t - \eta_j$, between E_t and each environmental factor. Thus formulated, the function E_t could be a linear or non-linear function of environmental factors or even proxies (e.g. temperature).

Fitting stock–recruitment functions

It is worth noting that most of the literature on classical models and assumptions of recruitment are dominated by empirical data and analyses of Salmonid stocks (see, e.g. Allen, 1973; Peterman, 1978, 1981). For such stocks (in particular, the sockeye salmon *Oncorhynchus nerka* population in the Skeena River), Walters and Hilborn (1976) cite Allen (1973) in arguing that empirical evidence in support of the lognormal distribution exists, while Hilborn (1985) cites Peterman (1981) for the same reason. It is also worth pointing out that the distribution of recruitment can vary for many reasons, including the influence of stock size or structure (Hsieh *et al.*, 2006; Anderson *et al.*, 2008) or due to productivity regimes (Vert-pre *et al.*, 2013). It is conceivable, however, that this assumption may fail to hold, for instance, for stocks with stable stock sizes, which are less prone to extreme fluctuations. In the statistical description of recruitment for 18 stocks, Hennemuth *et al.* (1980) report that Georges Bank cod and Peruvian anchovy (*Engraulis ringens*) exhibit normal rather than lognormal distributions, while multimodal distributions appeared to be more descriptive of both Georges Bank haddock and Northeast Arctic haddock. According to Hennemuth *et al.* (1980), "most well-known standard, single or even joint probability distribution functions do not account for the frequency of occurrence in the long right-hand tails observed in some cases". If it cannot be established, the assumption of a lognormal distribution may lead to biased parameter estimates, although it affords computational ease (see, e.g. Walters and Hilborn, 1976).

The issue has additional complexity associated with how the error structure in the dependent (recruitment) variable relates to that in the independent variable (spawning stock). Although not explicitly stated in most cases, the assumption in the literature is that the error term on the right-hand side of the recruitment function captures the uncertainty in the dependent variable, which also includes the stochasticity observed in the independent variable. It is not unusual, however, that both the recruitment and SSB are characterized by different variances and error distributions. It is not uncommon that the SR data are characterized by a general lack of contrast in the independent variable (usually SSB) coupled with highly variable recruitment. Then given the assumption that the two variables are independent and normally distributed, the solution obtained will very much depend on the ratio of their variances.

In the fisheries literature, the issue of separate error structures in the recruitment and spawner data failed to receive further attention

after it was first discussed in Ludwig and Walters (1981b). If $(\hat{R}_t, \hat{S}_t)$ represent observations of the true recruitment and SSB values (R_t, S_t), then separable error structures imply that

$$\hat{R}_t = R_t + \varepsilon_t, \qquad (14)$$

$$\hat{S}_t = S_t + \eta_t, \qquad (15)$$

where ε_t and η_t can, for instance, be assumed to be independent and normally distributed. It must be recognized that this is a potential source of bias when deriving SR functions and that standard linear regression methods in such cases will yield suboptimal solutions (Cornbleet and Gochman, 1979); see also discussion in Ludwig and Walters (1981b). The problem is best solved using an errors-in-variables model or methodology, which seeks to account for errors in observations on both the independent ($\hat{S}_t$) and the dependent ($\hat{R}_t$) variables. An example methodology is the Deming regression approach (Martin, 2000). The methodology is equivalent to a maximum likelihood estimation procedure in which the errors for the two variables are assumed to be independent and normally distributed, while the ratio of their variances, $\delta_t = (\varepsilon_t/\eta_t)$, is known (e.g. estimated from data) (Linnet, 1993).

Of all the possible sources of uncertainty, structural uncertainty remains one of the most challenging issues in fisheries science (Hammond and O'Brien, 2001) and, in particular, to the task of fitting stock–recruitment functions to observation data. The choice of recruitment function (e.g. Ricker, Beverton–Holt) may be wrong (Patterson *et al.*, 2001), or there may be aspects of the dynamics that the rigid analytical functions are incapable of capturing (Mace and Sissenwine, 2002). This is understandable given the absence of data and rigorous statistical techniques to explore the basic assumptions of the nature of density dependence and other biological processes that underpin the SRRs. Thus, the choice of recruitment model remains a major source of uncertainty in the fitting process.

For almost identical goodness-of-fit values, Myers *et al.* (1994) obtained different functional results for the Ricker and Beverton–Holt models (Barrowman and Myers, 2000; Williams and Shertzer, 2003). With respect to analytical stock–recruitment functions, a more central and basic issue is the fact that models with a small number of parameters are fit to data. Regardless of whether the Deriso, Beverton–Holt, or Ricker models are used, the fact remains that these analytical functions do not possess enough flexibility in their definitions to allow any deviations from the presumed recruitment–spawning stock relationship, even if the data dictate otherwise. Though the diagnostics of the fitted stock–recruit model should indicate that the fitted model is incorrect, this conclusion may not be exclusively valid (see, e.g. de Valpine and Hastings, 2002), given the nature of the data.

In the literature, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (see Burnham and Anderson, 2002) have been employed in the selection of SRRs (see, e.g. Wang and Liu, 2006; Shimoyama *et al.*, 2007). Indications from the literature are that the use of AIC or BIC as selection criteria can be potentially misleading. While Wang and Liu (2006) compared and found both the AIC and BIC as valid selection criteria, simulation studies by de Valpine and Hastings (2002) indicate that the Ricker model gave a better fit to the data than the Beverton–Holt model, when the AIC was used to select the best-fit model. Furthermore, though the Beverton–Holt model generated the data, the AIC selected the

Ricker model as best fit. For typical fishery data, Zhou (2007) reports the probability of selecting the correct model based on information criteria to be 0.70 and 0.61, respectively, for the Ricker and Beverton–Holt models. In general, the task of choosing the *best* model among a variety of candidates is a statistically challenging and non-trivial problem. For a review, see de Gooijer *et al.* (1985). When short time-series are used as input, it is hard to distinguish between closely related models (based, e.g. on AIC, BIC) since selection indices tend to be very close to each other. A change in, for instance, the length of the input data may result in a different model choice, and consequently in the forecast. See a detailed discussion in Zou and Yang (2004).

A completely different approach to the use of rigid analytical functions will be to posit (either the Beverton–Holt or Ricker type) dynamics behind the SRR relationship, based on some *a priori* knowledge (e.g. ecological/biological), then allows the data to "speak for itself". To illustrate, suppose we introduce a parameter transformation $u = R/S$ (for $S > 0$) into the Beverton–Holt and Ricker functions (this also holds for the Deriso function for which the Ricker and Beverton–Holt models are special cases), we obtain

$$u(S) = \begin{cases} e^{\alpha - \beta S}, & \text{Ricker,} \\ \dfrac{\alpha}{1 + \beta S}, & \text{Beverton-Holt.} \end{cases} \qquad (16)$$

We can then prove (maintaining the original constraints on α and β) that u is both monotone and convex. An immediate consequence of this transformation is that one can write

$$u(S) = \sum_{i=1}^{N} c_j B_j(S),$$

where $B_j(S)$ represents a set of basis functions with some local support (e.g. low degree splines) and c_j are a set of parameters to be determined. It is then possible to solve for u by imposing constraints on c_j such that the derived solution has the required properties. Once u is obtained, R can be recovered. Bravington *et al.* (2000) and later, Cadigan (2013) who fitted a non-parametric stock–recruitment model using the R-SCAM package (Pya, 2010, 2013), are based on such an approach. This approach (reparametrization using B-spline basis functions) has the attraction that the spline knots (design points) can be chosen independent of the observation. There are two potential challenges. The choice of an arbitrary high number N of basis functions may result in oscillatory (non-monotonic), and biologically implausible solutions. In the R-SCAM package, regularization of the problem is addressed by adding a term in the objective function with a smoothing parameter, which penalizes divergence from smoothness. The actual choice of the smoothing parameter, which can be determined by, for example, generalized cross-validation (GCV), is usually data dependent (see, e.g. Wahba, 1985). More generally, the choice of penalizing term and how it is determined will dictate the type of solution obtained. For instance, methods such as GCV and unbiased risk have been developed under the assumption that the data are from independent observations (Wang, 1998). When the independent observation assumption is violated, the results obtained are underestimates of the optimal smoothing parameter. If this approach (reparametrization using basis functions) is to be truly generic, the following must be considered (i) since the fitting problem involves more than data interpolation and smoothing, it must admit other basis functions which possess the

same flexibility and continuity properties as B-spline, e.g. Hermite or Bernstein polynomial (see, e.g. El Attar, 2006); (ii) rather than being explicitly imposed on the solution, the degree of smoothness of the SR function must be an emergent property of the derived function; and (iii) given the nature of the SR data, the parameter estimation must be robust to outliers and computationally feasible, even for a very small number of observation. Following the discussion from the previous paragraph, non-parametric recruitment models of the type described in this paragraph (despite their attractiveness) must still be regarded as candidates in an ensemble of several possible explanatory models (both parametric and non-parametric).

Forecasting stock–recruitment

The problem of forecasting stock–recruitment shares several common characteristics with that of finding a functional fit to, for example, recruit-spawner data. The shared characteristics include (but are not limited to) decision on model choice (type and degree of complexity), issues with bias-variance trade-off, and effect of underlying statistical assumptions on estimated model parameters. However, recruitment forecasting presents an additional challenge since it seeks to address the problem of making statements about the most likely outcome of future values of a process (or time-series variable) whose actual outcomes are unknown. The outstanding problems with recruitment forecasting are therefore many, and deal with issues concerning the process, accuracy, and relevance of forecasts.

Why has the SST-based forecast of California sardine recruitment (Jacobson and Mac-Call, 1995) been successful, while the potential of a similar SST-based recruitment forecasting of, for example, North Sea cod (Planque et al., 2003), been considered ill-advised? This is because the assumed SST–recruitment relationship for the California sardine turned out to be a (non-linearly) causal relationship (Sugihara et al., 2012). Since recruitment is forced by several drivers (including temperature), it is conceivable that not all factors will have direct (linear/non-linear) causal link to recruitment, and that one or several links may be transitive. The correlation between SST and recruitment may be high (even in the absence of a direct causal link) when temperature represents the most viable proxy for the group of recruitment driving forces or the common denominator for a series of transitive relationships. On the other hand, correlations between covariates could be weak although a causal relationship exists. For instance, the observed correlations in the relationship between the 3-year running averages of the Scripps Pier SST–sardine recruitment and spawning stock size by Jacobson and MacCall (1995) appeared to vanish, when the analysis was extended to include assessment results from 1992 to 2009 (see McClatchie et al., 2010; Sugihara et al., 2012). Another good illustration is provided by an example model by Stiansen et al. (2005), which provides a two year-ahead forecasts for 3-year-old recruits of Northeast Arctic cod,

$$R_t \sim (\text{Water Temp.})_{t-3} + (\text{Age} - 1 \text{ cod})_{t-2}$$
$$+ \log(\text{Biomass of Matured Capelin})_{t-2}. \qquad (17)$$

Retrospective analysis (from 1984 to 2004) showed that the model could explain ~85% of the recruitment variation. A recent (2012–2014) re-examinations show that the model forecasts for the past 5 years account for <50% of the annual recruitment variability (see ICES, 2012). The model's poor performance can be linked to the inability of the environmental indices (age-1 cod,

biomass of matured capelin, and perhaps water temperature) to capture components of the state-dependent dynamics of the ecosystem that appears to force the stock–recruitment.

In practice, one is often faced with an array of environmental indices (or indicators) as candidates covariates for the forecast model. Two outstanding challenges in developing recruitment forecast models then emerge. First, relevant causal indices must be identified among candidate environmental covariates using robust methodologies other than correlations. For fisheries data, this challenge has an additional component because detecting causalities may call for longer time series than is usually available. Second, models must be formulated such that the environmental variables are state-dependent rather than fixed indices (see Sugihara et al., 2012). This second consideration alludes to the development of models that incorporate environmental indices in a state-space framework. The state-space framework is particularly attractive for recruitment forecasting because it separates parameter (and observation) uncertainty from demographic variability. The variance of unexplained demographic variation is necessary to accurately represent the total stock–recruitment uncertainty. In general, parametric uncertainty can be reduced by improving the information content (or accuracy) of the data used in estimating model parameters. Unfortunately, uncertainty in projections due to demographic uncertainty generally cannot be reduced, particularly in medium and long-term projections.

The existence of multiple environmental indicators of recruitment, however, implies the possible development of several candidate forecast models on subsets of the pool of data. For example, about eight different regression models currently exist for forecasting age-3 recruits of Northeast Arctic cod. The models take as input different combinations of data from a time-series pool consisting of oxygen saturation at bottom layers of the Kola section transect (stations 3–7), the air temperature at the Murmansk (Russia) station, water temperature (stations 3–7) of the Kola section (layers 0–200 m), ice coverage in the Barents Sea, abundance indices from acoustic and bottom-trawl surveys, and biomass of fish. In the absence of an absolute truth for future estimates of recruitment, each of the model forecasts constitutes a plausible truth. Further, when candidate models use different datasets or different combinations of such in the modelling process, choosing a *best* candidate model becomes an even more challenging task especially when the variable to be forecasted is inherently highly uncertain. Combining individual model forecasts as introduced by Bates and Granger (1969) is often considered as a successful alternative to using just an individual best model. Furthermore, there is theoretically proven advantage of a proper combining over any selection method (see Yang, 2004). Specifically for time-series forecasting, predictive performance increases (Makridakis and Winkler, 1983; Armstrong, 1989; Clemen, 1989). A viable alternative to model selection has been to admit a plurality of candidate models (Ricker, Beverton–Holt, etc.). In Jiao et al. (2009), a representative value for recruitment was found by Bayesian Model Averaging (Hoeting et al., 1999) applied to an ensemble of model solutions. It must be kept in mind that the variance across a number of models is related to the risk of selecting among these models (see Vapnik, 2000; Bousquet and Elisseeff, 2002). Hence, for any measure of variance, the goal of combining individual forecasts will be to reduce the variance of the performance across the combinations relative to the variance across the individual methods (Breiman, 1996; Evgeniou et al., 2004).

Even if the computational challenges and complexities involved in developing good recruitment forecast models are addressed, the

fact remains that the models forecasts are inherently uncertain. This begs the ultimate question: "what is an acceptable degree of uncertainty in a recruitment forecast?" Perhaps the answer lies in the goal of the recruitment forecast. From a managerial point of view, this question translates into how accurate environmental-dependent forecast models have to be to be considered as management decision tools (De Oliveira and Butterworth, 2005). In this context, an acceptable degree of uncertainty translates into a tolerable management decision risk in setting TACs and other harvest rules. In a controlled numerical experiment, De Oliveira and Butterworth (2005) concluded that an environmental index need to explain at least 50% of the total variation in recruitment before the management procedure starts showing benefits, measured in terms of the summary statistics for risk and average catch. There is, however, room for further research in this area, when model forecasts are based on short-time-series data. In particular, there is a requirement for methodologies for predicting, quantifying uncertainty and risk associated with rare events (recruitment spikes).

Concluding remarks

As stated by Hilborn and Walters (1992), "Analysis of stock–recruitment data provides an enormous number of traps for the unwary—good luck". This statement becomes even more poignant when attempting to include the effects of environmental variables in recruitment modelling and prediction. While theoretical reasons (Bradford, 1992; Mertz and Myers, 1995) establish doubt over whether recruitment will ever be predicted with the degree of accuracy needed for management purposes, incremental progress has been made in the past 100 years; techniques are now available (e.g. linked biophysical IBMs, conceptual modelling, more robust analysis of empirical information) that provide a firm groundwork for further investigations. For instance, within the domain of coupled biophysical IBMs, three broad categories of hypotheses have been identified to explain and predict recruitment variability in fish populations (Miller, 2007). Studies show that recruitment variability may be related to food and growth, transport, or predation, and most of the hypotheses (see, e.g. Cowan and Shaw, 2002; Govoni, 2005) clearly point to coupling between physical and biological processes as key to understanding recruitment variability. Ecological concepts continue to be translated with success into methodologies and models, which when pooled together, are capable of explaining and interpreting different aspects of recruitment variability (see, e.g. Minto, 2011). Several statistical tools have evolved (e.g. randomization tests) that provide robust analysis of empirical information. Such tests can be used to check the statistical characteristics of data, such as independence of observations, or to screen goodness-of-fit tests (Stephens, 1974). Power (1996), for instance, applied randomization and goodness-of-fit tests to recruitment dataseries for several hundreds of North Atlantic fish stocks to establish the suitability of the exponential, lognormal, and Weibull distributions as representing appropriate descriptions of the data.

The central question remains: "will stock–recruitment modelling and forecasting be recognized as being relevant to fisheries science and management in the next one hundred years?" While believing this is possible, there are prerequisites for success. The purpose of recruitment models needs to be identified and clarified before development, i.e. the models must be developed and tuned to specific goals and objectives. This calls for two distinct lines of recruitment modelling, dictated by whether the goal is to address policy-relevant issues or to understand cause–effect mechanisms.

Management decision models for instance, must be required to satisfy some decision calculus, i.e. a set of numerical procedures for data processing and judgement, to qualify as decision tools. These model requirements include (see, e.g. Little, 2004) (i) completeness and parsimony—incorporating only the most important drivers influencing decision parameters; (ii) robustness—providing plausible, non-absurd results; (iii) controllability—knowledge of type of input data required to produce a desired output; (iv) adaptiveness—seamless incorporation and accommodation of new information with the goal to improve model performance; and (v) possession of a simple interface for evaluation of risk and decision effects, e.g. of how a change in model parameters may impact management decision. While parsimony remains a guiding modelling principle, models built purposely to understand recruitment mechanisms may need to be adequately complex, i.e. incorporate community dynamics and environmental factors, to fully integrate recruitment drivers. Such process models must also be able to produce a larger spectrum of results other than observations, which only constitute a snapshot of natural variability. Thus all model scenarios that do not violate first biological/physical principles may be considered as plausible (contrast with (ii)–(iii) for management decision models). Finally, for process models, risk and decision effects are secondary to the primary quest to understand the different configurations of driver alignments and how these may result in different observation states. In the recognition and pursuance of this dichotomy lies the path to relevance and success for stock–recruitment modelling and forecasting.

Do we need the whole life cycle for assessment purposes? The answer may be "probably not" for applied purposes. Keeping the analyses as simple as possible while understanding the limitations may be good enough to assessment and management. Recruitment models for such purposes (e.g. determining reference points or MSE) could be formulated, for instance, in terms of management parameters (e.g. maximum sustainable catch and harvest rate) rather than productivity parameters, making them more relevant to policy (see, e.g. Schnute and Kronlund, 1996). On a cautionary note, the user should be extremely aware of how the recruitment data were generated (see Dickey-Collas et al., 2014). Often recruitment time series are outputs from standard stock assessments and are therefore the consequence of the assumptions in the assessment model. Further, since the outputs from different models give different perceptions, quantifying uncertainty in model predictions must be an integral component of the modelling exercise. Shorter time series are available from survey indices, which, while they are free from the assumptions in the assessment model, are influenced by assumptions in survey design and strategy.

Apart from providing support to policy decisions, recruitment modelling could be the result of a quest to understand cause–effect mechanisms underpinning stock variability, and this may require considering life histories and the drivers of early life-history dynamics, including parental effects, as well as endogenous and exogenous influences on survivorship through to sexual maturity. For this purpose, a life-stage model may be essential; here, success is required in further development of predictive modelling tools to fully link recruitment variability to perturbations in biotic and abiotic conditions. This idea is in itself not new (see, e.g. Myers and Drinkwater, 1989; Myers, 1998), however, in this particular context, the term "predictive modelling" is used in a broader sense to describe two major tasks, namely (i) the ability to develop models based on established causal links in the mechanisms (biotic and abiotic, across populations, and on different time and

data resolutions) during different life stanzas of fish growth and (ii) to demonstrate that these mechanisms determine the dynamics of recruitment, based on more robust methodologies other than correlations (Myers, 1998). Many have even gone as far as questioning the utility of funding research in recruitment, which incorporates environmental indices (see Myers, 1998; Walters and Collie, 1988). The crux of the problem, however, lies in the statement by Myers (1998): "The emphasis on the search for environmental *correlations* of recruitment may have led to the neglect of other important processes" (emphasis on correlations by us). Further, although recruitment forcing may be better understood when studied across populations (Myers and Barrowman, 1995; Myers, 1998), the basic challenge remains: the information available for most exploited stocks is inadequate for precise modelling, even for those species which have long-time series (e.g. North Sea or Barents Sea gadoid stocks). It will therefore be essential to harness efforts being made in various fronts, including linked biophysical IBMs, conceptual modelling, and field experiments to understand mechanisms regulating interannual survival at the egg and larval stages.

There is no doubt that, in an ever changing climate, recruitment modelling and forecasting will remain central to fisheries science in the next 100 years. Recruitment models will be required to develop and evaluate harvest control rules in management plans, set precautionary and MSY reference points, and predict likely changes to stock productivity caused by changes to habitat, regime, or local climate. While the task involved is far from simple, developments in recent years (in field experiments and conceptual modelling) give grounds for optimism.

References

Allee, W. C., Emerson, A. E., Park, O., Park, T., and Schmidt, K. P. 1949. Principles of Animal Ecology. Saunders, Philadelphia.

Allen, K. R. 1973. The influence of random fluctuations in the stock-recruitment relationship on the economic return from salmon fisheries. Rapports et Proces-Verbaux des Reunions, Conseil International Pour L'Exploration de la Mer, 164: 350–359.

Anderson, C. N., Hsieh, C.-h., Sandin, S. A., Hewitt, R., Hollowed, A., Beddington, J., May, R. M., *et al.* 2008. Why fishing magnifies fluctuations in fish abundance. Nature, 452: 835–839.

Andrews, D. W. 1993. Exactly median-unbiased estimation of first order autoregressive/unit root models. Econometrica, 61: 139–165.

Armstrong, J. S. 1989. Combining forecasts: the end of the beginning or the beginning of the end? International Journal of Forecasting, 5: 585–588.

Bakun, A. 1998. Ocean triads and radical interdecadal stock variability: bane and boon for fishery management science. *In* Reinventing Fisheries Management, pp. 331–358. Ed. by T. J. Pitcher, P. J. B. Hart, and D. Pauly. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bakun, A., and Weeks, S. J. 2006. Adverse feedback sequences in exploited marine systems: are deliberate interruptive actions warranted? Fish and Fisheries, 7: 316–333.

Barrowman, N. J., and Myers, R. A. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. Canadian Journal of Fisheries and Aquatic Sciences, 57: 665–676.

Basson, M. 1999. The importance of environmental factors in the design of management procedures. ICES Journal of Marine Science, 56: 933–942.

Bates, J. M., and Granger, C. W. J. 1969. The combination of forecasts. Operational Research, 20: 451–468.

Beaugrand, G., Brander, K. M., Lindley, J. A., Souissi, S., and Reid, P. C. 2003. Plankton effect on cod recruitment in the North Sea. Nature, 426: 661–664.

Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. Fishery Investigation Series II, Vol. XIX. UK. 533 pp.

Beverton, R. J. H., and Iles, T. C. 1992a. Mortality rates of 0-group plaice (*Platessa platessa* L.), dab (*Limanda limanda* L.) and turbot (*Scophthalmus maximus* L.) in European waters: III. Density-dependence of mortality rates of 0-group plaice and some demographic implications. Netherlands Journal of Sea Research, 29: 61–79.

Beverton, R. J. H., and Iles, T. C. 1992b. Mortality rates of 0-group plaice (*Pleuronectes platessa* L.) dab (*Limanda limanda* L.) and turbot (*Scophthalmus maximus* L.) in Euro pean waters: II. Comparison of mortality rates and construction of life table for 0-group plaice. Netherlands Journal of Sea Research, 29: 49–59.

Birnbaum, A. 1964. Median-unbiased estimators. Bulletin of Mathematical Statistics, 11: 25–34.

Bochkov, Y. A. 1982. Water temperature in the 0–200 m layer in the Kola–Meridian in the Barents Sea, 1990–1981. Sb. Nauchn. Trud. PINRO, Murmansk (In Russian), 46: 113–122.

Bogstad, B., Røttingen, I., Sandberg, P., and Tjelmeland, S. 2000. The use of Medium-Term Forecasts in advice and management decisions for the stock of Norwegian spring spawning herring (*Clupea harengus* L.). ICES CM 2000/V:01. ICES.

Borja, A., Uriarte, A., Egaña, J., Motos, L., and Valencia, V. 1998. Relationships between anchovy (*Engraulis encrasicolus*) recruitment and environment in the Bay of Biscay (1967–1996). Fisheries Oceanography, 7: 375–380.

Bousquet, O., and Elisseeff, A. 2002. Stability and generalization. The Journal of Machine Learning Research, 2: 499–526.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. 1994. Time Series Analysis: Forecasting and Control. Prentice Hall, New Jersey.

Bradford, M. J. 1992. Precision of recruitment predictions from early life stages of marine fishes. Fishery Bulletin, 90: 439–453.

Bravington, M. V., Stokes, T. K., and O'Brien, C. M. 2000. Sustainable recruitment: the bottom line. Marine and Freshwater Research, 51: 465–475.

Breiman, L. 1996. Bagging predictors. Machine Learning, 24: 123–140.

Brooks, E. N., and Powers, J. E. 2007. Generalized compensation in stock-recruit functions: properties and implications for management. ICES Journal of Marine Science, 64: 413–424.

Burnham, K. P., and Anderson, D. R. 2002. Model selection and multi-model inference: a practical information-theoretic approach, 2nd edn. Springer-Verlag, New York.

Cadigan, N. G. 2013. Fitting a non-parametric stock–recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty. ICES Journal of Marine Science, 70: 56–67.

Chen, D. G. 2004. Bias and bias correction in fish recruitment prediction. 969 North American Journal of Fisheries Management, 24: 724–730.

Chen, D. G., and Irvine, J. R. 2001. A semiparametric model to examine stock-recruitment relationships incorporating environmental data. Canadian Journal of Fisheries and Aquatic Sciences, 58: 1178–1186.

Chen, D. G., Irvine, J. R., and Cass, A. J. 2002. Incorporating allee effects in fish stock recruitment models and applications for determining reference points. Canadian Journal of Fisheries and Aquatic Sciences, 59: 242–249.

Chen, D. G., and Ware, D. M. 1999. A neural network model for forecasting fish stock recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2385–2396.

Clemen, R. T. 1989. Combining forecasts: A review and annotated bibliography. International Journal of Forecasting, 5: 559–583.

Cochrane, K., and Hutchings, L. 1995. A structured approach to using biological and environ mental parameters to forecast anchovy recruitment. Fisheries Oceanography, 4: 102–127.

Cochrane, K., and Starfield, A. 1992. The potential use of predictions of recruitment success in the management of the South African

anchovy resource. South African Journal of Marine Science, 12: 891–902.

Cook, R. M. 1998. A sustainability criterion for the exploitation of North Sea cod. ICES Journal of Marine Science, 55: 1061–1070.

Cornbleet, P. J., and Gochman, N. 1979. Incorrect least-squares regression coefficients in method-comparison analysis. Clinical Chemistry, 25: 432–438.

Corten, A. 2013. Recruitment depressions in North Sea herring. ICES Journal of Marine Science, 70: 1–15.

Cowan, J., and Shaw, R. 2002. Recruitment. *In* Fishery Science: The 992 unique contributions of early life stages, pp. 88–111. John Wiley & Sons.

Cury, P., and Roy, C. 1989. Optimal environmental window and pelagic fish recruitment success in upwelling areas. Canadian Journal of Fisheries and Aquatic Sciences, 46: 670–680.

Cushing, D. H. 1973. Dependence of recruitment on parent stock. Journal of the Fisheries Board of Canada, 30: 1965–1976.

Cushing, D. H. 1982. Climate and Fisheries. Academic Press, London.

Cushing, D. H. 1984. The gadoid outburst in the North Sea. Journal du Conseil International Pour L'Exploration de la Mer, 41: 159–166.

Cushing, D. H. 1990. Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. Advances in Marine Biology, 26: 249–293.

de Gooijer, J. G., Abraham, B., Gould, A., and Robinson, L. 1985. Autoregressive-moving average process: a survey. International Statistical Review, 53: 301–329.

De Oliveira, J., and Butterworth, D. 2005. Limits to the use of environmental indices to reduce risk and/or increase yield in the South African anchovy fishery. African Journal of Marine Science, 27: 191–203.

De Oliveira, J. A., Uriarte, A., and Roel, B. A. 2005. Potential improvements in the management of Bay of Biscay anchovy by incorporating environmental indices as recruitment predictors. Fisheries Research, 75: 2–14.

de Valpine, P., and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs, 72: 57–76.

Deriso, R. B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences, 37: 268–282.

Deriso, R. B., Barnes, J. T., Jacobson, L. D., and Arenas, P. R. 1996. Catch-at-age analysis for Pacific sardine (*Sardinops sagax*), 1983 1995. ColCOFl Rep. vol. 37, USA.

Dickey-Collas, M., Hintzen, N. T., Nash, R. D. M., Schön, P., and Payne, M. R. 2014. Quirky patterns in time-series of estimates of recruitment could be artefacts. ICES Journal of Marine Science, doi:10.1093/icesjms/fsu022

Dingsør, G. E., Ciannelli, L., Chan, K.-S., Ottersen, G., and Stenseth, N. C. 2007. Density dependence and density independence during the early life stages of four marine fish stocks. Ecology, 88: 625–634.

Drinkwater, K. F. 2005. The response of Atlantic cod (*Gadus morhua*) to future climate change. ICES Journal of Marine Science, 62: 1327–1337.

Duplisea, D. E., and Robert, D. 2008. Prerecruit survival and recruitment of northern Gulf of St Lawrence Atlantic cod. ICES Journal of Marine Science, 65: 946–952.

El Attar, R. 2006. Special Functions and Orthogonal Polynomials. Mathematical Series Volume 3. Lulu Press, Morrisville, NC, USA.

Evans, G. T., and Rice, J. C. 1988. Predicting recruitment from stock size without the mediation of a functional relation. Journal du Conseil International pour l'Exploration de la Mer, 44: 111–122.

Evgeniou, T., Pontil, M., and Elisseeff, A. 2004. Leave one out error, stability, and generalization of voting combinations of classifiers. Machine Learning, 55: 71–97.

Fitzhugh, G. R., Shertzer, K. W., Kellison, G. T., and Wyanski, D. M. 2012. Review of size and age-dependence in batch spawning: implications for stock assessment of fish species exhibiting indeterminate fecundity. Fishery Bulletin, 110: 413–425.

Frank, K. T., and Brickman, D. 2000. Allee effects and compensatory population dynamics within a stock complex. Canadian Journal of Fisheries and Aquatic Sciences, 57: 513–517.

Gjøsæter, H., and Bogstad, B. 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.

Govoni, J. J. 2005. Fisheries oceanography and the ecology of early life histories of fishes: a perspective over fifty years. Scientia Marina, 69: 125–137.

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.

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.

Hammond, T. R., and O'Brien, C. M. 2001. An application of the Bayesian approach to stock assessment model uncertainty. ICES Journal of Marine Science, 58: 648–656.

Hamre, J. 2003. Capelin and herring as key species for the yield of North-east Arctic cod. Results from multispecies model runs. Scientia Marina, 67: 315–323.

Hennemuth, R. C., Palmer, J. E., and Brown, B. E. 1980. A statistical description of recruitment in eighteen selected fish stocks. Journal of Northwest Atlantic Fishery Science, 1: 101–111.

Hilborn, R. 1985. Simplified calculation of optimum spawning stock size from Ricker's stock recruitment curve. Canadian Journal of Fisheries and Aquatic Sciences, 42: 1833–1834.

Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Reviews in Fish Biology and Fisheries, 2: 177–186.

Hjermann, D. Ø., Bogstad, B., Eikeset, A. M., Ottersen, G., Gjøsæter, H., and Stenseth, N. C. 2007. Food web dynamics affect Northeast Arctic cod recruitment. Proceedings of the Royal Society B: Biological Sciences, 274: 661–669.

Hjort, J. 1914. Fluctuations in the great fisheries of Northern Europe viewed in the lights of biological research. Rapport et Proces-Verbaux des Réunion Conseil International pour l'Exploration de la Mer, 20: 1–228.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. 1999. Bayesian model averaging: tutorial with discussion. Statistical Science, 14: 382–417.

Hsieh, C.-h., Reiss, C. S., Hunter, J. R., Beddington, J. R., May, R. M., and Sugihara, G. 2006. Fishing elevates variability in the abundance of exploited species. Nature, 443: 859–862.

Hutchinson, W. F. 2008. The dangers of ignoring stock complexity in fishery management: the case of the North Sea cod. Biology Letters, 4: 693–695.

ICES. 2003. Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. ICES CM 2003/ACFM:07, ICES, Copenhagen, Denmark.

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

ICES. 2011b. Report of the Working Group for Celtic Seas Ecoregion (WGCSE). ICES CM 2011/ACOM:12, ICES, Copenhagen, Denmark.

ICES. 2012. Report of the Study Group on Recruitment Forecasting 1080 (SGRF). ICES CM 2012/ACOM:24, ICES, Barcelona, Spain.

ICES. 2013a. Report of the Arctic Fisheries Working Group (AFWG). ICES CM 2013/ACOM:05, ICES, Copenhagen, Denmark.

ICES. 2013b. Report of the Herring Assessment Working Group for the Area South of 62 N (HAWG). ICES CM 2013/ACOM:06, ICES, Copenhagen, Denmark.

Iles, T. C. 1994. A review of stock-recruitment relationships with reference to flatfish populations. Netherlands Journal of Sea Research, 32: 399–420.

Iles, T. C., and Beverton, R. J. H. 1998. Stock, recruitment and moderating processes in flatfish. Journal of Sea Research, 39: 41–45.

Jacobson, L. D., and MacCall, A. D. 1995. Stock-recruitment models for Pacific sardine (*Sardinops sagax*). Canadian Journal of Fisheries and Aquatic Sciences, 52: 566–577.

Jiao, Y., Reid, K., and Smith, E. 2009. Model selection uncertainty and Bayesian model averaging in fisheries recruitment modeling. *In* The Future of Fisheries Science in North America, chap. 26, pp. 505–524. Springer.

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

Kimoto, A., Mouri, T., and Matsuishi, T. 2007. Modelling stock–recruitment relationships to examine stock management policies. ICES Journal of Marine Science, 64: 870–877.

Kjesbu, O. S., Solemdal, P., Bratland, P., and Fonn, M. 1996. Variation in annual egg production in individual captive Atlantic cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences, 53: 610–620.

Lasker, R. 1981. The role of a stable ocean in larval fish survival and subsequent recruit ment. *In* Marine fish larvae, Washington Sea Grant Program, pp. 80–87. Ed. by R. Lasker. University of Washington Press, Seattle, USA.

Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? Fisheries Research, 125: 254–261.

Liermann, M., and Hilborn, R. 1997. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. Canadian Journal of Fisheries and Aquatic Sciences, 54: 1976–1984.

Linnet, K. 1993. Evaluation of regression procedures for methods comparison studies. Clinical Chemistry-Washington, 39: 424–424.

Little, J. D. 2004. Models and managers: the concept of a decision calculus. Management Science, 50: 1841–1853.

Ludwig, D., and Walters, C. J. 1981a. Measurement errors and uncertainty in parameter estimates for stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 38: 711–720.

Ludwig, D., and Walters, C. J. 1981b. Optimal harvesting with imprecise parameter estimates. Ecological Modelling, 14: 273–292.

Lungarella, M., Ishiguro, K., Kuniyoshi, Y., and Otsu, N. 2007. Methods for quantifying the causal structure of bivariate time series. International Journal of Bifurcation and Chaos, 17: 903–921.

Mace, P. M., Doonan, I., and Fisheries, M. 1988. A Generalised Bioeconomic Simulation Model for Fish Population Dynamics. MAFFish, NZ Ministry of Agriculture and Fisheries, Wellington.

Mace, P. M., and Sissenwine, M. P. 2002. Coping with uncertainty: evolution of the relationship between science and management. *In* Incorporating Uncertainty into Fishery Models, pp. 9–28. Ed. by J. M. Berkson, L. L. Kline, and D. J. Orth. American Fisheries Society, Symposium 27, Bethesda, MD.

Makridakis, S., and Winkler, R. L. 1983. Averages of forecasts: some empirical results. Management Science, 29: 987–996.

Marshall, D. J., Heppell, S. S., Munch, S. B., and Warner, R. R. 2010. The relationship between maternal phenotype and offspring quality: do older mothers really produce the best offspring? Ecology, 91: 2862–2873.

Martin, R. F. 2000. General deming regression for estimating systematic bias and its confidence interval in method-comparison studies. Clinical Chemistry, 46: 100–104.

Maunder, M. N., and Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1204–1216.

Maunder, M. N., and Starr, P. J. 2001. Bayesian assessment of the SNA1 snapper (*Pagrus auratus*) stock on the northeast coast of New Zealand. New Zealand Journal of Marine and Freshwater Research, 35: 87–110.

Maunder, M. N., and Watters, G. M. 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and examples. Fishery Bulletin, 101: 89–99.

McBride, R. S., Somarakis, S., Fitzhugh, G. R., Albert, A., Yaragina, N. A., Wuenschel, M. J., Alonso-Fernández, A., et al. 2013. Energy acquisition and allocation to egg production in relation to fish reproductive strategies. Fish and Fisheries, doi:10.1111/faf.12043

McClatchie, S., Goericke, R., Auad, G., and Hill, K. 2010. Re-assessment 1147 of the stock-recruit and temperature-recruit relationships for Pacific sardine (*Sardinops sagax*). Canadian Journal of Fisheries and Aquatic Sciences, 67: 1782–1790.

Mertz, G., and Myers, R. A. 1995. Estimating the predictability of recruitment. Fishery Bulletin, 93: 657–665.

Methot, R. D., and Taylor, I. G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1744–1760.

Meyer, R., and Millar, R. B. 1999. Bayesian stock assessment using a state–space implementation of the delay difference model. Canadian Journal of Fisheries and Aquatic Sciences, 56: 37–52.

Miller, T. J. 2007. Contribution of individual-based coupled physical-biological models to understanding recruitment in marine fish populations. Marine Ecology Progress Series, 347: 127–138.

Minto, C. 2011. Ecological inference from variable recruitment data. PhD thesis, Dalhousie University, Halifax, NS.

Morgan, M. J., Perez-Rodriguez, A., Saborido-Rey, F., and Marshall, C. T. 2011. Does increased information about reproductive potential result in better prediction of recruitment? Canadian Journal of Fisheries and Aquatic Sciences, 68: 1361–1368.

Morgan, M. J., Wright, P. J., and Rideout, R. M. 2013. Effect of age and temperature on spawning time in two gadoid species. Fisheries Research, 138: 42–51.

Mukhina, N. V., Marshall, C. T., and Yaragina, N. A. 2003. Tracking the signal in year-class strength of Northeast Arctic cod through multiple survey estimates of egg, larval and juvenile abundance. Journal of Sea Research, 50: 57–75.

Munch, S. B., Kottas, A., and Mangel, M. 2005. Bayesian nonparametric analysis of stock recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences, 62: 1808–1821.

Myers, R. A. 1998. When do environment–recruitment correlations work? Reviews in Fish Biology and Fisheries, 8: 285–305.

Myers, R. A., and Barrowman, N. 1995. Time series bias in the estimation of density dependent mortality in stock-recruitment models. Canadian Journal of Fisheries and Aquatic Sciences, 52: 223–232.

Myers, R. A., and Drinkwater, K. 1989. The influence of Gulf Stream warm core rings on recruitment of fish in the northwest Atlantic. Journal of Marine Research, 47: 635–656.

Myers, R. A., Rosenberg, A. A., Mace, P. M., Barrowman, N., and Restrepo, V. R. 1994. In search of thresholds for recruitment overfishing. ICES Journal of Marine Science, 51: 191–205.

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

Nash, R. D. M., and Dickey-Collas, M. 2005. The influence of life history dynamics and environment on the determination of year class strength in North Sea herring (*Clupea harengus* L.). Fisheries Oceanography, 14: 279–291.

Nash, R. D. M., Dickey-Collas, M., and Kell, L. T. 2009. Stock and recruitment in North Sea herring (*Clupea harengus*); compensation and depensation in the population dynamics. Fisheries Research, 95: 88–97.

Nash, R. D. M., and Geffen, A. J. 2012. Mortality through the early life history of fish: What can we learn from European plaice (*Pleuronectes platessa* L.)? Journal of Marine Systems, 93: 58–68.

Needle, C. L. 2001. Recruitment models: diagnosis and prognosis. Reviews in Fish Biology and Fisheries, 11: 95–111.

Nelson, C. R., and Plosser, C. R. 1982. Trends and random walks in macroeconmic time series: some evidence and implications. Journal of Monetary Economics, 10: 139–162.

O'Brien, C. M., Fox, C. J., Planque, B., and Casey, J. 2000. Fisheries: climate variability and North Sea cod. Nature, 404: 142–142.

Olsen, E. M., Ottersen, G., Llope, M., Chan, K.-S., Beaugrand, G., and Stenseth, N. C. 2011. Spawning stock and recruitment in North Sea cod shaped by food and climate. Proceedings of the Royal Society B: Biological Sciences, 278: 504–510.

Ottersen, G., Hjermann, D., and Stenseth, N. C. 2006. Changes in spawning stock structure strengthen the link between climate and recruitment in a heavily fished cod (*Gadus morhua*) stock. Fisheries Oceanography, 15: 230–243.

Ottersen, G., Kim, S., Huse, G., Polovina, J. J., and Stenseth, N. C. 2010. Major pathways by which climate may force marine fish populations. Journal of Marine Systems, 79: 343–360.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., *et al.* 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, 2: 125–157.

Paulik, G. J. 1973. Studies of the possible form of the stock-recruitment curve. Rapports et Proces-Verbaux des Réunions du Conseil International pour l'Exploration 1214 de la Mer, 164: 302–315.

Payne, M. R., Hatfield, E. M. C., Dickey-Collas, M., Falkenhaug, T., Gallego, A., Gröger, J., Licandro, P., *et al.* 2009. Recruitment in a changing environment: the 2000s North Sea herring recruitment failure. ICES Journal of Marine Science, 66: 272–277.

Peck, M. A., and Hufnagl, M. 2012. Can IBMs tell us why most larvae die in the sea? Model sensitivities and scenarios reveal research needs. Journal of Marine Systems, 93: 77–93.

Peterman, R. M. 1978. Testing for density-dependent marine survival in Pacific salmonids. Journal of the Fisheries Research Board of Canada, 35: 1434–1450.

Peterman, R. M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Canadian Journal of Fisheries and Aquatic Sciences, 38: 1113–1119.

Planque, B., Fox, C. J., Saunders, M. A., and Rockett, P. 2003. On the prediction of short term changes in the recruitment of North Sea cod (*Gadus morhua*) using statistical temperature forecasts. Scientia Marina, 67: 211–218.

Planque, B., and Frédou, T. 1999. Temperature and the recruitment of Atlantic cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences, 56: 2069–2077.

Power, M. 1996. The testing and selection of recruitment distributions for North Atlantic fish stocks. Fisheries Research, 25: 77–95.

Prager, M. H., and MacCall, A. D. 1993. Detection of contaminant and climate effects on spawning success of three pelagic fish stocks of southern California: northern anchovy *Engraulis mordax*, Pacific sardine *Sardinops sagax*, and chub Mackerel *Scomber japonicus*. Fishery Bulletin, 91: 310–327.

Pya, N. 2010. Additive models with shape constraints. PhD thesis, University of Bath.

Pya, N. 2013. Package scam@. http://cran.r-project.org/ (last accessed December 2013).

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

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11: 559–623.

Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada, 191: 382.

Rideout, R. M., and Tomkiewicz, J. 2011. Skipped spawning in fishes: more common than you might think. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science, 3: 176–189.

Roel, B. A., O'Brien, C. M., and Basson, M. 2004. Management options for the Blackwater herring, a local spring-spawning stock in the Thames Estuary. ICES Journal of Marine Science, 61: 297–307.

Rothschild, B. J. 1986. Dynamics of Marine Fish Populations. Harvard University Press, USA.

Rothschild, B. J. 2000. Fish stocks and recruitment: the past thirty years. ICES Journal of Marine Science, 57: 191–201.

Rothschild, B. J., and Fogarty, M. 1989. Spawning-stock biomass: a source of error in recruitment/stock relationships and management advice. ICES Journal of Marine Science, 45: 131–135.

Saila, S. B., Recksiek, C. W., and Prager, M. H. (Eds) 1988. Basic fishery science programs. *In* A Compendium of Microcomputer Programs and Manual of Operations. Developments in Aquaculture and Fisheries Science, 18. Elsevier, Amsterdam, 230 pp.

Sakuramoto, K. 2005. Does the Ricker or Beverton and Holt type of stock-recruitment relationship truly exist? Fisheries Science, 71: 577–592.

Schaefer, M. B. 1991. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bulletin of Mathematical Biology, 53: 253–279.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences, 42: 414–429.

Schnute, J. T., and Kronlund, A. 1996. A management oriented approach to stock recruitment analysis. Canadian Journal of Fisheries and Aquatic Sciences, 53: 1281–1293.

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: 67–75.

Shepherd, J. G., Pope, J. G., and Cousens, R. D. 1984. Variations in fish stocks and hypotheses concerning their links with climate. Rapports et Procés-verbaux des Réunions du Conseil International pour l'Exploration de la Mer, 185: 255–267.

Shimoyama, S., Sakuramoto, K., and Suzuki, N. 2007. Proposal for stock-recruitment relationship for Japanese sardine *Sardinops melanostictus* in North-western Pacific. Fisheries Science, 73: 1035–1041.

Skjæraasen, J. E., Nash, R. D. M., Korsbrekke, K., Fonn, M., Nilsen, T., Kennedy, J., Nedreaas, K. H., *et al.* 2012. Frequent skipped spawning in the worlds largest cod population. Proceedings of the National Academy of Sciences of the United States of America, 109: 8995–8999.

Stephens, M. A. 1974. EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association, 69: 730–737.

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

Stige, L. C., Hunsicker, M. E., Bailey, K. M., Yaragina, N. A., and Hunt, G. L., Jr 2013. Predicting fish recruitment from juvenile abundance and environmental indices. Marine Ecology Progress Series, 480: 245–261.

Stocker, M., Haist, V., and Fournier, D. 1985. Environmental variation and recruitment of Pacific herring (*Clupea harenugs pallasi*) in the Strait of Georgia. Canadian Journal of Fisheries and Aquatic Sciences, 42(Suppl. 1): 174–180.

Sugihara, G., May, R., Ye, H., Hsieh, C-H., Deyle, E., Fogarty, M., and Munch, S. 2012. Detecting causality in complex ecosystems. Science, 338: 496–500.

Sun, L., Xiao, H., Li, S., and Yang, D. 2009. Forecasting fish stock recruitment and planning optimal harvesting strategies by using neural network. Journal of Computers, 4: 1075–1082.

Swain, D. P., and Sinclair, A. F. 2000. Pelagic fishes and the cod recruitment dilemma in the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences, 57: 1321–1325.

Trippel, E. A. 1999. Estimation of stock reproductive potential: history and challenges for Canadian Atlantic gadoid stock assessments. Journal of Northwest Atlantic Fishery Science, 25: 61–82.

Ulltang, Ø. 1996. Stock assessment and biological knowledge: can prediction uncertainty be reduced? ICES Journal of Marine Science, 53: 659–675.

Vapnik, V. 2000. The Nature of Statistical Learning Theory, 2nd edn. Springer, New York.

Vert-pre, K. A., Amoroso, R. O., Jensen, O. P., and Hilborn, R. 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceedings of the National Academy of Sciences of the United States of America, 110: 1779–1784.

Wahba, G. 1985. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. The Annals of Statistics, 13: 1378–1402.

Walters, C. J. 1985. Bias in the estimation of functional relationships from time series data. Canadian Journal of Fisheries and Aquatic Sciences, 42: 147–149.

Walters, C. J., and Collie, J. S. 1988. Is research on environmental factors useful to fisheries management? Canadian Journal of Fisheries and Aquatic Sciences, 45: 1848–1854.

Walters, C. J., and Hilborn, R. 1976. Adaptive control of fishing systems. Journal of the Fisheries Board of Canada, 33: 145–159.

Walters, C. J., and Ludwig, D. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences, 38: 704–710.

Wang, Y. 1998. Smoothing spline models with correlated random errors. Journal of the American Statistical Association, 93: 341–348.

Wang, Y., and Liu, Q. 2006. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock–recruitment relationships. Fisheries Research, 77: 220–225.

Williams, E. H., and Shertzer, K. W. 2003. Implications of life-history invariants for biological reference points used in fishery management. Canadian Journal of Fisheries and Aquatic Sciences, 60: 710–720.

Wootton, R. J. 1990. Introduction. *In* Ecology of Teleost Fishes, pp. 1–14. Springer.

Wright, P. J., and Trippel, E. A. 2009. Fishery-induced demographic changes in the timing of spawning: consequences for reproductive success. Fish and Fisheries, 10: 283–304.

Yang, Y. 2004. Combining forecasting procedures: some theoretical results. Econometric Theory, 20: 176–222.

Zhou, S. 2007. Discriminating alternative stock–recruitment models and evaluating uncertainty in model structure. Fisheries Research, 86: 268–279.

Zou, H., and Yang, Y. 2004. Combining time series models for forecasting. International Journal of Forecasting, 20: 69–84.

Handling editor: Mark Maunder