

Comparative analysis of statistical tools to identify recruitment–environment relationships and forecast recruitment strength

Bernard A. Megrey, Yong-Woo Lee, and S. Allen Macklin

Megrey, B. A., Lee, Y-W., and Macklin, S. A. 2005. Comparative analysis of statistical tools to identify recruitment–environment relationships and forecast recruitment strength. – ICES Journal of Marine Science, 62: 1256–1269.

Many of the factors affecting recruitment in marine populations are still poorly understood, complicating the prediction of strong year classes. Despite numerous attempts, the complexity of the problem often seems beyond the capabilities of traditional statistical analysis paradigms. This study examines the utility of four statistical procedures to identify relationships between recruitment and the environment. Because we can never really know the parameters or underlying relationships of actual data, we chose to use simulated data with known properties and different levels of measurement error to test and compare the methods, especially their ability to forecast future recruitment states. Methods examined include traditional linear regression, non-linear regression, Generalized Additive Models (GAM), and Artificial Neural Networks (ANN). Each is compared according to its ability to recover known patterns and parameters from simulated data, as well as to accurately forecast future recruitment states. We also apply the methods to published Norwegian spring-spawning herring (*Clupea harengus* L.) spawner–recruit–environment data. Results were not consistently conclusive, but in general, flexible non-parametric methods such as GAMs and ANNs performed better than parametric approaches in both parameter estimation and forecasting. Even under controlled data simulation procedures, we saw evidence of spurious correlations. Models fit to the Norwegian spring-spawning herring data show the importance of sea temperature and spawning biomass. The North Atlantic Oscillation (NAO) did not appear to be an influential factor affecting herring recruitment.

© 2005 International Council for the Exploration of the Sea. Published by Elsevier Ltd. All rights reserved.

Keywords: Artificial Neural Networks, environment–recruitment models, GAM, herring, NAO, spawner–recruit models.

Received 2 July 2004; accepted 8 May 2005.

B. A. Megrey: National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA. Y-W. Lee: Joint Institute for the Study of the Atmosphere and the Oceans, PO Box 354235, University of Washington, Seattle, WA 98195, USA. S. A. Macklin: National Oceanic and Atmospheric Administration, Pacific Marine Environmental Laboratory, 7600 Sand Point Way NE, Seattle, WA 98115, USA. Correspondence to B. A. Megrey: tel: +1 206 526 4147; fax: +1 206 526 6723; e-mail: bern.megrey@noaa.gov.

Introduction

Marine ecosystems are notoriously difficult to study. Trophic relationships are multidimensional, relevant biophysical factors vary widely in their spatial and temporal scales of influence, and process linkages are complex and highly non-linear (Wooster and Bailey, 1989; Mullin, 1993). Walters and Ludwig (1981) showed that the problem is further compounded by inaccuracies in measuring environmental variability, as well as the biotic response. Consequently, applied ecological investigations attempting to relate oceanic physics,

atmospheric physics, and marine biology to variations in fish stock–recruitment are difficult to carry out. Nonetheless, the collective impacts of regime shifts, large multi-decadal-scale forcings of marine ecosystems (such as those attributed to the NAO), and natural and man-made influences on variability in fish populations and future states of ecosystems are widely recognized as important areas of study.

One of the major outstanding questions of fisheries science concerns the existence of a relationship between the size of the breeding population (spawners), their offspring (recruits), and the influence of the surrounding environment

(Hilborn and Walters, 1992). Many of the theories on fisheries yield optimization and management policy of fish populations in marine ecosystems are based on analysis of spawner–recruit relationships. However, for most fish species, recruitment, being itself highly variable, often appears to fluctuate independently of stock size. The most common explanation for the apparent failure of spawner–recruit theory is that the relationship is obscured by extremely high mortality rates during the pre-recruit period (Goodyear and Christensen, 1984; Cushing, 1988; Walters and Collie, 1988). Presumably, population dynamics during early life stages are strongly influenced by the environmental conditions experienced by planktonic fish larvae and eggs. This counterintuitive result has led others to advocate the development of management strategies that do not assume an underlying spawner–recruit relationship (Evans and Rice, 1988; Koslow, 1992; Mackinson *et al.*, 1999; Chen and Irvine, 2001; Voges *et al.*, 2002).

The estimation and forecasting of recruitment has been a persistent preoccupation of fisheries scientists for close to a century because of the critical importance of recruitment to the exploitable segment of the stock (Needle, 2002). Linkages between recruitment and external variables are often explored with various parametric statistical techniques in the hope that the important process controlling recruitment variability will be identified, and that useable recruitment forecasting tools can be developed. Despite numerous attempts, the complexity of the problem often seems beyond the capabilities of traditional statistical analysis paradigms. There may be limitations in theoretical development, inadequate length of time-series, the need to partition already short time-series into segments representing unique production regimes, lack of degrees of freedom, or the inability to meet required assumptions. Typical parametric procedures like linear regression and correlation analysis require restrictive assumptions such as normally distributed errors, dependent and independent variables measured without error, linear relationships, and non-auto-correlated data. As biologists, we know these assumptions are not valid in marine systems. These methods also suffer from common restrictions such as: standard multiple linear models do not allow interactions between variables unless they are pre-specified and of a particular multiplicative form; functional relationships must be specified *a priori* before analysis proceeds; and analytical methods do not deal well with gaps (missing values) in the data time-series. In most circumstances, these conditions cripple the ability of analysts to make sense of the data.

Modelling paradigms and analytical tools that effectively identify environment-dependent recruitment relationships and provide the capability for recruitment prediction are not well established, despite the need to understand recruitment dynamics and the compelling economic, biological, and political arguments for prudent management. Many of the factors affecting recruitment in marine populations are still poorly understood, complicating the prediction of strong

year classes. Still, recruitment prediction remains a vital information component required by resource management decision-makers dealing with exploited marine ecosystems.

The objectives of this study are to investigate the utility of four separate statistical procedures to identify relationships between recruitment and the environment. Because we can never really know the parameters or underlying relationships of actual data, we chose to use simulated data with known properties and different levels of measurement error to test and compare the methods, especially their ability to forecast future recruitment states. We also apply the methods to published spawner–recruit–environment data for a selected North Atlantic fish stock, namely Norwegian spring-spawning herring (*Clupea harengus* L.). Methods examined include traditional linear (LR) and non-linear regressions (NLR), Generalized Additive Models (GAM), and Artificial Neural Networks (ANN). Each is compared according to its ability to recover known patterns and parameters from the simulated data, describe the observed data from a North Atlantic fish stock as well as to accurately forecast future recruitment states.

Material and methods

Data sets

Simulated data

We wanted to simulate a recruitment time-series with known properties, and then to use these data to test the performance of statistical procedures typically used to examine recruitment–environment relationships. The simulated data (R_s) should have statistical properties similar to the observed recruitment time-series (R_o), and we used the Gulf of Alaska walleye pollock recruitment time-series, estimated from the annual stock assessment model applied to this stock (Dorn *et al.*, 2003) to help guide us in this aspect of the study. Age-2 recruitment (in billions) was estimated for the 1961–2002 year classes for a total of 42 data points.

By design, simulated recruitment (R_s) was a response to three explanatory factors, each related to recruitment by a predefined functional relationship. Simulated recruitment was related to the first factor in a non-linear way and to the second in a linear way. There was no relationship between recruitment and the third factor.

Factor 1, SB. This factor, (X_1), is similar to a relationship between recruitment and spawning biomass. We used the actual spawner biomass (million tonnes) from the Gulf of Alaska walleye pollock stock assessment model (Dorn *et al.*, 2003) to generate random data with similar statistical properties. These data were skewed to the right, and we attempted to fit non-symmetrical distributions to the data. The gamma distribution best described the data using residual sum of squares criteria.

The gamma distribution, which is a family of curves based on two parameters, is given by

$$X_1 = f(x|a, b) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b} \quad (1)$$

where a and b are parameters > 0 , x is the spawning biomass data, and $\Gamma(\cdot)$ is the Gamma function. When a is large, the gamma distribution closely approximates a normal distribution with the advantage that the gamma distribution has density only for positive real numbers. The gamma distribution was fit to the spawning biomass data, and the parameters, $a = 3.0224$ and $b = 0.115$, were estimated with maximum likelihood.

The functional relationship relating Factor 1 to recruitment (R_s) was defined as being similar to a non-linear Ricker-type spawner–recruit relationship (Ricker, 1975), using the random variable X_1 as the explanatory variable. To parameterize this relationship, we fit a Ricker spawner–recruit curve to the Gulf of Alaska spawner and recruit time-series using non-linear regression. Parameter estimates were $\alpha = 4.171$ and $\beta = 1.212$, with an $R^2 = 10.4\%$.

Factor 2, SST. This factor, (X_2), is related to recruitment by a linear functional relationship between recruitment and X_2 , similar to a relationship between recruitment and sea surface temperature (SST) anomalies (see Equation (2)). We used actual March SST anomaly data from the Gulf of Alaska. A normal distribution was fit to the March SST anomaly data, and the parameters $\mu = 0.0943$ and $\sigma^2 = 9.29$ were estimated with maximum likelihood.

Factor 3, wind anomaly. This factor, (X_3), was intended to have no relationship to recruitment. This feature (i.e. no relation to recruitment) is often an important characteristic of actual data, and we include it here to examine the ability of the statistical procedures to identify this variable as unimportant. We call this variable wind anomaly. X_3 was a random variable defined as

$$X_3 \sim \text{lognormal}(\mu, \sigma^2)$$

where $\mu = 0.0$, and $\sigma^2 = 1.8943$. For any value of the random variable X_3 , the explanatory variable used to simulate recruitment in Equation (2) was set to a constant equalling the log of the mean of observed recruitment R_0 , $\ln(\mu_{R_0})$.

Random noise. Random noise was generated by defining a random variable as

$$\varepsilon \sim N(\mu, \sigma^2)$$

where $\mu = 0.0$, and $\sigma^2 = 0.74$. The value of σ^2 was approximated from the spawner–recruit fit by taking the

residual sum of squares and dividing by 40 degrees of freedom.

Generation of R_s . Simulated recruitment was generated by the equation

$$R_s = \alpha X_1 e^{[-\beta X_1 + \phi X_2 + \ln(\mu_{R_0}) + \varepsilon]} \quad (2)$$

where X_1 , X_2 , and X_3 , all random variables, are the data we typically measure and try to relate to the recruitment variation; $\alpha = 4.171$ and $\beta = 1.212$ are the spawner–recruit parameters; and ϕ is an arbitrary linear coefficient where $\phi = 0.25$. This approach is similar to that proposed by Hilborn and Walters (1992, p. 285). Thus, R_s is a combination of a non-linear response to a gamma random variable X_1 similar to spawning biomass, a linear relationship to a normal random variable X_2 , a constant mean level of recruitment ($\ln(\mu_{R_0}) = 2.0423$) unrelated to the lognormal random variable X_3 , and a multiplicative, lognormally distributed measurement error term. All recruitment effects occur in the birth year (i.e. there are no lag effects).

The random noise component in Equation (2) was generated under three different error levels: no error, $(1/2)\sigma^2$, and σ^2 . These error assignments are somewhat arbitrary, but we were generally guided on the range of magnitude in the error after examining the coefficient of variation of 42 different fish stock–recruitment time-series.

Norwegian spring-spawning herring

Data for Norwegian spring-spawning herring, potentially the largest of the herring stocks in the Northeast Atlantic, were taken from information presented in Torensen and Østvedt (2000). Biological data consisted of spawning-stock biomass SB (in million tonnes) and recruitment measured as three-year olds (R3 in billions) and age-0 (R0 in billions). The SB, R0, and R3 time-series were available for the periods 1907–1998, 1907–1997, and 1904–1995, respectively. We used the R3 time-series for recruitment for the period 1907–1998 (lagged 3 years to match SB), since the R0 and R3 time-series were highly correlated ($r = 0.992$, $p < 0.001$).

North Atlantic Oscillation (NAO) data represent the winter (December–March) index, which is based on the difference of normalized sea level pressure (SLP) between Lisbon, Portugal, and Stykkisholmur/Reykjavik, Iceland, from 1864 through 2000. The average winter SLP data at each station were normalized by division of each seasonal pressure by the long-term (1864–1983) mean standard deviation (data supplied by James Hurrell are available at <http://www.cgd.ucar.edu/~jhurrell/nao.stat.winter.html#winter>). The index was smoothed with a three-year moving average to reduce the influence of fluctuations with periods less than four years. The NAO time-series was available from 1864 to 2003.

Water temperature, available from the Russian Kola meridian transect (known as the Kola Line), represents the

average sea temperature ($^{\circ}\text{C}$) between 0 and 200 m in an area which stretches from $70^{\circ}30' \text{N}$ to $72^{\circ}30' \text{N}$ along the $33^{\circ}30' \text{E}$ meridian (Bochkov, 1982; Tereshchenko, 1996). The Kola Line time-series was available from 1904 to 1998 (R. Toresen, pers. comm.). These data will be called Kola sea temperature (KST) in subsequent discussion.

No time-lags were explored when fitting the different statistical models to the Norwegian spring-spawning herring and environmental data.

Statistical methods

Four statistical methods were applied to the simulated and actual recruitment time-series. Three were applied to recruitment on the absolute scale. These were non-linear regression (Hilborn and Walters, 1992), Generalized Additive Models (GAM) (Hastie, 1991), and Artificial Neural Networks based on the Generalized Regression Neural Network method of Specht (1991). Linear regression and GAMs were applied to herring recruitment time-series on the log scale.

Non-linear regression is a standard way of fitting spawner–recruit–environment models where the user has to specify *a priori* the functional relationships between the independent and dependent variables. Parameters are estimated by minimizing the squared difference between the observed and the predicted recruitment.

GAMs model the data non-parametrically by using scatterplot data smoothers that allow the GAM to suggest the functional relationship between independent and dependent variables. Smoothers are tools for summarizing the trend of a response measure as a function of one or more predictor measurements. An important property of a smoother is its non-parametric nature: it does not assume a rigid form for the dependence of the response variable on the predictor variable. Whereas linear models assume that the response is linear in each predictor, additive models assume only that each predictor affects the response in a smooth way. The response is modelled as a sum of smooth functions in the predictors, where the smooth functions are estimated automatically using scatterplot smoothers.

The Generalized Additive Model (GAM) can be written as

$$E(R_s | X_1, \dots, X_p) = \alpha + \sum_{j=1}^p f_j(X_j) + \varepsilon \quad (3)$$

which is the expectation that the response variable (R_s) is related to the covariates (X_1, \dots, X_p) by the additive predictor, $\alpha + \sum f_j(X_j)$. The amount of smoothing is determined by the number of degrees of freedom applied to the smoothing spline function applied to each covariate. We used cubic spline scatterplot smoothers in our application. More information on non-parametric GAM regression models can be found in Hastie and Tibshirani (1990) and Chambers and Hastie (1992). The use of GAMs in fisheries

applications is growing (Cury *et al.*, 1995; Swartzman *et al.*, 1995; Borchers *et al.*, 1997; Augustine *et al.*, 1998; Beare and Reid, 2002; Ciannelli *et al.*, 2004), and useful applications to spawner–recruit–environment models have been demonstrated (Myers *et al.*, 1995; Jacobson and MacCall, 1995; Daskalov, 1999).

Neural networks are computer applications devised in the 1940s that attempt to mimic the neurophysiology of the human brain. They are capable of “learning” patterns found in the data. Often they are referred to as “black box” models because their parameters are generally uninterpretable, so the emphasis is on prediction rather than on model building or process understanding. For some reason, neural networks are not very common in fisheries applications, although there has been a recent increase in their use (Komatsu *et al.*, 1994; Brey *et al.*, 1996; Akoi and Komatsu, 1997; Huse and Gjosæter, 1999; Laë *et al.*, 1999). Their utility specifically addressing the recruitment–environment problem has also recently been demonstrated (Chen and Ware, 1999; Huse and Ottersen, 2003).

Multi-layer feed-forward neural network models are multivariate statistical models that can be viewed as non-linear regression models. Sometimes these are called multi-layer perceptron (MLP) models (Sarle, 1994). MLPs are general purpose, flexible non-linear models that can approximate virtually any function to any degree of accuracy. White (1992) refers to them as universal approximators. MLPs are especially useful in situations where there is little knowledge about the mathematical form of the relationship between independent and dependent variables.

A feed-forward neural network model used in this study is based on the General Regression Artificial Neural Networks developed by Specht (1991). A schematic of a generalized Artificial Neural Network model is given in Figure 1.

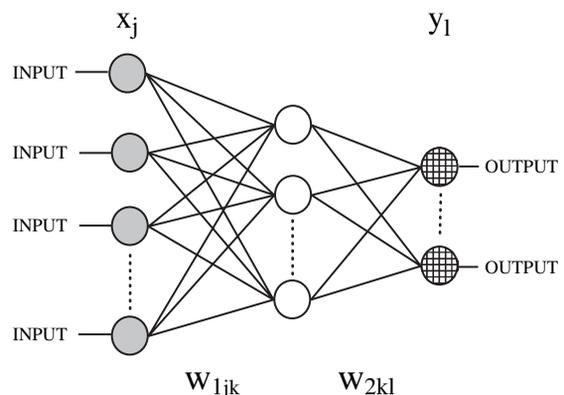


Figure 1. A schematic of a generalized neural network model showing input neurons (x_j) in grey, hidden layer neurons in white, output neurons (y_l) in cross-hatched pattern, and the input weight (W_{1jk}) and the output weight connections (W_{2kl}).

The equations describing the neural network model can be given by

$$y_1 = \tilde{g} \left(\sum_{k=1}^K w_{2kl} g \left(\sum_{j=1}^J w_{1jk} x_j + \theta_{1j} \right) + \theta_{2k} \right) \quad (4)$$

$$\tilde{g}(x) = g(x) = (1 + e^{-x})^{-1}$$

where the hidden neurons are a weighted sum of the input weights (w_{1jk}) and the input neurons (x_j), the output neurons (y_1) are a weighted sum of the hidden neurons, and the output weights (w_{2kl}) and θ_{1j} and θ_{2k} are constant terms often referred to as the bias terms. The response, y_1 , is a transformed linear combination of transformed linear combinations of the predictors using the logistic transformation $g(x)$. It is this flexible form, with many parameters, that gives the network its universal approximation property, i.e. the ability to fit a wide variety of functions.

The input and output weights are adjusted by an iterative scheme to minimize the residual sum of squares (RSS) using the objective function

$$RSS = \sum_i \sum_l (y_{li} - \hat{y}_{li})^2 \quad (5)$$

where y_{li} are the observed data, and \hat{y}_{li} are the data predicted from the neural network. Relative weights for each of the input variables can be used to judge the relative importance of the input variables similar to the use of estimated coefficients in a linear regression. Often neural networks have more than one output neuron, indexed by l . We have only one output neuron (recruitment, R_s) in our analysis so, in our application, Equation (5) is evaluated over only one summation (year, i).

For each neural net, two types of training were used: a non-linear least squares optimization procedure and a genetic algorithm (Goldberg, 1989; Davis, 1991; Michalewicz, 1992). The genetic algorithm is a derivative-free, “brute force”, stochastic optimization method recently being used in fisheries applications (Chen *et al.*, 2000) that systematically searches the entire parameter space by mimicking the evolution of populations. In the genetic algorithm, several possible solutions to a problem are generated. For each solution, the vector of parameters is encoded into “genes” and “chromosomes”. Each is tested for its performance (fitness), then a fraction of the “good” solutions is selected, while others are eliminated (survival of the fittest). New (more fit) generations of possible solutions are created using the process of reproduction, crossover, and mutation. New generations are produced and evaluated until convergence is achieved. Genetic algorithms are very robust, but the penalty for generalization is computational intensity. This is not much of an issue for relatively small problems implemented on fast computers.

Ward Systems Group[®] Neuroshell Easy Predictor[™] (Release 2.0) was used to implement the Generalized Artificial Neural Network; S Plus[®] for Windows (version 6.1

Professional, Release 1) was used to implement the linear and non-linear regressions and GAM statistical procedures; and MATLAB[®] (version 6.5.0.180913a, Release 13) was used to fit the statistical distribution to data and generate the simulated data.

Testing and evaluation procedures

Each time-series was split into a testing segment and a forecasting segment. The last five observations in the simulated time-series were “held back” to evaluate each statistical procedure’s ability to forecast recruitment. In the observed herring time-series, the last ten observations were held back for the forecasting segment, so that we had a complete set of contrasting recruitment levels (low, average, and high). The decision to keep the last five or ten observations was arbitrary, but based on the procedure described in Saila (1996). Alternatively, we could have selected a random set of five (or ten) observations.

The testing segment was used to fit each model (i.e. estimate the parameters for each method). Then the fit model was applied to the forecasting segment of the time-series, and the observed and forecast recruitment values were compared. Relationships were tested using untransformed and transformed (natural logarithm) recruitment time-series.

LR and NLR models used the Gauss–Newton minimization routine to solve for the model parameters. GAM models used the cubic spline smoother of varying degrees of freedom. ANNs used the non-linear minimization method via the back propagation technique called Turboprop[®], which is a proprietary algorithm of Ward Systems Group[®], and genetic algorithms to solve the network. Model fit and parsimony were evaluated through analysis of deviance using approximate F-tests and the Akaike information criteria (AIC: Chambers and Hastie, 1992). Suitable GAM models were identified by a stepwise selection procedure in which each independent variable was started with one degree of freedom with a maximum of four degrees of freedom allowed by the selection procedure. Selection of the best model was based on the AIC criteria.

Results

Simulated data did possess statistical characteristics of actual recruitment time-series (Table 1), showing close

Table 1. Descriptive statistics comparing simulated data and actual recruitment time-series data.

	Mean	Median	Minimum	Maximum	Autocorrelation*
Simulated data	1.05	0.92	0.05	3.07	0.632
Actual data	0.86	0.43	0.08	3.51	0.539

*Autocorrelation calculated at a lag of one year.

correspondence in terms of the values of mean, median, minimum, maximum, and autocorrelation at a lag of one. The Norwegian spring-spawning herring spawner biomass and age-3 recruitment time-series, along with the Kola sea temperature (KST) and NAO time-series, are presented in Figure 2. After a period of high abundance, the spawning biomass collapsed towards the end of the 1960s (Dragesund *et al.*, 1980; Jakobsson, 1980). During the 1970s and early 1980s, the stock was gradually rebuilt, and the first strong year class after the collapse came in 1983. By the early 1990s, the stock had regained much of its previous distribution area, was found feeding in the Norwegian Sea, and has since been rebuilding (Dommasnes *et al.*, 2004). Recruitment time-series from both the simulated data and the herring data demonstrate the typical property of the somewhat random appearance of strong year classes. Comparison of recruitment time-series for herring and simulated data plotted against spawning biomass and environmental covariates (Figure 3) show similar patterns of variability. Also evident is the usual lack of an apparent relationship between recruitment and spawner biomass.

Results of fitting GAM, ANN, and NLR models to the simulated data with different levels of error are presented in Table 2. Comparisons among methods can be made for similarly configured models by comparing the models for each error level. For example, compare the GAM models (GAM2, GAM4, GAM6), with the NLR models (NLR1,

NLR3, NLR5), and with the ANN models (ANN7–ANN12). Reviewing the models demonstrates the ability of the method to identify important variables and screen out unimportant variables. The NLR model fit to data with no error (NLR1), described the data perfectly, and showed no error in the training and forecast data segments. This is as we expected, because we had to specify the functional relationship before fitting. In the case just described, we fit the NLR model to the data using the formula employed to generate the data. Even though WIND was not related to recruitment, the NLR procedure incorrectly identified WIND as a significant variable (NLR2 and NLR4) when the error level was high. WIND explained 1.5% of the variation in recruitment for error level 1 (NLR2). More disturbingly, WIND explained 13.9% of the variation in recruitment for error level 2 (NLR4).

The statistically significant GAM models, when fit using a stepwise selection procedure (GAM1, GAM3, GAM5), did screen out WIND as unimportant in all cases except when the error level was high. The GAM5 model incorrectly selected WIND as an important variable, explaining 13.1% of the variation in recruitment similar to the non-linear regression model (NLR4). Results of applying the GAM model (GAM1) to the Norwegian spring-spawning herring data are shown in Figure 4. The GAM plot shows a very weak and statistically insignificant linear relationship of NAO to herring recruitment. The NAO was correctly

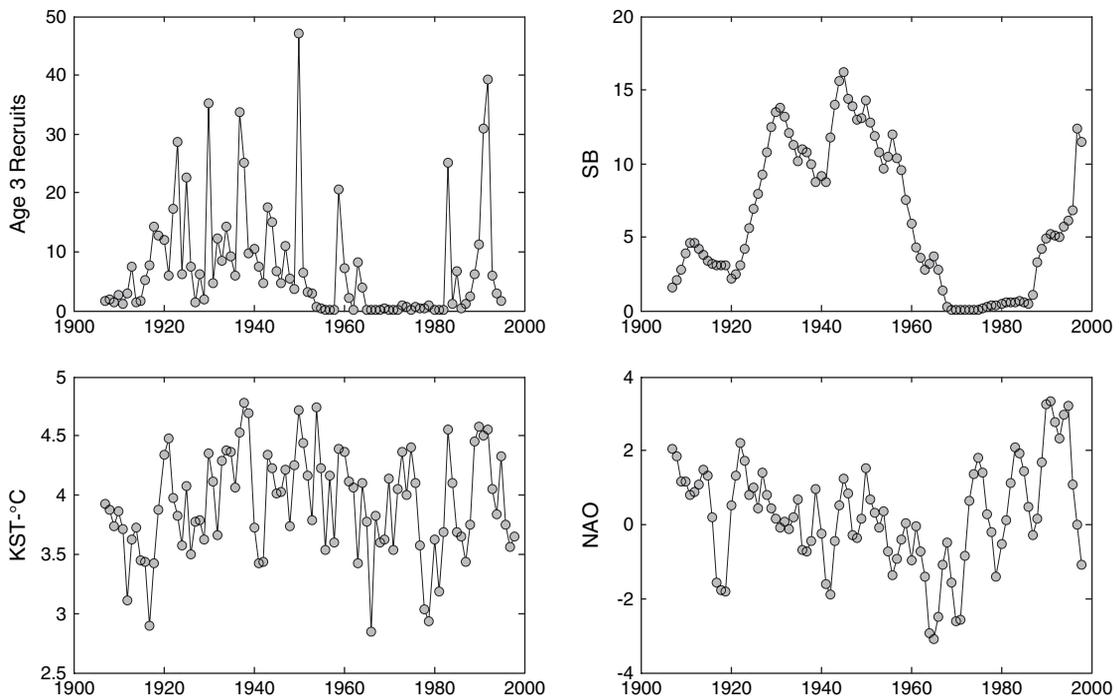


Figure 2. Time-series of Norwegian spring-spawning herring data, 1907–1999, showing the time-series of age-3 recruits (R in billions), spawning biomass (SB in million tonnes), Kola sea temperature (KST in °C), and North Atlantic Oscillation index (NAO, normalized sea level pressure anomaly).

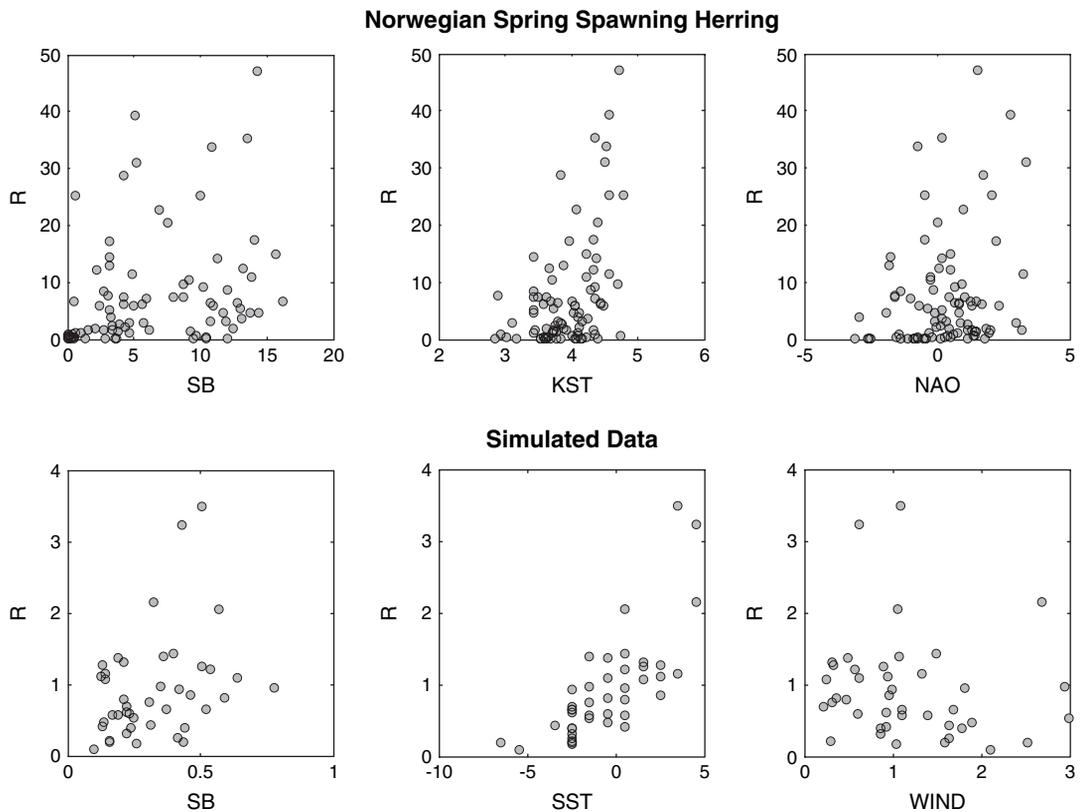


Figure 3. Relationship of recruitment to spawning biomass and environmental covariates for Norwegian spring-spawning herring (top row) and simulated data (bottom row). For herring the units are age-3 recruitment (R in billions), spawning biomass (SB in million tonnes), Kola sea temperature (KST in $^{\circ}C$), and North Atlantic Oscillation index (NAO , normalized sea level pressure anomaly). For simulated data the units are age-3 recruits (R in billions), spawning biomass (SB in million tonnes), sea surface temperature (SST in $^{\circ}C$), and wind anomaly ($WIND$ in $m\ s^{-1}$).

identified by the ANNs as not being influential, as indicated by the low weighting factor, and the ANN results were consistent with results from the GAM model that appear in Figure 4. The GAM model also suggested a statistically significant curvilinear relationship between SB and recruitment, although the overall trend hints at a weak positive linear relationship between SB and recruitment over the observed range of SB . Also, the GAM model showed that temperature did not affect recruitment for Kola sea temperatures below about $4^{\circ}C$. Thereafter, temperature had a positive linear relationship with recruitment.

ANN models fit and forecast better using the standard non-linear estimation method compared with using the genetic algorithm to solve for the model parameters. In general, the 3-hidden neuron ANNs performed better when compared with the 2-hidden neuron ANN, but the 3-hidden neuron models carried more parameters. ANNs that included $WIND$ fit the training segment better than a corresponding model without $WIND$ (i.e. ANN1 vs. ANN7), but the ANN model without $WIND$ (ANN7) forecasts better, as indicated by a lower MSE for the forecast segment. Relative weights

coming out of the ANN models showed $WIND$ as the least important, even in models that included the variables (ANN1, ANN2, ANN3, ANN4, ANN5, and ANN6). As indicated by the relative weighting factors, SST was consistently identified as important. Also of interest is the influence of error level on the estimated weighting factors of the three-variable ANN models (ANN1–ANN6). At error level 0, SST was identified as the most important variable, followed by SB . $Wind$ was consistently identified as the least important. As the error level increased from 0 to 2, the original large contrast between the SB and SST weighting factors at error level 0 diminished. At error level 2, the ability of the ANN to identify important variables appeared compromised, with the weighting factors for SB and SST being almost equal.

Of the four methods examined, NLR was the least accurate at forecasting. GAM and ANN both performed better at forecasting regardless of the error level (e.g. NLR5 vs. GAM6 vs. ANN12). The ANN marginally outperformed a similarly configured GAM in forecast accuracy (ANN8 vs. GAM2), with the performance of each degrading as more error was added to the data (e.g. GAM2, GAM4,

Table 2. Summary information from fitting General Additive Models (GAM), Artificial Neural Networks (ANN), and Non-linear Regression (NLR) statistical methods to the simulated data set. Data are tabulated separately for training and forecasting data segments. The dependent variable, recruitment, was examined on the arithmetic scale only. Error level codes are explained in the text. Variables in bold were incorrectly identified as statistically significant.

Model	Error	Terms	# parameters	Weights*	Training (# obs = 37)		Forecast (# obs = 5)
					R ²	MSE	MSE
GAM1	0	R = SB(1) + SST(3)†	5	—	0.944	0.031	0.018
GAM2	0	R = SB(2) + SST(2)	5	—	0.919	0.045	0.021
GAM3	1	R = SB(1) + SST(2)	4	—	0.683	0.129	0.153
GAM4	1	R = SB(2) + SST(2)	5	—	0.702	0.121	0.168
GAM5	2	R = SB(3) + SST(3) + WIND(4)	11	—	0.752	0.205	0.366
GAM6	2	R = SB(2) + SST(2)	5	—	0.555	0.368	0.258
NLR1	0	R = SB × SST	3	—	1.0	0.0	0.0
NLR2	1	R = SB × SST × WIND	4	—	0.750	0.101	0.201
NLR3	1	R = SB × SST	3	—	0.735	0.108	0.239
NLR4	2	R = SB × SST × WIND	4	—	0.634	0.303	0.329
NLR5	2	R = SB × SST	3	—	0.494	0.419	0.305
ANN1	0	ANN([SB,SST,WIND],3)‡	14	W ₁ = 0.139 W ₂ = 0.731 W ₃ = 0.130	0.933	0.036	0.042
ANN2	0	ANN([SB,SST,WIND],2)	10	W ₁ = 0.169 W ₂ = 0.714 W ₃ = 0.125	0.932	0.037	0.039
ANN3	1	ANN([SB,SST,WIND],3)	14	W ₁ = 0.388 W ₂ = 0.621 W ₃ = 0.041	0.691	0.125	0.193
ANN4	1	ANN([SB,SST,WIND],2)	10	W ₁ = 0.340 W ₂ = 0.627 W ₃ = 0.033	0.691	0.125	0.192
ANN5	2	ANN([SB,SST,WIND],3)	14	W ₁ = 0.408 W ₂ = 0.431 W ₃ = 0.161	0.660	0.323	0.317
ANN6	2	ANN([SB,SST,WIND],2)	10	W ₁ = 0.408 W ₂ = 0.395 W ₃ = 0.197	0.607	0.325	0.357
ANN7	0	ANN([SB,SST],3)	11	W ₁ = 0.288 W ₂ = 0.712	0.902	0.054	0.010
ANN8	0	ANN([SB,SST],2)	8	W ₁ = 0.275 W ₂ = 0.725	0.876	0.068	0.006
ANN9	1	ANN([SB,SST],3)	11	W ₁ = 0.195 W ₂ = 0.805	0.699	0.122	0.209
ANN10	1	ANN([SB,SST],2)	8	W ₁ = 0.353 W ₂ = 0.647	0.686	0.128	0.203
ANN11	2	ANN([SB,SST],3)	11	W ₁ = 0.561 W ₂ = 0.439	0.50	0.413	0.236
ANN12	2	ANN([SB,SST],2)	8	W ₁ = 0.546 W ₂ = 0.454	0.50	0.414	0.240

*Relative weighting factors for each of the independent variables: W₁ for SB, W₂ for SST, W₃ for WIND.

†Degrees of freedom for GAM spline smoothers are indicated in parentheses for each independent variable.

‡Artificial Neural Network configuration. Independent variables included in each model are in brackets. Second number in parenthetic pair represents the number of hidden neurons.

GAM6, and ANN8, ANN10, ANN12). Compared with GAM, the improved performance of the ANN did come at the cost of the model having more parameters. It is worthwhile noting that the improvement was still statistically significant as

indicated by the AIC criteria, which imposes a penalty for the number of parameters in a model.

Results of fitting GAM, ANN, LR, and NLR models to the spring-spawning herring data are presented in Table 3.

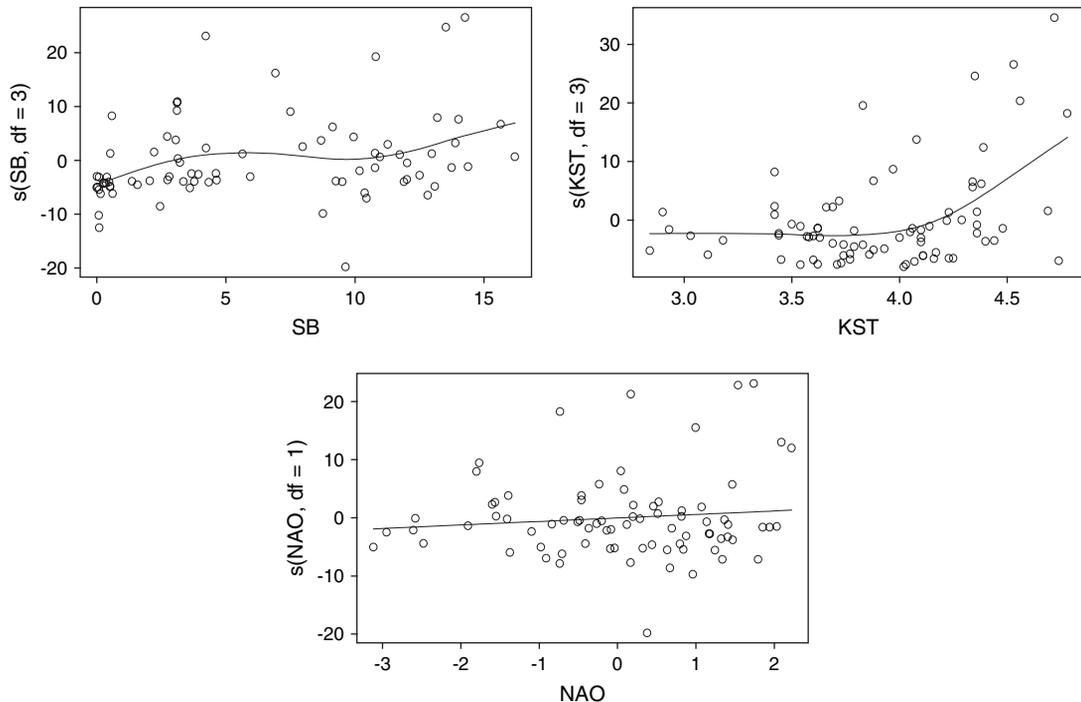


Figure 4. Plots from applying GAM1 to the Norwegian spring-spawning herring data.

This data set was very long, and issues of the number of parameters estimated relative to the number of observations were not as much of a concern as with the simulated data. The GAM model with only spawning biomass (SB) and Kola sea temperature (KST) performed best (GAM4), accounting for 39% of the variation in herring recruitment. This was not the model with the lowest R^2 , but it did have a low R^2 compared with GAM1 and also did well in the forecasting segment, having one of the lowest MSEs on the arithmetic recruitment scale.

The three-variable NLR model fit and forecast better compared with multiple linear regression with recruitment on the log scale (NLR1 vs. LR1 and LR2). The NLR model accounted for 33% of the variation in herring recruitment. This value was similar to the GAM fit (GAM4), but the GAM had a much lower MSE in the forecasting segment.

The ANN models also performed well, accounting for about 32% of the variation in herring recruitment when all three independent variables were included in the model. The ANN as well as the GAM did not perform in the forecasting segment (GAM1 vs. ANN2), but the GAM and ANN performed equally well in the training segment.

Relative weighting factors from the ANN fits, using recruitment on the arithmetic scale, indicated that the Kola sea temperature (KST) was the most important variable influencing recruitment, followed by SB, and finally the NAO was least important (ANN1, ANN2, ANN3). The relative magnitudes of coefficients estimated from non-linear

regression (NLR1) show similar trends in influence of the covariates. However, this trend was different when recruitment was on the log scale. In these cases, SB was the most important variable, followed by the NAO, with KST being the least important (ANN4, ANN5, ANN6).

One final observation compares a linear regression model (LR2) with an ANN with no hidden neurons (ANN7). In this ANN configuration, the input neurons are connected directly to the output neurons, and it can be seen that this simplified ANN is equivalent to a basic linear regression model.

In general, the data in Table 3 suggest that methods that fit the training segment on the log scale, then forecast on the log scale and back-transformed the forecasts to the arithmetic scale, did not perform well. Small deviations between the observed and the predicted values on the log scale were amplified when back-transformed.

Discussion

Studies in the fisheries literature attempting to explain the effect of various environmental factors on recruitment are numerous, e.g. Sutcliffe (1972, 1973); Leggett *et al.* (1984); Drinkwater (1987), and Drinkwater and Myers (1987). The complexity of the marine ecosystem and the inherent non-linearities often cause apparent relationships to fail when retested with additional data (Myers, 1998). Still, further work on the problem has been advocated (Kope and

Table 3. Summary information from fitting General Additive Models (GAM), Artificial Neural Networks (ANN), Non-linear Regression (NLR) and Linear Regression (LR) statistical methods to the Norwegian spring-spawning herring data. Data are tabulated separately for training and forecasting data segments on both the natural log and arithmetic scale. The dependent variable, age-3 recruitment, was examined on the natural log and arithmetic scale.

Model	Dependent variable	Terms	# parameters	Weights*	Training (# obs = 79)				Forecast (# obs = 10)	
					Log		Arithmetic		Log	Arithmetic
					R ²	MSE	R ²	MSE	MSE	MSE
GAM1	R	R = SB(3) + KST(3) + NAO(1)†	8	—	—	—	0.3947	54.30	—	93.78
GAM2	R	R = SB(3) + KST(3) + NAO(2)	9	—	—	—	0.4176	49.36	—	100.29
GAM3	R	R = SB(3) + KST(3) + NAO(3)	10	—	—	—	0.4397	47.49	—	140.46
GAM4	R	R = SB(3) + KST(3)	7	—	—	—	0.3908	51.63	—	92.93
GAM5	log R	log R = SB(3) + KST(2) + NAO(1)	7	—	0.5474	2.122	0.2685	62.00	1.966	401.08
LR1	log R	log R = SB + KST + NAO	4	—	0.3649	2.970	0.0489	80.60	1.092	151.41
LR2	log R	log R = SB + NAO	3	—	0.3631	2.986	0.0057	85.23	1.154	159.92
NLR1	R	R = SB + KST + NAO	4	—	—	—	0.3295	56.82	—	195.97
ANN1	R	ANN([SB,KST,NAO],3)‡	14	W ₁ = 0.344 W ₂ = 0.553 W ₃ = 0.103	—	—	0.3288	56.88	—	139.71
ANN2	R	ANN([SB,KST,NAO],2)	10	W ₁ = 0.351 W ₂ = 0.561 W ₃ = 0.088	—	—	0.3261	57.11	—	122.30
ANN3	R	ANN([SB,KST,NAO],1)	8	W ₁ = 0.421 W ₂ = 0.574 W ₃ = 0.005	—	—	0.2568	62.99	—	125.16
ANN4	log R	ANN([SB,KST,NAO],3)	14	W ₁ = 0.753 W ₂ = 0.049 W ₃ = 0.216	0.3985	2.825	0.0472	80.75	1.961	224.23
ANN5	log R	ANN([SB,KST,NAO],2)	10	W ₁ = 0.690 W ₂ = 0.072 W ₃ = 0.237	0.3964	2.830	0.0980	76.44	1.599	212.92
ANN6	log R	ANN([SB,KST,NAO],1)	6	W ₁ = 0.571 W ₂ = 0.068 W ₃ = 0.360	0.3658	2.973	0.0593	79.73	1.087	159.40
ANN7	log R	ANN([SB,NAO],0)	3	W ₁ = 0.601 W ₃ = 0.399	0.3631	2.986	—	—	—	—

*Relative weighting factors for each of the independent variables: W₁ for SB, W₂ for KST, W₃ for NAO.

†Degrees of freedom for GAM spline smoothers are indicated in parentheses for each independent variable.

‡Artificial Neural Network configuration. Independent variables included in each model are in brackets. Second number in parenthetic pair represents the number of hidden neurons.

Botsford, 1990; Tyler, 1992), despite some criticisms (Walters and Collie, 1988).

Our results demonstrate that “model free” approaches (i.e. no *a priori* specification of a functional relationship between recruitment and predictor variables) such as GAM and ANN can be quite flexible and provide useful assistance in identifying relationships between recruitment and influential factors. GAMs, in particular, have the benefit of identifying or suggesting the form of the functional relationship. We recommend first applying non-parametric GAMs to the predictor and response variables so as to ascertain the functional form empirically from the data without the need for *a priori* assumptions, and then exploring these relationships further with standard parametric techniques or more parsimonious models.

Artificial Neural Networks also show promise, especially in the forecasting step of the analysis. The relative weighting factors coming out of ANN models provide useful information, similar to information coming out of stepwise multiple linear regression models, in that they provide some sense of ranking of the independent variables according to their contribution to the dependent variable and the degree to which they explain recruitment variability. Caution should be used when applying ANNs, since the number of parameters can grow large quickly, and it is easy to overparameterize the model. Note that the number of hidden neurons (h) relative to the number of input (i) and output variables (o) all directly determine the number of parameters in the model according to the formula ($i \times h + h \times o + 2$). The more hidden neurons, the better the fit to the estimation segment of the data. This comes at a cost during the forecasting step, since an ANN that is overfit loses its ability to generalize to unseen data, a situation that impedes the ability of the ANN to forecast. When considering estimation in conjunction with forecasting, it is better to consider a balance between the best method for estimation vs. the best method for forecasting. Often the choice of one compromises the other.

Despite the promising performance of the non-parametric methods, we did see examples of spurious correlations in the analysis of the simulated data. There was more than one example of a method indicating that a variable was important to recruitment when, in fact, it had no relationship to recruitment at all. As noted by Tyler (1992), “Spurious correlations are the first enemy of recruitment biologists”, and we did see evidence of this unpleasant result increase as the noise in the data increased. When the data have a high level of variability, we showed that it is possible for a method to identify a variable as statistically related to recruitment when, in fact, it does not have any relationship to recruitment, even under relatively controlled data conditions. Unfortunately, when analysts are examining real data, they do not have the benefit of having this knowledge beforehand, as in the case of looking at simulated data with known properties. This underscores the need for building good, conceptual models first, then, guided by

hypotheses, the judicious application of the appropriate statistical models (or suite of statistical models) to quality data sets.

Compared with parametric methods, non-parametric methods such as GAM and ANN performed better using recruitment on the arithmetic scale in both estimation and forecasting. Log transforms are often performed in order to make data conform to assumptions of the statistical method. The transformation stabilizes variances, turns skewed data normal, and turns multiplicative errors into additive errors. However, our results show that estimating on the log scale, forecasting on the log scale, and then back-transforming the forecast onto the arithmetic scale turns small deviations between the observed and the predicted values on the log scale into large differences on the arithmetic scale. In general, methods using recruitment on the log scale did not perform well when forecasting. Ultimately, the test of a model is its ability to predict (Tyler, 1992), especially in an operational sense.

Results from applying the three methods to the herring data showed the strong influence of temperature, but did not indicate that the NAO was important. Hjermann *et al.* (2004) found a lagged response of herring to temperatures. It is interesting to note that their observation, that abundance declined for temperatures below $\sim 3.4^\circ\text{C}$, is very similar to the GAM relationship between recruitment and temperature revealed in this study. Our results suggest that a temperature threshold affects recruitment beginning at 4°C . Mikkelsen and Pedersen (2004) also showed that herring recruitment is positively correlated with higher water temperatures. Others (Sætre *et al.*, 2002) showed that strong year-class strength at the 0-group and as three-year olds occurs in warm years when the mean temperature during winter (December–March) at the Kola section is above average. They concluded that a significant relationship exists between temperature at the Kola section and the 0-group index, both for mean annual temperatures and on a monthly basis. In an analysis of stock–environment recruitment models for Norwegian spring-spawning herring, Fisker and Slotte (2002) concluded that inclusion of sea temperature in the model improved the explanatory ability of the model and removed autocorrelation from the residual variability.

Ottersen and Loeng (2000) offered an explanation for the mechanism underlying the temperature–recruitment relationship. They found that on average, 0-group herring were longer in length in warm years than in cold ones. They speculated that high temperatures caused increased prey production, which led to higher growth rates and higher survival through vulnerable larval and juvenile stages. Higher temperatures also contributed to increased development rates, which also reduced the time spent in vulnerable life stages. This *in situ* observation supports laboratory experiments and modelling on herring larvae that show the positive effects of adequate food and increased temperatures (Fisker and Folkvord, 1999). Despite these intriguing suggestions, the role of temperature in the mid-1960s collapse of Barents Sea herring remains

controversial, and it cannot be determined if the decline was a result of heavy fishing, a long-term drop in temperature, or the unfortunate simultaneous coincidence of these two influences (Toresen and Østvedt, 2000).

Currently, there are contrasting philosophies regarding the utility of cause and effect of environment—recruitment models. Because of the mixed success in establishing clear connections between variability in the environment and recruitment, two attitudes flourish regarding the utility of environment—recruitment studies. In one camp, Walters and Collie (1988) criticized correlative environment—recruitment studies as futile because of biases, measurement error, and the near certainty of spurious correlations. Others are optimistic: Kope and Botsford (1990) claim that correlative studies provide information on patterns that lead to the formulation of testable hypotheses.

The high dimensionality of the problem presents its own difficulties, since, in many cases, personal judgement and intuition are the tools used to collapse large multidimensional data sets into a relevant subset amenable to statistical analysis. Statistical issues aside, we should be guided by the careful use of conceptual models.

Admittedly, the recruitment—environment problem is a difficult one, but it does not mean that we should stop exploring models and techniques to help understand the factors that control recruitment dynamics and their spatial and temporal scales of influence. Simple statistical approaches still have their place if used appropriately. For example, it is known that stepwise regression is not a good tool for prediction, because its selection criterion for variable selection is geared towards error rates rather than predictive power. Nevertheless, the utility of linear regression, despite the widespread criticisms levelled against it, is still being demonstrated today (Carscadden *et al.*, 2000; Stein and Borovkov, 2004; Yatsu *et al.*, 2005).

The type of data we collect, as they relate to recruitment variability and the factors that influence it, probably will not change dramatically in the near future. As advocated by Beverton (1989), we should endeavour to treat the data differently in a statistical sense. That is what we have attempted to do in this study. We advocate more comparative analyses such as we have attempted, so that researchers applying statistical methods to recruitment—environment data have a better appreciation of the strengths and weaknesses of each method, as well as the method's ability to adequately perform in the face of common data pathologies.

Acknowledgements

The authors extend to Reidar Toresen and Ole Johan Østvedt our sincere appreciation for generously sharing the Norwegian spring-spawning herring and Kola sea temperature data. We also thank two anonymous reviewers for their useful comments and suggestions. An earlier version of the manuscript was improved by comments offered by David

Somerton and Susan Picquelle. This research is contribution FOCI-0505 to NOAA's Fisheries-Oceanography Coordinated Investigations.

References

- Akoi, I., and Komatsu, T. 1997. Analysis and prediction of the fluctuation of sardine abundance using a neural network. *Oceanologia Acta*, 20: 81–88.
- Augustine, N. H., Borchers, D. L., Clarke, E. D., Buckland, S. T., and Walsh, M. 1998. Spatiotemporal modelling for the annual egg production methods of stock assessment using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 2608–2621.
- Beare, D. J., and Reid, D. G. 2002. Investigating spatio-temporal change in spawning activity by Atlantic mackerel between 1977 and 1998 using Generalized Additive Models. *ICES Journal of Marine Science*, 59: 711–724.
- Beverton, R. J. H. 1989. Closing address: the symposium in perspective. *Journal of Fish Biology*, 35(Suppl. A): 355–363.
- Bochkov, Y. A. 1982. Water temperature in the 0–200 m layer in the Kola-Meridian in the Barents Sea, 1900–1981. *Sbornik Nauchnykh Trudov Polar Institute of Marine Fisheries and Oceanography of Murmansk*, 46: 113–122.
- Borchers, D. L., Buckland, S. T., Priede, I. G., and Ahmadi, S. 1997. Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 54: 2727–2742.
- Brey, T., Jarre-Teochmann, A., and Borlich, O. 1996. Artificial neural networks versus multiple linear regression: predicting P/B ratios from empirical data. *Marine Ecology Progress Series*, 140: 251–256.
- Carscadden, J. E., Frank, K. T., and Leggett, W. C. 2000. Evaluation of an environment—recruitment model for capelin (*Mallotus villosus*). *ICES Journal of Marine Science*, 57: 412–418.
- Chambers, J. M., and Hastie, T. J. (Eds). 1992. *Statistical Models*. S. Wadsworth & Brooks/Cole Advanced Books and Software, Pacific Grove, CA. 608 pp.
- Chen, D. G., Hargreaves, N. B., Ware, D. M., and Liu, Y. 2000. A fuzzy logic model with genetic algorithm for analyzing fish stock-recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, 57: 1878–1887.
- Chen, D. G., and Irvine, J. R. 2001. A semi-parametric model to examine stock-recruitment relationships incorporating environmental data. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 1178–1186.
- 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.
- Ciannelli, L., Chan, K-S., Bailey, K. M., and Stenseth, N. C. 2004. Non-additive effects of the environment on the survival of a large marine fish population. *Ecology*, 85: 3418–3427.
- Cury, P., Roy, C., Mendelssohn, R., Bakin, A., Husby, D. M., and Parrish, R. H. 1995. Moderate is better: exploring non-linear climatic effects on the California northern anchovy (*Engraulis mordax*). In *Climate Change and Northern Fish Populations*, pp. 417–424. Ed. by R. J. Beamish. *Canadian Special Publication in Fisheries and Aquatic Sciences*, 121.
- Cushing, D. H. 1988. The study of stock and recruitment. In *Fish Population Dynamics*, 2nd edn, pp. 105–128. Ed. by J. A. Gulland. John Wiley & Sons Ltd, Chichester, UK.
- Daskalov, G. 1999. Relating fish recruitment to stock biomass and physical environment in the Black Sea using generalized additive models. *Fisheries Research*, 41: 1–23.
- Davis, L. D. (Ed). 1991. *Handbook of Genetic Algorithms*. Van Nostrand and Reinhold, New York. 385 pp.

- Dommasnes, A., Melle, W., Dalpadado, P., and Ellertsen, B. 2004. Herring as a major consumer in the Norwegian Sea. *ICES Journal of Marine Science*, 61: 739–751.
- Dorn, M., Barbeaux, S., Guttormsen, M., Megrey, B. A., Hollowed, A., Brown, E., and Spalinger, K. 2003. Assessment of walleye pollock in the Gulf of Alaska. *In* Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska, 2003. pp. 33–124. North Pacific Fisheries Management Council, Anchorage, AK.
- Dragesund, O., Hamre, J., and Ulltang, Ø. 1980. Biology and population dynamics of the Norwegian spring-spawning herring. *In* The Assessment and Management of Pelagic Fish Stocks, pp. 43–71. Ed. by A. Saville. *Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, 177.
- Drinkwater, K. F. 1987. "Sutcliffe revisited": previously published correlations between fish stocks and environmental indices and their recent performance. *In* Environmental Effects on Recruitment in Canadian Atlantic Fish Stocks, pp. 41–61. Ed. by R. I. Pery and K. T. Frank. Canadian Technical Report in Fisheries and Aquatic Sciences, 1556.
- Drinkwater, K. F., and Myers, R. A. 1987. Testing predictions of marine fish and shellfish landings from environmental variables. *Canadian Journal of Fisheries and Aquatic Sciences*, 44: 1568–1573.
- 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.
- Fisken, Ø., and Folkvord, A. 1999. Modelling growth and ingestion processes in herring *Clupea harengus* larvae. *Marine Ecology Progress Series*, 184: 273–289.
- Fisken, Ø., and Slotte, A. 2002. Stock-environment recruitment models for Norwegian spring spawning herring (*Clupea harengus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 211–217.
- Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley, Reading, MA. 412 pp.
- Goodyear, C. P., and Christensen, S. W. 1984. On the ability to detect the influence of spawning stock on recruitment. *North American Journal of Fisheries Management*, 4: 186–193.
- Hastie, T. J. 1991. Generalized additive models. *In* Statistical Models, pp. 249–308. Ed. by J. M. Chambers, and T. J. Hastie. S. Wadsworth & Brooks/Cole Advanced Books and Software, Pacific Grove, CA. 608 pp.
- Hastie, T. J., and Tibshirani, R. J. 1990. Generalized Additive Models. *Monographs on Statistics and Applied Probability* 43. Chapman & Hall/CRC, New York, NY. 335 pp.
- Hilborn, R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York, NY. 570 pp.
- Hjermann, D. Ø., Stenseth, N. C., and Ottersen, G. 2004. Indirect climatic forcing of the Barents Sea capelin: a cohort model. *Marine Ecology Progress Series*, 273: 229–238.
- Huse, G., and Gjøsæter, H. 1999. A neural network approach for predicting stock abundance of the Barents Sea capelin. *Sarsia*, 84: 457–464.
- Huse, G., and Ottersen, G. 2003. Forecasting recruitment and stock biomass of Northeast Arctic cod using neural networks. *Scientia Marina*, 67(Suppl. 1): 325–335.
- 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.
- Jakobsson, J. 1980. The North Icelandic herring fishery and environmental conditions, 1960–1968. *Rapports et Procès-Verbaux des Réunions du Conseil International pour l'Exploration de la Mer*, 177: 460–465.
- Komatsu, T., Aoki, I., Mitani, I., and Ishi, T. 1994. Predication of catch of Japanese sardine larvae in Sagami Bay using a neural network. *Fisheries Science*, 60: 385–391.
- Kope, R., and Botsford, L. W. 1990. Determination of factors affecting recruitment of chinook salmon *Oncorhynchus tshawytscha* in central California. *Fishery Bulletin*, U.S., 88: 257–269.
- Koslow, J. A. 1992. Fecundity and the stock-recruitment relationship. *Canadian Journal of Fisheries and Aquatic Sciences*, 49: 210–217.
- Laë, R., Lek, S., and Moreau, J. 1999. Predicting fish yield of African lakes using neural networks. *Ecological Modelling*, 120: 325–335.
- Leggett, W. C., Frank, K. T., and Carscadden, J. E. 1984. Meteorological and hydrographic regulation of year-class strength in capelin (*Mallotus villosus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 41: 1193–1201.
- Mackinson, S., Vasconcellos, M., and Newland, N. 1999. A new approach to the analysis of stock-recruitment relationships: "model-free estimation" using fuzzy logic. *Canadian Journal of Fisheries and Aquatic Sciences*, 56: 686–699.
- Michalewicz, Z. 1992. Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag, New York. 250 pp.
- Mikkelsen, N., and Pedersen, T. 2004. How can the stock recruitment relationship of the Barents Sea capelin (*Mallotus villosus*) be improved by incorporating biotic and abiotic factors? *Polar Research*, 23: 19–26.
- Myers, R. A. 1998. When do environment–recruitment correlations work? *Reviews in Fish Biology and Fisheries*, 8: 285–305.
- Myers, R. A., Bridson, J., and Barrowman, N. J. 1995. Summary of worldwide spawner and recruitment data. Canadian Technical Report in Fisheries and Aquatic Sciences, 2024. 327 pp.
- Mullin, M. M. 1993. Webs and Scales: Physical and Ecological Processes in Marine Fish Recruitment. University of Washington Press, Seattle. Washington Sea Grant Program. 131 pp.
- Needle, C. L. 2002. Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries*, 11: 95–111.
- Ottersen, G., and Loeng, H. 2000. Covariability in early growth and year-class strength of Barents Sea cod, haddock, and herring: the environmental link. *ICES Journal of Marine Science*, 57: 339–348.
- Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada*, 191: 382 pp.
- Sætre, R., Toresen, R., and Anker-Nilssen, T. 2002. Factors affecting recruitment variability of the Norwegian spring-spawning herring (*Clupea harengus* L.). *ICES Journal of Marine Science*, 59: 725–736.
- Saila, S. 1996. Guide to some computerized artificial intelligence methods. *In* Computer in Fisheries Research, pp. 8–40. Ed. by B. A. Megrey, and E. Moksness. Chapman and Hall, New York, NY. 254 pp.
- Sarle, W. S. 1994. Neural networks and statistical models. *Proceeding of the Nineteenth Annual SAS Users Group International Conference*. 24 pp.
- Specht, D. 1991. A general regression neural network. *IEEE Transactions on Neural Networks*, 6: 568–576.
- Stein, M., and Borovkov, V. A. 2004. Greenland cod (*Gadus morhua*): modeling recruitment variation during the second half of the 20th century. *Fisheries Oceanography*, 13: 111–120.
- Sutcliffe, Jr, W. H. 1972. Some relations of land drainage, nutrients, particulate material, and fish catch in two eastern Canadian bays. *Journal of the Fisheries Research Board of Canada*, 21: 661–680.
- Sutcliffe, Jr, W. H. 1973. Correlations between seasonal river discharge and local landings of American lobster (*Homarus americanus*) and Atlantic halibut (*Hippoglossus hippoglossus*) in the Gulf of St. Lawrence. *Journal of the Fisheries Research Board of Canada*, 30: 856–859.

- Swartzman, G., Silverman, E., and Williamson, N. 1995. Relating trends in walleye pollock (*Theragra chalcogramma*) abundance in the Bering Sea to environmental factors. *Canadian Journal of Fisheries and Aquatic Sciences*, 52: 369–380.
- Tereshchenko, V. V. 1996. Seasonal and year-to-year variations of temperature and salinity along the Kola meridian transect. ICES CM 1994/C:11.
- Toresen, R., and Østvedt, O. J. 2000. Variation in abundance of Norwegian spring-spawn herring (*Clupea harengus*, *Clupeidae*) throughout the 20th century and the influence of climatic fluctuations. *Fish and Fisheries*, 1: 231–256.
- Tyler, A. V. 1992. A context for recruitment correlations: why marine fisheries biologists should still look for them. *Fisheries Oceanography*, 1: 97–107.
- Voges, E., Gordo, A., Bartholomae, C. H., and Field, J. G. 2002. Estimating the probability of different levels of recruitment for Cape hake *Merluccius capensis* off Namibia using environmental indices. *Fisheries Research*, 58: 333–340.
- 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 Ludwig, D. 1981. Effects of measurement errors on the assessment of stock-recruitment relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, 38: 704–710.
- White, H. 1992. *Artificial Neural Networks: Approximation and Learning Theory*. Blackwell, Oxford, UK.
- Wooster, W. S., and Bailey, K. M. 1989. Recruitment of marine fishes revisited. *In* Effects of Ocean Variability on Recruitment and an Evaluation of Parameters used in Stock Assessment Models, pp. 153–159. Ed. by R. J. Beamish and G. A. McFarlane, Canadian Special Publication in Fisheries and Aquatic Sciences, 108.
- Yatsu, A., Watanabe, T., Ishida, M., Sugisaki, H., and Jacobson, L. 2005. Environmental effects on recruitment and productivity of Japanese sardine *Sardinops melanostictus* and chub mackerel *Scomber japonicus* with recommendations for management. *Fisheries Oceanography*, 14: 1–16.