Abstract—We present a method to integrate environmental time series into stock assessment models and to test the significance of correlations between population processes and the environmental time series. Parameters that relate the environmental time series to population processes are included in the stock assessment model, and likelihood ratio tests are used to determine if the parameters improve the fit to the data significantly. Two approaches are considered to integrate the environmental relationship. In the environmental model, the population dynamics process (e.g. recruitment) is proportional to the environmental variable, whereas in the environmental model with process error it is proportional to the environmental variable, but the model allows an additional temporal variation (process error) constrained by a log-normal distribution. The methods are tested by using simulation analysis and compared to the traditional method of correlating model estimates with environmental variables outside the estimation procedure. In the traditional method, the estimates of recruitment were provided by a model that allowed the recruitment only to have a temporal variation constrained by a log-normal distribution. We illustrate the methods by applying them to test the statistical significance of the correlation between sea-surface temperature (SST) and recruitment to the snapper (*Pagrus auratus*) stock in the Hauraki Gulf–Bay of Plenty, New Zealand. Simulation analyses indicated that the integrated approach with additional process error is superior to the traditional method of correlating model estimates with environmental variables outside the estimation procedure. The results suggest that, for the snapper stock, recruitment is positively correlated with SST at the time of spawning.

Manuscript accepted 20 September 2002.
Fish. Bull. 101:89–99 (2003).

A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example

Mark N. Maunder
George M. Watters
Inter-American Tropical Tuna Commission
Scripps Institution of Oceanography
8604 La Jolla Shores Drive
La Jolla, California 92037-1508
E-mail address (for M. N. Maunder): mmaunder@iattc.org

Identifying a clear relationship between an environmental variable and processes in the dynamics of the population (recruitment, natural mortality, growth) or the fishery (catchability) would allow improved estimation and prediction of model parameters and derived quantities. It is well known that the environment plays a large role in the population dynamics and catchability of fish stocks. Many researchers (Green, 1967; Joseph and Miller, 1989; Hinton and Nakano, 1996; Lehodey et al., 1997; Shepherd et al., 1984) have identified correlations between population processes and environmental factors, and others (Hunter, 1983; Bertignac et al., 1998; Lehodey et al., 1998) have suggested hypotheses for the underlying causes of these correlations. Incorporation of environmental time series into stock assessment models may provide additional information to help estimate model parameters, particularly when fishing observations (catch, effort, length-frequencies) are missing. For the management of fish stocks, it can be advantageous to be able to predict future catch rates and population sizes. Because there is often a delay due to the propagation of the recruitment signal in the population structure or because statistical and numerical models can provide predictions for some environmental variables (e.g. temperature) (or for both reasons), the relationship can be used to predict future catch rates or population sizes.

Statistical catch-at-age analysis (e.g. Fournier and Archibald, 1982; Deriso et al., 1985; Methot, 1990) is more appropriate than cohort analysis (virtual population analysis) to include relationships between an environmental variable and processes in the dynamics of the population. In cohort analysis, if there are missing data, they are simply extrapolated without any statistical methods, which may cause bias in the parameter estimates. Also, the potential correlation with an environmental series is calculated outside of the estimation procedure, producing several disadvantages, including the loss of information and the difficulty of propagating uncertainty (Maunder, 1998a, 2001a, 2001b). However, in statistical catch-at-age analysis, there are robust statistical methods (maximum likelihood, with all the parameters estimated together by obtaining the best fit between predicted and observed data) that allow inclusion of multiple data sets and the integration of the environmental series into the stock assessment model. These methods automatically allow for missing data and provide confidence intervals, and the hypotheses can be easily incorporated and tested.

The methods used to integrate the environmental series into the stock assessment model can be applied to different processes in the population, but are illustrated here with the case of recruitment. Recruitment is the fundamental process in the population dynamic that is responsible for the fluctuations of the stock size. Many studies (e.g. Francis, 1993) show that environmental variables affect the recruitment. In statistical catch-at-age analysis, recruitment combines an average value with an annual deviate, constrained by using a distributional assumption (e.g. Maunder and Starr, 2001). This constraint allows the estimation when there is no information (i.e. missing data). Traditional methods that relate recruitment to environmental factors use correlation analysis of an environmental time series with estimates of recruitment from a stock assessment model. For example, cohort analysis is first used to generate a time series of recruitment. Then the time series of recruitment is regressed against sea-surface temperature (SST). This two-step procedure has a number of disadvantages (Maunder, 1998a, 2001a, 2001b), including the loss of information and the difficulty of propagating uncertainty.

We introduce a method suggested by Maunder (1998a; see Maunder and Starr, 2001) that incorporates environmental time series into stock assessment models and tests the significance of the correlation between the population processes and the environmental time series. We test the model with simulated data and compare the results to correlating model estimates with environmental variables outside the estimation procedure. We illustrate this method with an application that investigates the correlation between SST and recruitment within the context of a statistical catch-at-age analyses used to assess snapper (*Pagrus auratus*) in the Hauraki Gulf–Bay of Plenty, New Zealand (Maunder and Starr, 2001).

Materials and methods

Integrating environmental indices into stock assessment models

Parameters that relate the environmental time series to population processes were included in the statistical catch-at-age stock assessment model. We added additional structure to the stock assessment model for each parameter of the stock assessment model (X) that was hypothesized to 1) have temporal variation, 2) be correlated with an environmental time series, and 3) have sufficient information in the data to be estimated for multiple time periods. This structure included a mean value for the stock assessment model parameter (μ), temporal deviations in the stock assessment model parameter (ε_t), a parameter that relates the environmental series to the stock assessment model parameter (β), and a scaling parameter (α) that ensures that μ is the mean value for the stock assessment model parameter over the time period used in the model.

$$X_t = \mu \exp(\alpha + \beta I_t + \varepsilon_t), \tag{1}$$

where t = time, and
I_t = the value of the environmental time series at time t.

The parameter α ensures that μ is equal to the mean over the whole time period (Gilbert[1]; see Maunder and Starr, 2001). Therefore, α removes the log normal bias and bias caused by an unnormalized environmental time series and is defined as

$$a = \ln\left(\frac{n}{\sum \exp(\varepsilon_t + \beta I_t)}\right), \tag{2}$$

where n is the number of time periods.

The additional structure requires that a set of parameters (ε_t) that are constrained by a distributional assumption and two free parameters (μ, β) be estimated. The distributional assumption (referred to as a "prior" in the following description and represented by the negative logarithm of the prior probability, see Eq. 3) is a prior on the degree of temporal variation in the stock assessment model parameter. The default assumption is a normal distribution (assuming that the stock assessment model parameter is lognormally distributed) with mean zero and given standard deviation. Information about this distribution can be obtained from estimates for similar species (e.g. Myers et al., 1995). The prior

$$-\ln \text{ Prior } (\varepsilon \mid \sigma) = \sum_t \frac{(\varepsilon_t)^2}{2\sigma^2} \tag{3}$$

keeps the temporal deviations close to zero if there is no information in the data to the contrary. It is important to note that the prior is also needed to avoid making β a redundant parameter.

The parameters μ and β and the set of parameters ε, are estimated simultaneously with the other parameters of the stock assessment model, and the negative logarithm of the prior is added to the negative log-likelihood function of the stock assessment model. The parameter estimates are really the mode of the posterior distribution, but we treat them in a likelihood context. The influence of the environmental time series can be removed from the analysis by fixing β at zero. Therefore, likelihood ratio tests can be used to determine if the β parameter significantly improves the fit to the data. If the addition of β reduces the total negative log likelihood by more than about 1.92 units ($\chi^2_{1,\alpha=0.05}$), then the additional parameter significantly improves the fit to the data at the 0.05 level, and there is a statistically significant correlation between the population process and the environmental time series. Similar tests can be performed to test the significance of the set of temporal deviation parameters, ε_t, by taking into account the number of additional parameters. Hilborn and Mangel (1997) provided a simple description of the likelihood ratio test. (The Akaike information criterion or the Bayes information criterion could also be used.) Therefore, by fixing, or not, β or ε at zero we can define three types of statistical models:

1 Traditionalmodel
β is fixed at zero, the parameter set ε_t is estimated, and a significant relationship is determined by testing if the correlation coefficient between ε_t and the environmental time series is significantly different from zero.

[1] Gilbert, D. J. 1999. Personal commun. National Institute of Water and Atmospheric Research Limited, P.O. Box 14-901, Wellington, New Zealand.

2 Environmental model
β is estimated, each value in the parameter set ε_t is fixed at zero, and a significant relationship is determined by testing if $\beta = 0$, using a likelihood ratio test.

3 Environmental model with process error
Both β and ε_t are estimated and a significant relationship is determined by testing if $\beta = 0$, with a likelihood ratio test.

Simulation testing

Simulation analysis was carried out to test the performance of the integrated approach and to compare this approach to the traditional model. A simple age-structured model (Appendix I) was set up to simulate a population for 20 years, starting from an unexploited population and generating catch, effort, and catch-at-age data. The simulated recruitment was generated as having a component based on an environmental time series and a random component. Each component was given the same variance (0.6^2). The environmental time series was randomly generated with $\beta = 1$ for each simulation. The standard deviation of the observation error in the CPUE index, σ_{CPUE}, was set at 0.6, and the sample size of the catch-at-age data was set to 50. The same age-structured model was then fitted to the data to estimate the model parameters. The three models defined in the previous section (traditional model, environmental model, and environmental model with process error) were tested with the simulated data. In addition to the parameters outlined in the description of the three models, average recruitment, the catchability coefficient, and the standard deviation of the fit to the CPUE data were also estimated. We also used a model that had constant recruitment to provide likelihood values to use in testing the significance of the environmental model.

The simulation analysis was repeated 500 times for four scenarios: 1) using catch-at-age data for all years, 2) using catch-at-age data for the first 10 years, 3) using catch-at-age data for the last 10 years, 4) using catch-at-age data for all years, but using $\beta = 0$ when generating the simulated data. Scenario 4 was used to investigate the probability of type-I error of the models when used in combination with the statistical tests. For each set of simulated data and each model, we determine how often a significant relationship between the logarithm of annual recruitment and the environmental time series is detected, the estimate of the slope of the relationship between the logarithm of annual recruitment and the environmental time series, the estimates of average recruitment, and the depletion level (ratio of current to unexploited biomass). We also calculate minimum-width 95% confidence intervals for average recruitment, using the likelihood profile method for the simulated data sets with catch-at-age data for all years.

Application: relating recruitment in the Hauraki Gulf–Bay of Plenty snapper stock to SST

Recruitment to the Hauraki Gulf snapper (*Pagrus auratus*) stock is correlated with temperature (Paul, 1976). The abundance of 1+ snapper in the Hauraki Gulf estimated by trawl surveys has been shown to have a positive correlation with SST (Francis, 1993) and air temperature (Gilbert, 1994) around or just after the time of spawning in the previous year. This relationship has also been shown with catch-at-age analysis to continue to hold as snapper enter the fishery at ages 4 and older (Maunder and Starr, 1998).

We applied the integrated approach described in this study in combination with the age-structured statistical catch-at-age model described in Maunder and Starr (2001) to the Hauraki Gulf–Bay of Plenty snapper stock. The model was fitted to catch-at-age data and biomass estimates. The biomass estimates were available for 1985 and 1994 and were obtained from analysis of tagging data. The majority of the catch-at-age data were available from 1990 to 1997, but there were some catch-at-age data of dubious quality, small sample size, and high variability for 1970 to 1973. The annual recruitment at age 1 was estimated for the time period of the model (1970–98) and also for 18 age classes (ages 2 to 19) that comprised the initial conditions in 1970.

Results

Simulation analysis

For all four sets of simulated data the environmental model had the highest probability of detecting a relationship between recruitment and the environmental time series (Table 1). This model had a very high probability of detecting a relationship even when there was no relationship in the simulated data (Table 1D). This indicates that the likelihood ratio test is not appropriate for the environmental model (see Appendix III). For all data sets, except that with only catch-at-age data in the first 10 years, the traditional model had a higher probability of detecting a relationship between recruitment and the environmental time series than did the environmental model with process error. The environmental model with process error had a lower probability of detecting a true relationship than the traditional model, but also had a slightly lower probability of type-I error (the probability of incorrectly accepting a nonexistent relationship) than the traditional model. The probability of detecting a relationship was reduced as the number of catch-at-age data sets was reduced.

The environmental model with process error did not show any bias in the estimate of the slope of the relationship between the logarithm of annual recruitment and the environmental time series, β (Table 1). For this model, the variation in the estimates of β increased when fewer years with catch-at-age data were available. The environmental model showed a small negative bias and slightly more error in the estimates of β. The traditional model showed a large negative bias in the estimate of the slope of the relationship between the logarithm of annual recruitment and the environmental time series, and this bias increased as the amount of catch-at-age data was reduced. The traditional method also had larger error, which increased as less catch-at-age data were available.

The errors in the estimates of average recruitment and B_{cur}/B_0 increased slightly with less catch-at-age data

Table 1

(**A**) Results from the simulation analysis in which all the catch-at-age data were used. % = percentage of data sets that produced a significant relationship between the environmental time series and recruitment; β = average (average absolute relative error) of the estimates of the slope of the relationship between the environmental time series and recruitment; R_0 = average (average absolute relative error) of the estimates of the average recruitment; B_{cur}/B_0 = average error (average absolute relative error) in the estimate of the ratio of current to unexploited biomass. β was set to 1 when generating the simulated data. (**B**) Results from the simulation analysis using only the first 10 years of catch-at-age data. β was set to 1 when generating the simulated data. (**C**) Results from the simulation analysis using only the last 10 years of catch-at-age data. β was set to 1 when generating the simulated data. (**D**) Results from the simulation analysis using all the catch-at-age data, but setting β = 0 when generating the simulated data. (Absolute, rather than relative, error was used for β.) EMwPE = environmental model with process error.

	%	β	R_0	B_{cur}/B_0
A				
Traditional	92	0.86 (0.25)	996 (0.05)	–0.01 (0.15)
Environmental	99	0.95 (0.25)	1,024 (0.10)	0.03 (0.24)
EMwPE	83	1.01 (0.23)	988 (0.05)	–0.04 (0.14)
B				
Traditional	57	0.45 (0.55)	1,011 (0.07)	–0.03 (0.17)
Environmental	95	0.94 (0.33)	1,078 (0.17)	0.01 (0.27)
EMwPE	60	1.05 (0.30)	1,004 (0.07)	–0.03 (0.15)
C				
Traditional	81	0.68 (0.36)	1,022 (0.09)	0.04 (0.21)
Environmental	97	0.95 (0.32)	1,035 (0.10)	0.04 (0.25)
EMwPE	74	1.00 (0.26)	1,007 (0.08)	–0.02 (0.19)
D				
Traditional	3	–0.01 (0.20)	1,005 (0.04)	0.01 (0.10)
Environmental	61	–0.02 (0.22)	1,038 (0.09)	0.07 (0.21)
EMwPE	2	–0.01 (0.21)	1,003 (0.04)	0.01 (0.10)

Table 2

Results related to the confidence intervals for average recruitment from the simulation analysis obtained by using all the catch-at-age data. EMwPE = Environmental model with process error.

	Lower bound average	Upper bound average	Lower bound SD	Upper bound SD	True is within	True is below	True is above
Traditional	888	1,120	59	132	0.91	0.03	0.08
Environmental	963	1,117	94	176	0.44	0.32	0.24
EMwPE	887	1,118	56	120	0.92	0.01	0.07

(Table 1). The errors in these estimates were slightly greater for the environmental model (more bias and larger absolute error) than for the environmental model with process error and traditional model.

The confidence intervals on R_0 were, on average, greater for the traditional and environmental model with process error than for to the environmental model (Table 2), which greatly underestimated the width of the confidence intervals. However, the confidence intervals for the traditional and environmental model with process error showed the true value falling below the confidence interval less often than it fell above it.

As expected, an environmental relationship was more difficult to correctly detect with the traditional model in situations with missing data (e.g. when catch-at-age data were missing in the last few years of the series), and, as stated above, using the environmental model is inappropriate because it has a high probability of detecting a significant relationship when none exists. The environmental model also has a tendency to under estimate the width of the confidence interval for R_0 and the true value frequently falls outside of this confidence interval. Therefore, the environmental model with process error is the model of choice.

Application: relating recruitment in the Hauraki Gulf–Bay of Plenty snapper stock to SST

The environmental model with process error has the lowest negative log-likelihood, but this model has many more parameters than the environmental model (Table 3).

Table 3

Results from applying the three methods (traditional, environmental, and environmental with process error) to the snapper application. Constant = recruitment is constant each year and equal to the average recruitment. EMwPE = environmental model with process error. n/a = not applicable.

	Average recruitment	β	–ln(Likelihood)	Number of parameters
Constant	13,315 (11,381–15,381)	n/a	482.5	3
Traditional	11,406 (8,500–14,603)	0.20	466.8	50
Environmental	13,530 (11,527–15,569)	0.48	473.5	4
EMwPE	12,029 (9,147–15,328)	0.55	464.4	51

Using the likelihood ratio test at the 0.05 level and compensating for the difference in the number of parameters being estimated in each model, we consider the environmental model to be the model of choice. If the likelihood ratio test were used, the environmental model with process error would be chosen over the traditional model, indicating a statistically significant correlation between SST and recruitment. Due to the weaknesses of the environmental model discussed above, we concentrated on the results of the traditional model and the environmental model with process error.

The time series of estimated recruitments from the traditional model showed very little annual variation in recruitment for the first half of the time series and for the last few years of the time series (Fig. 1A). This indicates that there is very little information in the data (catch-at-age) about annual recruitment for these time periods and that the prior on the recruitment residuals constrains the estimated recruitment to be close to the average recruitment. This result is consistent with the catch-at-age data, which, ignoring the inconsistent data from the 1970s, started in 1990. The greatest age in the catch-at-age data that had individual information was age 19; therefore the 1971 cohort is the first for which there is information. However, at the current exploitation rates, very few snapper live to be more than 10 years of age, so that there is very little information about cohort size for any of the cohorts produced during the 1970s.

The environmental model with process error indicated high variation in recruitment for the whole time period (Fig. 1B). This is due to the formulation of the recruitment submodel, for which the annual anomalies are anomalies from the temperature-recruitment relationship; if there is no information in the data about recruitment for a particular year, the recruitment will follow the temperature-recruitment relationship.

The correlation of the estimated recruitment from the traditional model with SST had a low r-square (0.26), but it was statistically significant at the 0.05 level when a two-tailed test was used. In addition, the slope of the relationship between recruitment and SST was much less for the traditional model than for to the environmental model with process error (Table 3). The estimates of recruitment from the traditional model included a large number of estimates that were close to the mean because there was no information in the data about these recruitments. Therefore, it was inappropriate to use these recruitments to correlate with SST and, if used, they would result in a poor fit. However, a significant correlation, as obtained in this application, suggests that the correlation is probably stronger than apparent from the analysis, which should give confidence that a relationship exists and provide an incentive to apply the integrated models.

The environmental model with process error did not show a statistically significant improvement over the environmental model because, ignoring the 1970s data, the catch-at-age data were available only for the last part of the time period. The recruitment anomalies were estimated for the whole time period, as well as for the initial conditions. Many of these recruitment anomalies had very

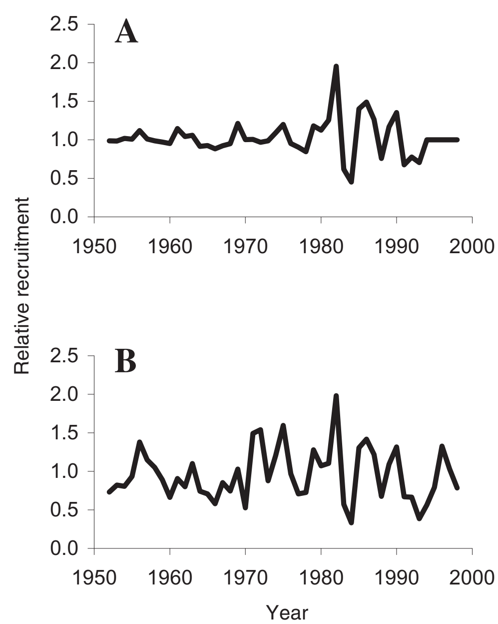

Figure 1

Annual estimates of relative recruitment strength at age 1 for the Hauraki Gulf–Bay of Plenty snapper stock from the traditional (**A**) and environmental model with process error (**B**) models little information associated with them and therefore did not add anything to the estimation procedure. However, they do add additional parameters, which reduce the possibility of accepting the model when using the likelihood ratio test. If the recruitment anomalies were estimated for only a limited number of years, it is likely that the environmental model with process error would be a statistically significant improvement over the environmental model. Statistical tests could be carried out to determine which annual recruitment anomalies should be estimated, but this would be very time consuming. Reducing the number of annual recruitment anomalies may also cause an underestimation of the confidence intervals. For the snapper example, removing the anomalies for the initial conditions may be a good compromise.

Discussion

We have developed a general framework for integrating environmental time series into stock assessment models that appears to perform better than traditional methods. The method is flexible and it can be used to model many different functional relationships between population or fishing processes and environmental time series and to include multiple environmental time series for any population model parameter (see Appendix II). Furthermore, it can be used with any statistical stock assessment model. The method can be used to test whether an environmental time series describes temporal variation in model parameters.

The traditional model, which estimates annual recruitment within the stock assessment model and subsequently correlates the recruitment with the environmental series outside the stock assessment model, performs poorly. It has a reasonable probability of detecting a relationship between recruitment and the environmental series, but this probability decreases rapidly as the number of years with missing catch-at-age data sets increases. The probability of incorrectly detecting a relationship when one is not present is low. This method has reasonable confidence-interval coverage for average recruitment and little bias or variance in the estimates of model parameters. The factor causing the poor performance of the traditional model is the large bias in the estimate of the slope of the relationship between recruitment and the environmental time series, which increases as the number of years with missing catch-at-age data increases. The bias occurs because the traditional model has a penalty on the absolute size of the annual recruitment deviations. This penalty constrains an annual recruitment anomaly to be close to the mean recruitment when there is little or no information about the recruitment in that year. Therefore, when the logarithm of the annual recruitment is correlated with the environmental time series, the estimated slope of the relationship is biased downward. Even in situations for which there is sufficient information for every recruitment anomaly, there will be a small tradeoff in the size of the anomaly, which reduces the contribution of the penalty to the objective function and the likelihood from the catch-at-age data. Unfortunately, if the penalty on the annual recruitment anomalies is removed, the estimation process can become unstable, particularly in data-poor situations for which the bias is greater. The amount of time that is required by the estimation algorithm also increases if the penalty is removed. When the penalty on the size of the recruitment anomalies is removed, the bias in estimates of the slope of the relationship between recruitment and the environmental time series is reduced when using all the catch-at-age data, but the variance in the estimates is greatly increased. In addition, when removing the penalty there was a large positive bias when using only the last 10 years of catch-at-age data and a large negative bias when using only the first 10 years of catch-at-age data. It is not known what results would be obtained if cohort analysis, which does not use a constraint on the annual recruitment anomalies, is used instead of the statistical catch-at-age analysis. It should be remembered that cohort analysis cannot be used or assumptions that are unlikely to be satisfied will have to be made when catch-at-age data are missing for some years.

The environmental model, which has a deterministic relationship between recruitment and the environmental time series that is integrated into the stock assessment model, also performs poorly. This method has poor confidence interval coverage for average recruitment because the size of the confidence intervals are greatly underestimated. The method has larger bias and variance in the estimates of model parameters compared to the other two methods. There is a small negative bias in the estimate of the slope of the relationship between recruitment and the environmental time series. The environmental model has a very high probability of detecting a relationship between recruitment and the environmental series, and this probability only decreases slightly as the number of missing years of catch-at-age data sets increases. However, this model has a very large probability of incorrectly detecting a relationship when one is not present. Therefore, when using the environmental model, the likelihood ratio test should not be used to determine if there is a significant relationship between recruitment and an environmental time series. The value used to compare to the χ^2 statistic in the likelihood ratio test for the environmental model is highly correlated with the catch-at-age sample size; therefore simulation analysis is needed to find the appropriate χ^2 statistic for the given sample size (see Appendix III). This is also important for calculating confidence intervals that are also based on the χ^2 statistic and is the reason for the poor coverage for R_0.

The environmental model with process error, which has a relationship between recruitment and the environmental time series that is integrated into the stock assessment model with additional process error, performs well. This model has a reasonable probability of detecting a relationship between recruitment and the environmental series, but this probability is lower than those of the other two models, and decreases as the amount of data is reduced. It has a low probability of incorrectly detecting a relationship when one is not present. These probabilities could be improved by using simulation analysis to find the appropriate χ^2 statistic (see Appendix III). This method has reasonable confidence interval coverage for average recruitment and has little bias or variance in the estimates of model parameters. There is very little bias in the estimate of the slope of the relationship between recruitment and the environmental time series.

For the environmental model with process error, when there is little or no information in the data to estimate the recruitment for that year, the penalty on the annual recruitment anomalies causes recruitment to be estimated close to the recruitment predicted by the relationship between recruitment and the environmental time series. Therefore, if there is a relationship between recruitment and the environmental time series, this model should provide better estimates because additional information is included in the estimation procedure. This model has the favorable property that if there is no relationship between recruitment and the environmental time series, the model estimates β to be small, eliminating any influence of the relationship between recruitment and the environmental time series, and still estimates the annual recruitment anomalies to represent the variation in annual recruitment. The likelihood ratio test can be used to detect a relationship between recruitment and the environmental time series, and if a relationship does not exist, the results with β fixed at zero can be used. However, including β in the estimation procedure, even when there was no relationship between recruitment and the environmental time series, did not increase the error in the parameter estimates in relation to the model with β fixed at zero (see the results for the traditional model, Table 1D).

The method we describe can be used to integrate environmental time series for parameters of the stock assessment model other than recruitment. The influence of the environment on catchability of the fish would be an obvious choice because there are numerous publications on the topic. For example, Green (1967) suggested that thermocline data would improve estimation of tuna abundance from catch and effort data, by allowing for the differentiation between changes in tuna abundance and catchability due to vertical distribution of tunas influenced by temperature. We have used a method similar to the method that is presented in the present study to incorporate SST into the purse-seine catchability parameters for yellowfin and bigeye tuna (Maunder and Watters, 2001; Watters and Maunder, 2001). Maunder (2001a) presented a general method to integrate the standardization of CPUE data into stock assessment models, including the integration of environmental variables. Growth rates have been observed to have temporal variation, and this variation has been correlated with environmental factors. Several authors have presented growth curves that include temperature data (e.g. Mallet et al., 1999). Movement is another process that may be influenced by the environment. Lehodey et al. (1997) showed that spatial shifts in the western Pacific skipjack tuna population are linked to the movement of a large pool of warm water and that the movements of this large pool are related to El Niño-Southern Oscillation events.

Once a correlation between the environmental time series and the population process has been determined, this relationship can be used to improve the predictive ability of the model. For example, if a relationship between SST at the time of spawning and recruitment has been determined, and the age at recruitment to the fishery is 3 years, recruitment to the fishery can be estimated 3 years in advance. One should be cautious about assuming that these relationships are valid and will continue to hold into the future, however. Hilborn and Walters (1992) cautioned about using environmental data because there are many environmental indices that one can try, and if the data set has a few large and a few small observations, it is likely that one of the environmental data sets will correlate with the data. Myers (1998) reviewed a number of published correlations between recruitment and environmental factors and found that few of the correlations held when retested at later dates. Maunder and Starr (1998) also advised caution because they found that a strong cohort may not enter the fishery when expected because of variations in growth rates. We have found that, when applying this method to the bigeye tuna data, there is an inconsistency in the pre-1997 data and the data for 1997 and 1998 caused by much stronger than expected year classes entering the fishery in 1997 and 1998. There is also difficulty in deciding on the management strategy if environmental regime shifts are influencing the productivity of the stock (Maunder, 1998b). An advantage of the integrated approach, particularly the environmental model with process error, is that it more fully describes the uncertainty in the relationship between the population process and the environmental time series, and therefore this uncertainty can be included in any management advice based on the relationship.

Conclusions

Integrating environmental relationships in a statistical stock assessment model is an improvement over the traditional statistical model when there are large gaps in the data. However, it is important to include process error to avoid the high probability of detecting spurious correlations seen in the environmental model when using the likelihood ratio test. Therefore, the environmental model with process error is the model of choice because 1) there is no bias in the estimates, 2) when there is no relationship with the environmental series, it is equivalent to the traditional model, 3) when such a relationship exists, the recruitment estimates are improved, particularly if there are important gaps in the data, 4) it may be used for prediction, and 5) uncertainty about the relationship can be modeled.

Acknowledgments

We thank Dave Fournier for advice on using AD Model Builder and related software, and Bill Bayliff, Rick Deriso, Shelton Harley, and Ransom Myers for commenting on the manuscript.

Literature cited

Bertignac, M., P. Lehodey, and J. Hampton.
1998. A spatial population dynamics simulation model of

tropical tunas using a habitat index based on environmental parameters. Fish. Oceanogr. 7:326–334.

Beverton, R. J. H., and S. J. Holt.
1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. II, Mar. Fish. G.B. Minist. Agric. Fish. Food 19, 533 p.

Deriso, R. B., T. J. Quinn II, and P. R. Neal.
1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815–824.

Forsbergh, E. D.
1989. The influence of some environmental variables on the apparent abundance of skipjack tuna, *Katsuwonus pelamis*, in the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. 19:433–569.

Fournier, D., and C. P. Archibald.
1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195–1207.

Francis, M. P.
1993. Does water temperature determine year class strength in New Zealand snapper (*Pagrus auratus*, Sparidae)? Fish. Oceanogr. 2:65–72.

Gilbert, D. J.
1994. A total catch history model for SNA 1. New Zealand Fisheries Research Document 94/24. 16 p. NIWA (National Institute of Water and Atmospheric Research), Wellington, New Zealand.

Granger, C. W. J.
1993. Forecasting in economics. *In* Time series prediction: forecasting the future and understanding the past (A. E. Weigend and N. A. Gershenfeld, eds.), p. 529–538. SFI Studies in the Sciences of Complexity, Proc. Vol. XV. Addison-Wesley, Boston, MA.

Green, R. E.
1967. Relationship of the thermocline to success of purse seining for tuna. Trans. Am. Fish. Soc. 96:126–130.

Hilborn, R., and M. Mangel.
1997. The ecological detective: confronting models with data, 315 p. Princeton Univ. Press, Princeton, NJ.

Hilborn, R., and C. J. Walters.
1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty, 570 p. Chapman and Hall, New York, NY.

Hinton, M. G., and H. Nakano.
1996. Standardizing catch and effort statistics using physiological, ecological, or behavioral constraints and environmental data, with an application to blue marlin (*Makaira nigricans*) catch and effort data from Japanese longline fisheries in the Pacific. Inter-Am. Trop. Tuna Comm. Bull. 21:171–200.

Hunter, J. R.
1983. On the determinants of stock abundance. *In* From year to year (W. S. Wooster, ed.), p. 11–16. Washington Sea Grant Program, Univ. Wash., Seattle, WA.

Joseph, J., and F. R. Miller.
1989. El Niño and the surface fishery for tunas in the eastern Pacific. Japan. Soc. Fish. Ocean., Bull. 53:77–80.

Lehodey, P., J. Andre, M. Bertignac, J. Hampton, A. Stones, C. Menkes, L. Memery, and N. Grima.
1998. Predicting skipjack tuna forage distributions in the equatorial Pacific using a coupled dynamical bio-geochemical model. Fish. Oceanogr. 7:317–325.

Lehodey, P., M. Bertignac, J. Hampton, A, Lewis, and J. Picaut.
1997. El Niño Southern Oscillation and tuna in the western Pacific. Nature 389:715–718.

Mallet, J. P., S. Charles, H. Persat, and P. Auger.
1999. Growth modelling in accordance with daily water temperature in European grayling (*Thymallus thymallus* L.). Can. J. Fish. Aquat. Sci. 56:994–1000.

Maunder, M. N.
1998a. Integration of tagging and population dynamics models in fisheries stock assessment. Ph.D. diss., 306 p. Univ. Washington, Seattle, WA.
1998b. Problems with using an environmental based recruitment index: examples from a New Zealand snapper (*Pagrus auratus*) assessment. *In* Fishery stock assessment models (F. Funk, T. J. Quinn II, J. Heifetz, J. N. Ianelli, J. E. Powers, J. J. Schweigert, P. J. Sullivan, and C. I. Zhang, eds.), p. 679–692. Alaska Sea Grant College Program Report No. AK-SG-98-01, Univ. Alaska, Fairbanks, AK.
2001a. A general framework for integrating the standardization of catch-per-unit-effort into stock assessment models. Can. J. Fish. Aquat. Sci. 58:795–803.
2001b. Integrated tagging and catch-at-age analysis (ITCAAN). *In* Spatial processes and management of fish populations (G. H. Kruse, N. Bez, A. Booth, M. W. Dorn, S. Hills, R. N. Lipcius, D. Pelletier, C. Roy, S. J. Smith, and D. Witherell, eds.), p. 123–146. Alaska Sea Grant College Program Report AK-SG-01-02. Univ. Alaska, Fairbanks, AK.

Maunder, M. N., and P. J. Starr.
1998. Validating the Hauraki Gulf snapper pre-recruit trawl surveys and temperature recruitment relationship using catch at age analysis with auxiliary information. New Zealand Fisheries Research Document 98/15, 23 p. NIWA (National Institute of Water and Atmospheric Research), Wellington, New Zealand.

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

Maunder, M. N., and G. M. Watters.
2001. Status of yellowfin tuna in the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Stock Assess. Rep. 1:5–86.

Methot, R. D.
1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. Inter. North Pacif. Fish. Comm. Bull. 50:259–277.

Myers, R. A.
1998. When do environment-recruitment correlations work? Reviews in fish biology and fisheries 8:285–305.

Myers, R. A., J. Bridson, and N. J. Barrowman.
1995. Summary of worldwide spawner and recruitment data. Can. Tech. Rep. Fish. Aquat. Sci. 2020, iv + 327 p.

Paul, L. J.
1976. A study on age, growth, and population structure of the snapper, *Chrysophrus auratus* (Forster), in the Hauraki Gulf, New Zealand. New Zealand Ministry of Fisheries and Agriculture, Fish. Bull. 13, 62 p.

Ricker, W. E.
1954. Stock and recuitment. J. Fish. Res. Board Can. 11: 559–623.

Shepherd, J. G., J. G. Pope, and R. D. Cousens.
1984. Variations in fish stocks and hypotheses concerning their links with climate. Rapp. P.-V. Reun., Cons. Int. Explor. Mer 185:255–267.

Watters, G. M., and M. N. Maunder.
2001. Status of bigeye tuna in the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Stock Assess. Rep. 1:109–210.

Appendix I: description of simulator and estimator

The following is a description of the model equations used for the data simulator and for the estimator. The model is run from an unexploited state at the start of the fishery for 20 years. The model includes 10 age classes, where the 10th age class is a plus group.

Dynamics

$$N_{y,1} = R_0 \exp(\beta I_y + \varepsilon_y^R + \alpha) \tag{I.1}$$

$$\alpha = \ln\left(\frac{n}{\sum \exp(\varepsilon_y^R + \beta I_y)}\right) \tag{I.2}$$

$$N_{y,a} = \left(N_{y-1,a-1}(1-u_{y-1}s_{a-1})\right)e^{-M_{a-1}} \text{ for } a < A \tag{I.3}$$

$$N_{y,A} = \left(N_{y-1,A-1}(1-u_{y-1}s_{A-1})\right)e^{-M_{A-1}} + \left(N_{y-1,A}(1-u_{y-1}s_A)\right)e^{-M_A} \tag{I.4}$$

$$u_y = \frac{C_y}{B_y} \tag{I.5}$$

$$B_y = \sum_a N_{y,a} s_a w_a \tag{I.6}$$

where Ny,a = the numbers in age class a at the beginning of year y;
R_0 = the average recruitment;
ε_y^R = the recruitment anomaly for year y;
σ_R^2 = the standard deviations for the recruitment anomalies;
M_a = the age specific natural mortality rate;
A = the maximum age used in the analysis;
u_y = the exploitation rate in year y;
s_a = the selectivity to the fishing gear for age a individuals;
C_y = the total catch in weight for year y;
B_y = the exploitable biomass for year y; and
w_a = the weight for an individual of age a.

Initial conditions

$$N_{1,a} = R_0 e^{\left(-\sum_{i=1}^{a-1} M_i\right)} \text{ for } 1 < a < A \tag{I.7}$$

$$N_{1,A} = \frac{N_{1,A-1}e^{(-M_{A-1})}}{1-e^{-M_A}} \tag{I.8}$$

Simulation

$$\varepsilon_y \sim N(0, \sigma_R^2) \tag{I.9}$$

$$I_a \sim N(0, \sigma_I^2) \tag{I.10}$$

$$\varepsilon_a^{CPUE} \sim N(0, \sigma^2_{CPUE}) \tag{I.11}$$

$$CPUE_y = qB_y \exp(\varepsilon_y^{CPUE} - 0.5\,\sigma^2_{CPUE}) \tag{I.12}$$

where q = the catchability coefficient for the CPUE index;
σ_I = the standard deviation for the variation in the recruitment index;
ε_a^{CPUE} = the observation error in the CPUE index; and
σ_{CPUE} = the standard deviation of the error in the CPUE index.

$$D_{y,a} \sim Multinomial\left(\frac{C_{y,a}^{numbers}}{\sum_a C_{y,a}^{numbers}}, n = 50\right) \tag{I.13}$$

$$C_{y,a}^{numbers} = N_a u_y s_a, \tag{I.14}$$

where $D_{y,a}$ = the number of individuals of age a in the catch-at-age same in year y; and
n = the number in the catch-at-age sample; and
σ_{CPUE} = the catch in numbers of age a individuals in year y.

Estimation

The likelihood values can be calculated by using the following equations:

$$-\ln L(\theta \mid I) = \sum_y \left[\ln(\sigma_{CPUE}) + \frac{\left(\ln(CPUE_y) - \ln(qB_y)\right)^2}{2\sigma^2_{CPUE}}\right] \tag{I.15}$$

$$-\ln L(\theta \mid D) = -\sum_{y,a} D_{y,a} \ln(p_{y,a}) \tag{I.16}$$

$$p_{y,a} = \frac{C_{y,a}^{numbers}}{\sum_a C_{y,a}^{numbers}} \tag{I.17}$$

The penalities (priors) on the annual recruitment anomalies can be calculated by using the following equation:

$$-\ln \text{ Prior } (\varepsilon^R) = \sum_y \left[\frac{(\varepsilon_y^R)^2}{2\sigma_R^2}\right]. \tag{I.18}$$

The following values were used for the biological parameters:

$$w_a = l_a^3, \quad \text{(I.19)}$$

$$l_a = 1 - e^{-0.1}, \quad \text{(I.20)}$$

and $M = 0.2$ and $s = 1$, where the a for M and s are subscripted.

Appendix II: extensions

The recruitment-environmental submodel that we used to analyze the snapper stock is simple, and other submodels may improve the fit to the data and the explanatory ability of the environmental time series. A dome-shaped relationship has been observed between abundance of tuna larvae and SST (Forsbergh, 1989), indicating that a quadratic or higher-order polynomial submodel may be more appropriate.

$$R_t = \exp(\alpha + \beta_1 I_t + \beta_2 I_t^2 + \ldots \beta_n I_t^n + \varepsilon_t). \quad \text{(II.1)}$$

Regime-switching models (Granger, 1993) that have the ability to favor two levels of values may be more appropriate for species that are hypothesized to experience two environmental regimes.

$$X_t = \left\{ (ub - lb)\left[1 + \exp\left(-\ln(19)\frac{I_t - I_{50}}{I_{95} - I_{50}}\right)\right]^{-1} + lb \right\} \exp(\varepsilon_t), \quad \text{(II.2)}$$

where lb = the lower bound of the model parameter (low regime);
ub = the upper bound of the model parameter (high regime);
I_{50} = the environmental time series value that gives a 50% influence; and
I_{95} = the environmental time series value that gives a 95% influence.

If I_{95} is only slightly higher than I_{50}, the model will have two regimes. Therefore, the model can be simplified by setting I_{95} as a small fixed value above I_{50}, allowing for the use of a regime-shifting model that requires estimation of only the lower bound, upper bound, and the value of the environmental time series at the point of change.

The method can be easily extended to include multiple environmental factors,

$$X_t = \mu \exp\left(\varepsilon_t + \sum_i \beta_i I_{i,t} + \alpha\right), \quad \text{(II.3)}$$

where i indexes the environmental factor.

The method we have used assumes that recruitment is independent of spawner biomass (i.e. we penalize the deviation from a mean recruitment modified by the relationship with the environmental time series). Maunder (1998a) suggested applying the method to stock-recruitment relationships, and the models described by Hilborn and Walters (1992, p. 285–287) could be used to integrate spawner-recruitment models and environmental time series into the stock assessment model.

$$R_t = f(S_t) \exp(\beta I_t + \varepsilon_t + \alpha), \quad \text{(II.4)}$$

where $f(S_t)$ = the function for the stock-recruitment relationship and
S_t = the spawning biomass at time t.

The equation for the Ricker (1954) and Beverton and Holt (1957) models would be

$$R_t = S_t \exp(a - bS_t) \exp(\beta I_t + \varepsilon_t + \alpha). \quad \text{Ricker} \quad \text{(II.5)}$$

$$R_t = \frac{aS_t}{b + S_t} \exp(\beta I_t + \varepsilon_t + \alpha), \quad \text{Beverton-Holt} \quad \text{(II.6)}$$

where a and b are parameters of the stock recruitment models.

Appendix III: the hypothesis test problem for the environmental model

Let, $\hat{p}_j = \dfrac{n_j}{N}$

where $N = \sum_j n_j$

is the sample size and n_j is the number from category j in the sample. The negative log-likelihood (ignoring constant) is

$$-\ln L = N \sum_j -\hat{p}_j \ln(p_j).$$

$$\chi^2 = 2(\ln L_1 - \ln L_0) = 2N\left(\sum_j \hat{p}_j \ln(p_{1,j}) - \sum_j \hat{p}_j \ln(p_{0,j})\right),$$

therefore χ^2 is proportional to N. LnL_1 (estimate β) has one more parameter than LnL_0 (β=0) and therefore will be at least slightly larger (two sets if independent random numbers always have a nonzero correlation). Therefore, there will be some value of N for which χ^2>3.84. Now, consider a simple example where

$$p_j = \frac{x_j}{\sum_j x_j} \text{ and } x_j = \mu \exp(\beta I_j + \varepsilon_j),$$

with the penalty $-\ln \text{ Prior }(\varepsilon \mid \sigma) = \sum_j \dfrac{(\varepsilon_j)^2}{2\sigma^2}$,

and σ is a constant. Consider two models: 1) $\varepsilon_j = 0$ and 2) estimate ε_j. For model 1, as N increases χ^2 increases in proportion to N, as explained above, because the penalty term is constant. However, for model 2, as N gets large, the relative size of the penalty compared to $-\ln L_0$ gets smaller and therefore the estimates of ε_j change so that p_i gets closer to $\hat{p}_i$. Therefore, for model 2, χ^2 does not increase proportionally with N.

An appropriate test for the environmental model would be to produce sets of random environmental indices that have the same variance and auto-correlation as the actual environmental index to determine the appropriate value of χ^2 that would give the desired type-I error. This test would overcome the sample size effect. The method could also be used to refine the test for the environmental model with process error.