

Near real-time forecasting of recruitment from larval surveys: application to Alaska pollock

Kevin M. Bailey^{1,*}, Tianyang Zhang², Kung-sik Chan², Steven M. Porter¹,
Annette B. Dougherty¹

¹Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, Washington 98115, USA

²Department of Statistics and Actuarial Sciences, University of Iowa, Iowa City, Iowa 52242, USA

ABSTRACT: The fishery for Alaska pollock *Theragra chalcogramma* in the western Gulf of Alaska is dependent on the strength of incoming year classes. Methods were developed to measure late-larval stage abundance and sizes at-sea and to use such data in a near real-time model to forecast recruitment about 4 yr in advance. Larval abundances were weighted for size- and temperature-dependent mortality to project the number attaining a critical size of 15 mm; this weighting reflects the greater survival value of larger larvae to the population. The method was applied to catches in the 2007 and 2008 larval surveys to show the feasibility of attaining near real-time estimates. The accuracy of the method to forecast recruitment was evaluated from historical data. Adjusting larval abundance for size- and temperature-dependent mortality improved the trend with recruitment, but the relationship was not significant over the entire time series (1979 to 2007) because of an increasing trend in predators of juvenile stages, decoupling the parameters. A threshold-based generalized additive recruitment forecast model was developed to account for autocorrelative structure caused by inter-cohort interactions and mortality-related factors occurring after the larval period. The model closely tracks past recruitment trends. Based on low larval abundances and high numbers of predators, low to moderate abundances of age-4 recruits were predicted for 2007 to 2011, which matched observed levels reasonably well.

KEY WORDS: Population dynamics · Fisheries · Environment · Ocean research · Regime shift · Ecological community

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

In the early 1900s, marine scientists studying fisheries were already thinking about forecasting abundance into the near future, but at that time such forecasts were considered premature (Hjort 1914). By 1928, Oscar Sette was making formal predictions of the Atlantic mackerel fishery based on juvenile abundances of incoming year classes (Sette 1943, Smith 1994). More recent studies have shown that information on environmental factors or abundances of early life stages can be useful for predicting recruitment levels (Van der Veer 1986, Meekan et al. 1993, Helle et al. 2000). Other studies point out the

difficulties involved in forecasting recruitment from environmental conditions (Walters & Collie 1988, Houde 2008) or from abundances of early life stages (Mukhina et al. 2003). For example, in a meta-analysis of marine fish populations, Bradford (1992) found that egg and larval abundances were imprecise as predictors of future recruitment, whereas juvenile abundances performed better. Walters & Collie (1988) found that most statistical recruitment–environment relationships eventually fail. Nevertheless, since most fish stocks depend on the strength of incoming year classes, advanced recruitment predictions would be useful, especially for species with a short lag between birth and recruitment, or stocks

*Email: kevin.bailey@noaa.gov

where harvests increasingly rely on young fish due to depletion of older age classes (Hsieh et al. 2010).

Ichthyoplankton surveys have been conducted by the Alaska Fisheries Science Center (AFSC) during the late May to early June period in the western Gulf of Alaska (GOA) since 1979, and at-sea counts of late larval abundance of walleye pollock *Theragra chalcogramma* have been made in the surveys since 1988. In the early years, the counts were used to find patches of larvae for studies of larval mortality. In recent years, the at-sea counts of larvae from the surveys have been utilized as a factor to forecast recruitment in the annual GOA pollock stock assessments (Dorn et al. 2009). These forecasting attempts have been marginally successful, and may have been limited by the inclusion of larvae of all sizes in the counts. Adjustments to the late larval survey abundances seem appropriate because larger and older larvae are of greater value to the population than smaller larvae since they have already run the gauntlet of the high larval mortality stanza. Therefore, abundances should be weighted to account for the size composition of the population. Until recently, these corrections have not been possible to make in a timely manner because size information on the larvae has not been available until processing of samples is completed, usually about 2 yr later. In 2007, we implemented methods to measure larvae at-sea and formulated a near real-time index of late larval abundance based on the at-sea counts of larvae, a weighting scheme to account for the larval size structure, and empirical relationships between sea surface temperature and larval mortality. In 2008, the index of larval abundance was estimated before the ship returned to port, a day after the last station was occupied.

Although larval supply is critical in the recruitment process, what happens to fishes during the juvenile stage is also important. A voracious predator of juvenile pollock, the arrowtooth flounder *Atheresthes stomias* (ATF), began to increase in abundance in the GOA after a climate regime shift in 1976 to 1977, coinciding with a dramatic shift in the fish community (Anderson & Piatt 1999). By the early 1990s, the biomass of ATF surpassed that of pollock as the major groundfish species (Bailey 2000). Since ATF is a major predator on juvenile pollock, their role in top-down control of pollock recruitment merits consideration in a scheme of recruitment forecasting (Zhang et al. 2010).

We demonstrate the utility of mortality- and size-adjusted larval counts based on analysis of historical survey and recruitment data. The annual adjusted

larval abundance was used as an independent covariate in a unique forecasting model, which attempts to account for what may happen to fish after the larval stage. We demonstrate that it is important to take into account the size structure of the larval population, and also changes in the composition and distribution of predators (such as ATF) and autocorrelative structure in the pattern of dynamics when considering the relationship of recruitment and larval abundance. We used the model, which is based on a threshold generalized additive model (TGAM; Ciannelli et al. 2007), to make preliminary forecasts for the number of age-4 recruits in 2007 to 2011 and point out several improvements and factors that may influence forecasting success.

MATERIALS AND METHODS

Data collection

Pollock spawn in the Shelikof Strait region of the western GOA, mostly from late March to late April. At ambient temperatures, eggs take about 2 wk to hatch. The late larval survey has been conducted from late May to early June each year in the Strait and downstream waters since 1979 (Fig. 1). Sampling was usually done with bongo gear equipped with a 500 μm mesh plankton netting from 100 m to the surface. When the gear was retrieved, one codend was preserved in formalin and the other was subsampled for length measurements and abundance of pollock larvae at-sea. Abundance was calculated from the mean number per 10 m^2 per station within the survey grid, multiplied by the number of 10 m^2 units in the grid. A more thorough discussion of the surveys and sampling can be found in Bailey et al. (1996).

From 2007, larvae were measured for standard lengths at-sea. Larvae were removed from the live codend and sorted over ice. They were photographed immediately after capture using a digital camera attached to a dissecting microscope. Lengths were determined from the photographs using a calibrated digital image-processing program. A length frequency was determined, sorting larvae into 1 mm bins. For the length measurements, a random subsample of up to 20 larvae at each station was measured. Larvae smaller than 5 mm were assumed to have shrunk significantly, a problem that is more acute for small larvae (Porter et al. 2001), and were conservatively assigned to the smallest size bin.

As the raw abundance does not reflect the age structure of the population and older larvae have

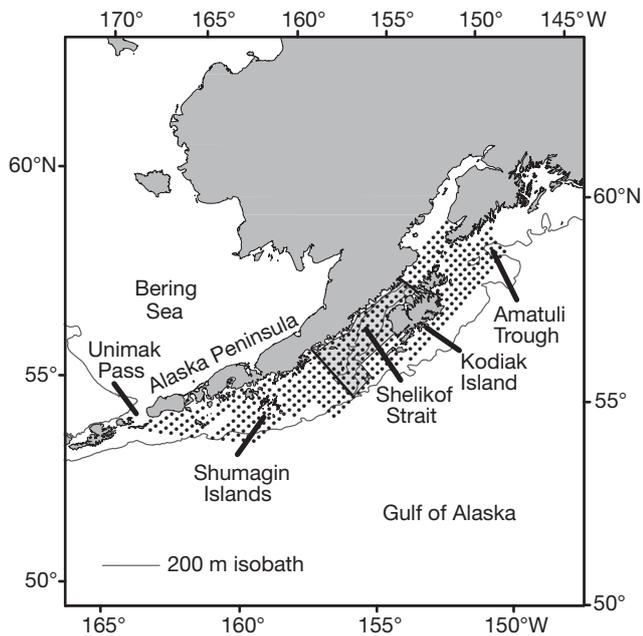


Fig. 1. Western Gulf of Alaska. The main survey grid region is shown by the shaded region. The main current in the sampling region is the Alaska Coastal Current which flows to the southwest along the peninsula

already passed through a critical period of high mortality, a weighting system for size and variability in the timing of the survey to give relative value to a larva's potential for recruitment was developed. This weighting system is an algorithm to estimate the number of larvae that will pass through a critical size of 15 mm based on mean historical growth and mortality rates (Bailey & Spring 1992). The mean growth rate was from a growth model (Schnute 1981) incorporating age (t in d) and length (Y in mm) data from 1983 to 2001 in the Shelikof region of the GOA (Dougherty et al. 2007):

$$Y(t) = \left([4.92^{-1.72} + (85.36^{-1.72} - 4.92^{-1.72})] \left[\frac{1 - \exp(-0.039(t))}{1 - \exp(-0.039(160))} \right] \right)^{1/1.72}$$

We compared 2 indices of larval abundance from the late larval surveys:

Index 1 is the abundance of larvae in the survey unadjusted for size or mortality.

Index 2 is the number of larvae estimated to reach the critical size of 15 mm after accounting for future growth and temperature-dependent mortality based on the size composition.

A mean mortality schedule was taken from Bailey et al. (1996). Temperature is an important factor in mortality (Bailey et al. 1996) and hence another mortality schedule for Index 2 was developed to account

for potential effect of temperature on recruitment (Table A1 in Appendix 1). From information presented in Bailey et al. (1996), we made an adjustment whereby for temperatures above the historical median, mortality was 0.5 times the average value, and for temperatures below the median, the mortality was 1.5 times the average value. We treated the mortality-temperature relationship nonparametrically by aggregating the data into above and below median values, as the data are too few for certainty about the linearity of the relationship or whether the variability is normally distributed. Historical average temperatures for May in the region were obtained from the NCEP Reanalysis data set provided by A. Macklin (Pacific Marine Environmental Laboratory, NOAA, pers. comm.; www.cdc.noaa.gov/cdc/reanalysis/). Mean sea surface temperature was interpolated across a longitude band in the GOA from 155.6° to 157.5° W centered at 56.2° N. For 2007 and 2008, CTD casts were taken at each station and the mean temperature for late May was calculated.

Data on pollock larval abundance were unavailable in 1980 and 1984. We computed the missing values as the product of the age-1 abundance times the geometric mean ratio of late larval abundance to age-1 abundance, with the latter estimated with the complete data set from 1982 to 2005.

Since larvae in the historical database were preserved in formalin, their standard lengths (SL) were corrected to a mean live SL based on the relation (Porter et al. 2001):

$$\text{Live SL} = 1.88 + 0.81 \times \text{Preserved SL}$$

We used TGAMs to examine the relationship between recruitment and environmental and biological factors, and to forecast future recruitment levels from the late larval abundances. The threshold structure is applied to capture the change of predation effects on pollock juveniles, and address the inverse trends of larval abundance with recruitment before and after the early 1990s. The abundance of age-4 recruits was obtained from Dorn et al. (2009), who estimated it from age-structured assessment models using catch-at-age data in the fishery. Pollock are caught as adults beginning at 4 to 5 yr of age. The age-4 recruits from 1983 to 2006, corresponding to the year classes from 1979 to 2002, were used to estimate the recruitment forecast model. The updated recruitment data in the following years from 2007 to 2009 (year classes from 2003 to 2005) were applied to evaluate the forecasting precision of the model. Abundance data were log-transformed to normalize the distribution and reduce conditional heteroscedasticity.

Potentially important environmental factors that may affect the survival of a late larval cohort to age-4 include the ATF abundance, drift out of Shelikof Strait (Transport), sea surface temperature (T), surface wind speed in Shelikof Strait (W), and the Pacific Decadal Oscillation (PDO). Historical ATF abundance was obtained from Turnock & Wilderbuer (2007). Although an exact estimate of ATF biomass was not available at the time of the larval survey, reasonable estimates can be obtained for the purpose of forecasting based on values from the previous year because of the low inter-annual variability in the trend trajectories and the high autocorrelation. Sea surface temperature data were interpolated across a longitudinal band from 155.6 to 157.5° W centered at 56.2° N. The wind speed covariates were computed from sea level pressure data collected twice per day in Shelikof Strait (56° N, 156° W). Transport was calculated from 'line 8', a transect of stations across Shelikof Strait. These data provided by A. Macklin and M. Spillane, Pacific Marine Environmental Laboratory. PDO was derived from surface temperature anomalies in the North Pacific Ocean, poleward of 20° N (jisao.washington.edu/pdo). Sea surface temperature and PDO are monthly data. Based on the values for daily Transport and daily W , we calculated their monthly averages respectively. To resolve which period within a year a particular environmental covariate affects survival, we computed the 12 monthly means, 5 seasonal means and yearly means for each covariate, thereby creating 18 variables per environmental factor. Later, we adopted model selection criteria to determine which of these mean covariates enter into the recruitment model. For example, the 12 monthly average sea surface temperatures were denoted by $T.Jan$ to $T.Dec$; the 5 seasonal averages were the mean temperature from January to March (JFM as the extension to the variable name), from April to June (AMJ), from July to September (JAS), from January to April (JFMA) and from May to June (MJ); and the annual mean was denoted by Year. The environmental data used for the recruitment model estimation and forecasting are available from 1979 to 2007.

Model development

As a strategy for developing a recruitment forecast, we started with the monitoring of real-time conditions, the adjusted larval abundance estimate, and examined the potential contribution of factors related to survival processes after the late larval stage. Since a near real-

time forecast was desired, a criterion for inclusion in the analysis was that reasonable estimates of the factors would be readily available at the time of the late larval survey. In the framework of a general model, environmental and predation factors in the same year of the late larval survey were used for the recruitment forecast. Significance levels were set at 5%.

We formulated modeled recruitment as the following general form:

$$y_{4,t+4} = \beta_0 + \beta_1 l_t + \beta_2 W_t + s(T_t) a_t 1_{(t>t_c)} + \xi_{t+4} \quad (1)$$

where the subscript t denotes year, $y_{4,t+4}$ denotes the log recruitment level of age-4 pollock spawned in year t and observed in year $t+4$, l_t denotes the late larval abundance of pollock transformed with the natural logarithm, and $y_{4,t+4}$ and l_t belong to the same spawning year class. Here, β_i ($i = 0, 1, 2$) are unknown coefficients. Additionally, W_t is the wind speed, T_t denotes the sea surface temperature, a_t the log ATF biomass, and $s(T_t)$ is a smooth function of T_t . Also, t_c is the threshold year to be estimated from the data, and the variable $1_{(t>t_c)}$ equals 1 in the years after t_c and zero otherwise. The term $s(T_t) a_t 1_{(t>t_c)}$ represents the threshold ATF interaction with temperature, where $s(T_t)$ is expected to be negative and its magnitude can be interpreted as the percent of juvenile pollock consumed per unit percent increase in ATF biomass at temperature T_t . With the indicator $1_{(t>t_c)}$, the ATF effect is restricted to be zero before the threshold year t_c . Furthermore, the structure of Model (1) specifies that the late larval abundance (l_t) and wind speed (W_t) are linearly correlated with the recruitment level ($y_{4,t+4}$), which is supported by preliminary data analysis. The error terms $\{\xi_i\}$ are of zero mean and finite variance, but they may be correlated over time with the auto-correlation structure modeled as some autoregressive (AR) process, e.g. AR(2). More specifically, $\xi_t = \phi_1 \xi_{t-1} + \phi_2 \xi_{t-2} + \varepsilon_t$, where $\{\varepsilon_i\}$ is a sequence of independent and identically distributed random variables of zero mean and finite variance, and ϕ_1 and ϕ_2 are the autoregressive parameters. Furthermore, the structure of Model (1) specifies that the late larval abundance (l_t) and wind speed (W_t) are linearly correlated with the recruitment level ($y_{4,t+4}$), which is supported by preliminary data analysis.

The mean structure of Model (1) contains a smooth function with unknown functional form, $s(T_t)$, while the regression errors are autocorrelated. Such a model structure can be fit in the framework of the generalized additive mixed models (GAMMs) (Lin & Zhang 1999, Wood 2006). Model (1) can be estimated by the method of penalized log-likelihood. The penalized log-likelihood is obtained by adding to the

log-likelihood function a roughness penalty term on the smooth function $s(\cdot)$, so that maximizing the penalized log-likelihood function allows a model to be found with good fit to the data and yet with a smooth function estimate. In practice, the penalty can have the form of the integrated squared second derivative of the smooth function $\lambda \int \{s(\cdot)\}^2$, where the non-negative smoothing parameter λ works as a tuning parameter to control the trade-off between model fit and smoothness of the function estimate. All models reported herein were fitted in R using functions from the `mgcv` library (Wood 2006).

Model (1) contains 2 environmental covariates, W_t and T_t . For each of them, there are 18 choices representing the monthly, seasonal and yearly average levels respectively, as discussed in the data subsection. We employed Akaike's information criterion (AIC) in selecting the combination of average levels for the environmental factors. We fitted Model (1) with each of the 18×18 combinations of W_t and T_t , with the threshold year searched from 1978 to 2000, and compared their AIC values. Models with lower AIC are preferred as they achieve a better balance between model fit and parsimony. Diagnostic checks on selected models were performed. If the model passed the diagnostic check, its choice of the environmental factors and threshold year was accepted; otherwise, the model with the next best AIC value was considered. The above model selection approach was continued until the best model satisfying all the model assumptions was found.

Forecasting

Age-4 recruitment forecasts were obtained by bootstrapping. The forecasts for the age-4 recruitment after 2006 were computed by drawing a future realization of the age-4 recruitments for 2007 to 2011 using a wild bootstrap (Wu 1986). The k -step ahead future recruitment point forecasts, $k = 1, \dots, 4$, were obtained by computing the median of the k -step ahead recruitment levels from the bootstrap samples; their associated prediction intervals comprised the 2.5 and 97.5 percentiles of the k -step ahead future recruitment. In the bootstrap approach, we simulated the future error terms by re-sampling the residuals from the fitted model. Such a bootstrap sampling approach provided valid predictions and associated prediction confidence intervals, even with an unknown error distribution.

A wild bootstrap method was applied in the recruitment forecast to assure that the simulated

white noise errors were symmetrical around zero. Since the error terms in the best recruitment model followed an AR(2) process, the bootstrap realizations of future recruitments were computed by the following equations:

$$\begin{aligned} \hat{Y}_{n+4+k}^* &= \\ &\hat{\beta}_0 + \hat{\beta}_1 I_{n+k} + \hat{\beta}_2 W.May_{n+k} + \hat{s}(T.Dec_{n+k}) a_{n+k} 1_{(n+k > \hat{t}_c)} + \hat{\xi}_{n+4+k}^* \\ \hat{\xi}_{n+4+k}^* &= \hat{\phi}_1 \hat{\xi}_{n+3+k}^* + \hat{\phi}_2 \hat{\xi}_{n+2+k}^* + \hat{\varepsilon}_{n+4+k}^*, \text{ for } k = 1, \dots, 4, \end{aligned}$$

where $\{\hat{\varepsilon}_{n+4+k}^*\}$ were randomly sampled from the residuals $\{\hat{\varepsilon}_t\}$ of Model (2) but with their signs randomly and independently flipped with probability $1/2$. Randomly flipping the signs of the residuals ensured that the bootstrap white noise had a distribution symmetric about zero, which is an advantage of the wild bootstrap. The error terms $\hat{\xi}_t^*$ were generated dynamically by the above equation for $t > n+4$, while $\hat{\xi}_t^* = \hat{\xi}_t = \hat{y}_t - [\hat{\beta}_0 + \hat{\beta}_1 I_{t-4} + \hat{\beta}_2 W.May_{t-4} + \hat{s}(T.Dec_{t-4}) a_{t-4} 1_{(t-4 > \hat{t}_c)}]$ are the regression residuals from Model (2) for $t \leq n+4$.

The above bootstrap procedure was repeated B times, with $B = 10\,000$ in our application, to get B replicates of $\{\hat{y}_{n+4+k}^*, k = 1, \dots, 4\}$. The median trajectory of these B future realizations provided the point predictions, with their associated 95% prediction bands enclosed by the 2.5 and 97.5 percentiles of the bootstrap future realizations.

RESULTS

Accuracy of at-sea larvae counts

Rough counts made at-sea were compared to verified counts from the laboratory from 9 cruises conducted in 1999 to 2005. This comparison indicated that the at-sea counts were proportional to the laboratory counts at a near 1:0.85 ratio, and 90% of the variability in the laboratory counts was accounted for by the at-sea counts (Fig. A1 in Appendix 1).

Larval abundance projections

Over the complete 1979 to 2003 time series, the linear relationship between recruitment and both the unadjusted and adjusted larval abundances was not significant ($r = 0.13$, $p = 0.54$). However, the adjusted value had a positive slope (as opposed to the unadjusted series which had a negative slope), indicating that it improved the relationship ($r = 0.23$, $p = 0.28$). We did not incorporate the esti-

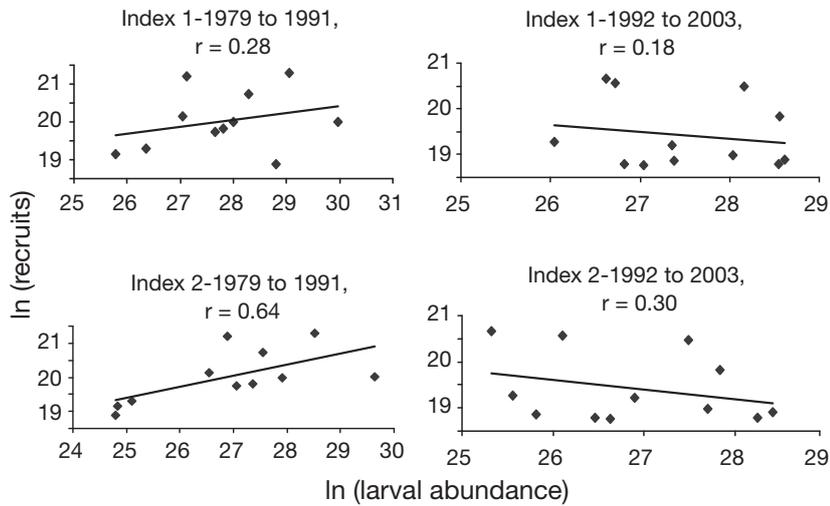


Fig. 2. *Theragra chalcogramma*. Larval abundance indices versus recruitment from 1979 to 1991 (left panels) and 1992 to 2003 (right panels). Upper and lower panels show, respectively, either no adjustment to counts (Index 1) or adjustment to counts weighting for larval size and temperature-dependent mortality (Index 2)

mates of the most recent year classes because they are preliminary, and they tend to shift in the age-structured models as more data on the year class accumulates.

The time series was deconstructed into 2 segments based on previous observations (Bailey 2000, Cianelli et al. 2005), indicating that the factors influencing recruitment have changed since the late 1980s to early 1990s. Consequently, in the period 1979 to 1991, there was a positive but statistically insignificant trend between unadjusted late larval numbers and recruitment (Fig. 2, Index 1 upper-left panel; $r = 0.28$, $p = 0.40$). Adjusting for size structure and temperature-dependent mortality (Index 2, lower-left panel) improved the relationship significantly ($r = 0.64$, $p = 0.03$). However during 1992 to 2003, the relationship changed, with both indices of larval abundance (Fig. 2, right upper and lower panels; $r = 0.18$, $p = 0.58$, and $r = 0.30$, $p = 0.35$, respectively), showing slight, but statistically insignificant, inverse trends.

Exploratory forecasting analysis

There was an obvious trend change in the relationship between adjusted late larval abundance and the age-4 recruitment (bottom 2 plots in Fig. 2). Age-4 recruitment was positively correlated to late larval abundance for the year classes before 1991, while such a relationship was much weaker after 1991. It suggests a change in the recruitment process starting

some time from the late 1980s to early 1990s. Indeed, 2 factors that may have brought about the aforementioned change were the increasing ATF biomass and the increasing overlap of the ATF and the juvenile pollock nursery areas since the late 1980s. In turn, the extent of overlap between the predators (ATF) and their prey (early juvenile-stage pollock) may be associated with bottom temperature. Sea surface temperature was used as a proxy for the bottom temperature (BT), assuming a mixed water column, because BT was not available for most years. Additionally, preliminary analysis also suggests that wind speed is significant for the pollock recruitment model, while the pollock biomass, Transport and PDO indices are not.

The fitted model includes late larval abundance, wind speed and an interaction term including ATF biomass and sea surface temperature. Moreover, it was found that the ATF predation effects took place after a threshold year, probably due to the fact that the spatial overlap between ATF and early-stage pollock was significant only from some time in the late 1980s or early 1990s.

In model development, the solid curve in Fig. 3 shows the profile of the best AIC among all the combinations of monthly, seasonal and yearly W_t and T_t discussed in 'Materials and methods', as a function of the candidate threshold year, and the dashed curve

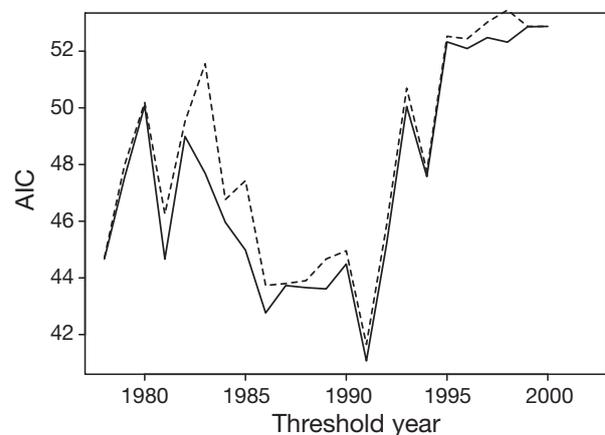


Fig. 3. Akaike's information criterion (AIC) profiles with different threshold years in Model (1). Solid curve: the best AIC among the environmental combinations at each threshold year; dashed curve: the second best AIC among the environmental combinations at each threshold year

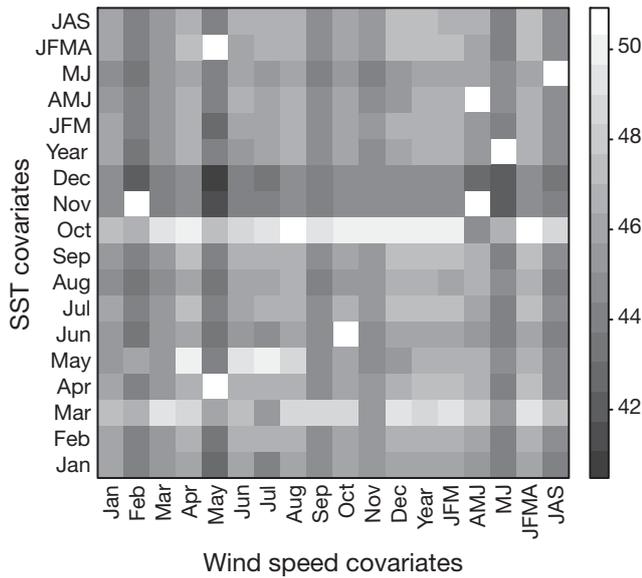


Fig. 4. Akaike's information criterion (AIC) heat map of Model (1) with various combinations of the environmental factors when the threshold year is 1988. Lower AIC is darker, and higher AIC is brighter; the white color represents a not applicable (NA) value in the AIC. SST is sea surface temperature. The abbreviations along the axes refer to the months used to determine the seasonal averages (see 'Materials and methods')

in Fig. 3 represents the second best AIC profile. These 2 curves show the best 2 AIC values obtained from all the choices of the environmental factors and threshold year, and these 2 choices comprise the first pool for the model selection discussed above. Additionally, since the lowest overall AIC is from a model with threshold year 1991 (Fig. 3; supporting the findings in Fig. 2), we studied the heat map of AIC for the combinations of environmental factors with that particular threshold year (Fig. 4). The darkest square in Fig. 4 suggests that Model (1) with the environmental covariates $W.May_t$ and $T.Dec_t$ has the lowest AIC. Since this model successfully passed the diagnostics check, it became the final fitted model:

$$Y_{4,t+4} = \beta_0 + \beta_1 I_t + \beta_2 W.May_t + s(T.Dec_t) a_t 1_{(t>t_c)} + \xi_{t+4} \quad (2)$$

Furthermore, the 4 best environmental combinations in Fig. 4 are related to the wind speed in May, June or both. Thus, there is evidence that the wind speed in late spring or early summer is correlated to the age-4 recruitment from their late larvae. We interpret this to mean that winds in May to June influence the future survival of larvae sampled then, either due to displacement by wind-driven currents or by altering feeding conditions.

As discussed in 'Materials and methods', Model (2) has a better fit to the data than the other models in the framework of Model (1) with a similar model structure but different choices of the environmental factors (Table 1). Additionally, Model (2) also shows its advantage over the other recruitment model structures according to the model selection criterion. First, an AR(2) error process in Model (2) provides a better fit with lower AIC (41.1) than other autoregressive structures in $\{\xi_{tj}\}$ (Table 2). Second, through comparing the fitted results from Model (2) and its constrained alternatives (Table 1), we found that all terms in Model (2) are significant (at 5%). The first alternative model for comparison does not take into account the ATF predation effect, and has the following form:

$$Y_{4,t+4} = \beta_0 + \beta_1 I_t + \beta_2 W.May_t + \beta_3 T.Dec_t + \xi_{t+4} \quad (3)$$

Strictly speaking, Model (3) is not a constrained version of Model (2), since it includes a linear temperature effect denoted by $\beta_3 T.Dec_t$. The second constrained model does not include any environmental factors, i.e. β_2 in Model (2) equals zero and $s(T.Dec_t)$ in the predation effect is restricted to be a slope parameter. The third constrained model only considers the larval abundance effect, but retains the AR(2)

Table 1. Comparison of Model (2) with its constrained models. AIC is Akaike's information criterion, edf is the effective degree of freedom and ATF is arrowtooth flounder

Fitted model	AIC	Adjusted R ² (%)	edf
Model (2)	41.1	74.0	8
No ATF predation effect	51.8	55.3	6
No environmental effects	41.3	69.4	6
With only late larval abundance, plus autocorrelated errors	52.0	51.6	4
With only larval abundance, and no autocorrelation in the errors	64.5	6.5	2

Table 2. Model comparison with various stochastic error processes in Model (2), where edf stands for the effective degree of freedom. Since lower Akaike's information criterion (AIC) levels are preferred in model selection, the model with autoregressive AR(2) error process has the best fit among 4 potential models

Autocorrelation structure	No autoregressive	Error process		
		AR(1)	AR(2)	AR(3)
AIC	63.3	64.3	41.1	41.9
edf	6	7	8	9

Table 3. Estimates of Model (2)

Parameter	Value	SE	t-value	p-value
β_0	17.3	1.43	12.1	<0.001
β_1	0.140	0.041	3.38	0.0038
β_2	-0.304	0.153	-1.99	0.0643
ϕ_1	0.519			
ϕ_2	-0.871			

structure in the error terms. The fourth model is only a simple linear regression on the (log) late larval abundance, without any autocorrelation structure in the errors. Overall, Model (2) has a better AIC (41.1) and adjusted R^2 (74 %) than the other 4 simplified recruitment models.

From the estimated $\hat{\beta}_1$ (with SE 0.041) in Table 3, the late larval abundance is positively related to the recruitment of age-4 pollock. The association between the late larval abundance and age-4 recruitment is weaker than the one between the age-1 pollock and their age-4 recruits (Zhang et al. 2010). This suggests that the survival of late larvae into the age-1 stage introduces more uncertainty into the recruitment process of age-4 pollock. This confirms the expectation that the recruitment forecast from late larval abundance will be more challenging than the forecast from age-1 abundance. The slope of the linear wind effect is -0.304 (with SE 0.153), which indicates that strong winds impede the recruitment of pollock.

The trend change in the relationship between adjusted late larval abundance and recruitment is explained by the threshold ATF predation in Model (2). The threshold year was determined from the

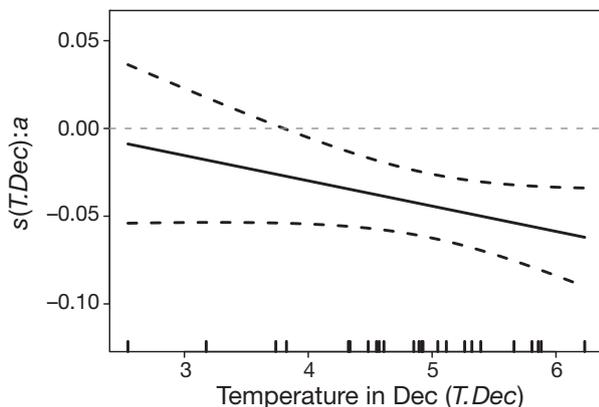


Fig. 5. Temperature interaction with the arrowtooth flounder (ATF) predation in Model (2)

minimum AIC to be 1991 (Fig. 3). This time coincides with a crossover from pollock to ATF as the major species biomass in the GOA. Fig. 5 reveals the predation effect on pollock recruitment in 2 aspects. First, the curve is below zero, which indicates that higher ATF biomass tends to reduce the recruitment of age-4 pollock. Second, the decreasing shape of the curve shows that the ATF predation effect is stronger in warmer years, which results in a lower recruitment of age-4 pollock. (The ATF predation effect is slightly nonlinear with estimated degrees of freedom being 1.07.) The ATF and temperature interaction is reasonable, because the ATF population apparently shifts toward the nursery area of juvenile pollock in warmer years. The temperature-ATF interaction and the threshold formulation make the model nonlinear. The temperature curve component of the interaction is slightly nonlinear with estimated degrees of freedom being 1.07.

Additionally, more information on recruitment and dynamics of pollock are revealed by the estimated autoregressive parameters, $\hat{\phi}_1 = 0.52$ and $\hat{\phi}_2 = -0.87$. The error process $\{\xi_t\}$ in Model (2) is stationary (Cryer & Chan 2008) and has a stochastic cycle close to 5 yr, which is consistent with stochastic structure in the age-4 pollock. The inter-group competition and cannibalism of the pollock provide potential mechanisms for the 5 yr quasi-periodicity, see Zhang et al. (2010).

The normalized residuals equal the standardized residuals pre-multiplied by the inverse square-root of the estimated error correlation matrix, and are assumed to follow an independent normal distribution with zero mean and constant variance. The linear trend in the Q-Q normal plot of the normalized residuals (upper left plot in Fig. 6) suggests that the normal distribution assumption for the errors is satisfied, which is also supported by the Shapiro-Wilks normality test result (with p-value of 0.37). The scatter plot of the normalized residuals against fitted values (upper right plot in Fig. 6) indicates that the variance of the normalized residuals is approximately constant. Additionally, both the ACF plot and Ljung-Box test suggest no autocorrelation among the normalized residuals. Furthermore, all the normalized residuals are within the lower and upper inner fences ($Q1 - 1.5 IQR = -3.1$ and $Q3 + 1.5 IQR = 3.2$, where $Q1$ and $Q3$ denote first and third quartiles respectively, and $IQR = Q3 - Q1$), so no outlier is found in the recruitment Model (2). Therefore, the assumptions of the error terms are satisfied for the fitted Model (2).

Moreover, the dynamic trend of the estimated pollock recruitment levels (dashed curve in Fig. 7) is

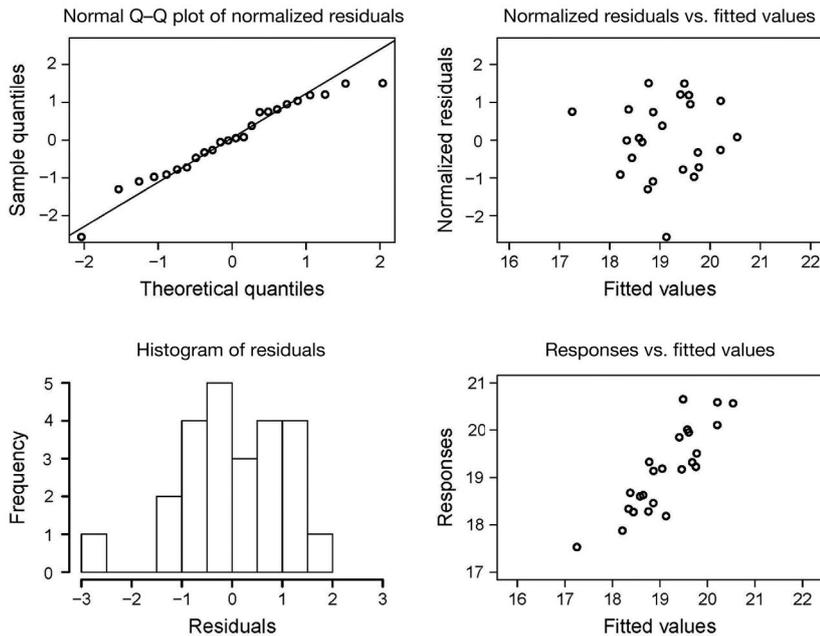


Fig. 6. Residual checks for Model (2)

consistent with the recruitment observations (gray solid curve in Fig. 7), so we conclude that Model (2) provides a good fit to the recruitment data through 2006.

Applying the bootstrap forecast discussed in 'Materials and methods', 10 000 bootstrap forecast trajectories of the age-4 recruitment have been generated over the period from 2007 to 2009. The

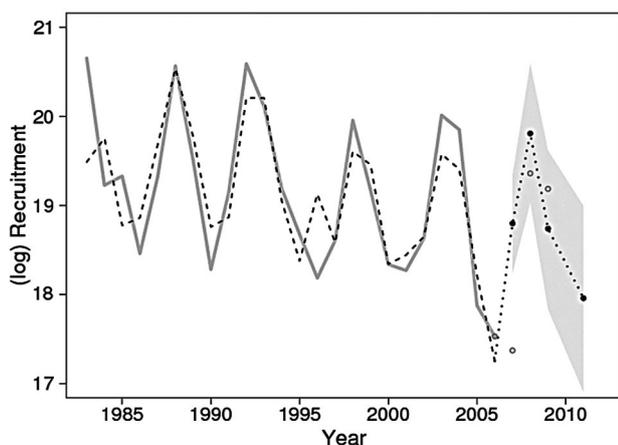


Fig. 7. Observed (log) recruitment and predicted (log) recruitment from Model (2). Numbers are abundance of 4 yr olds occurring in each year. (—): observed recruitment over the study period; (---): estimated recruitment; (·····): out-of-sample recruitment forecasts; ○: recent observed recruitments in 2007 to 2009. The 95% prediction band of the recruitment forecasts is shaded in gray

median forecast trajectory gives the point forecasts for future recruitment (dotted curve in Fig. 7). We also constructed the 95% prediction band (gray shaded area) based on the 2.5% and 97.5% quantiles of the empirical distribution of the bootstrap forecast paths. The 95% prediction band successfully covered the actual age-4 recruitment levels in 2008 and 2009, but missed the 2007 recruitment observation, although both the forecast and observed recruitment are below the average and median values in that year. There was a 5 yr dynamic cycle in the historical time series of the age-4 recruitment, which can be explained by the inter-competition and cannibalism of the pollock. The actual age-4 recruits in 2008 and 2009 were consistent with the 5 yr dynamic cycle, and were accurately predicted

by the forecast model. Therefore, the out-of-sample forecasts from Model (2) match fairly well with the actual age-4 recruitment levels.

While the recruitment Model (2) was fitted based on late larval abundance data of 2006 and earlier, the availability of the late larval indices of 2007 provides us with an opportunity to do genuine forecast for the yet unavailable age-4 abundance of 2011. Fig. 7 shows that the predicted age-4 recruitment in 2011 followed the 5 yr dynamic cycle in pollock recruitment, which is lower than the recruitment peaks around 2008 to 2009. The figure also illustrates that forecasts are increasingly uncertain with longer forecast horizons.

DISCUSSION

We propose a forecasting strategy based on larval surveys that includes: (1) adjustment of larval abundance to weight for size composition and temperature-dependent mortality, (2) potential predation occurring during juvenile life and interactions with environmental factors, (3) autocorrelative structure in the time series caused by inter-cohort interactions, and (4) recognition that shifts have occurred in factors influencing recruitment. There are only a few forecasting schemes that have used combined monitoring and environmental statistical models as we have done. One example, Axenrot & Hansson (2003),

attained relatively good results using larval abundance, spawning stock biomass and environmental factors occurring during larval life. In that study, however, since the environmental factors selected overlapped with the larval period, they served to adjust the fit of larval abundance with recruitment. The difference in our strategy is an approach that incorporates factors that account for variability in survival after the larval monitoring survey. Our results support the forecast model of Zhang et al. (2010) based on age-1 pollock abundance coupled with ATF biomass, but push back the forecast to the larval stage.

We recognize that recruitment is a complex process, which impacts the strategy to forecast year class strength (Bailey et al. 2005, Houde 2008). When dynamics are complex due to nonlinear systems and interactions, predictions become more uncertain as the time between initial conditions and the prediction target increases. Forecasts are most difficult to make years in advance strictly from environmental conditions because of the changes that occur in the relative abundance of a cohort as it develops. For pollock, there is a significant trend of abundance with recruitment by the age-1 juvenile stage (Bailey et al. 2005, Zhang et al. 2010). Due to interannual variability of mortality during juvenile life, it is expected that larval abundance, taken as a whole, is not an accurate or precise predictor of recruitment, as we found in this study. In fact, total larval abundance reflected the spawning stock abundance, even nearly 2 mo after peak spawning (K. M. Bailey unpubl. data). However, weighting the abundances of larvae for the parameters of size- and temperature-dependent mortality improved the forecast value during the period when larval survival and recruitment were coupled. We did not make an attempt to optimize the fit, which might be useful. We also believe that the larval abundance estimates could be improved, most significantly by expanding the area surveyed. Areal coverage outside of the main grid is inconsistent and in some years, relatively large numbers of larvae have been caught outside the main survey grid.

Changes in the ecosystem structure in the late 1980s to early 1990s had a remarkable effect on recruitment of pollock (Bailey 2000, Ciannelli et al. 2005, Zhang et al. 2010). The decoupling of larval abundance and recruitment that occurred in the 1990s could be explained by the gradual buildup in the abundance of ATF, finally reaching a threshold level of predation. This threshold occurred at about the same time that ATF surpassed pollock as the dominant biomass of groundfish in the GOA (Bailey 2000).

The ATF biomass has accumulated to an extent where it may consume a large portion of the pollock year class. We hypothesize a somewhat more complicated mechanism; along with the increase of ATF there has been a slow shift in their distribution. Whereas previously ATF were distributed across the western Gulf, it appears that in the early 1990s, a large portion of the population shifted towards the Shelikof Strait region nursery area, possibly in response to a predictable supply of juvenile pollock (www.afsc.noaa.gov/RACE/groundfish/default.php). It is furthermore intriguing that in 1999, the 1 pollock year class that was relatively strong during this period and a year when information on ATF distribution was available, the predator was relatively diminished in the Shelikof region, possibly blocked from the pollock nursery by cold water. In the Bering Sea, ATF avoid cold water (Spencer 2008). In the GOA, ATF also tend to avoid cold water, as in colder La Niña years they were found in warmer areas (Speckman et al. 2005). The interaction of ATF biomass and ocean temperature with pollock recruitment is supported by our exploratory TGAM analyses. A better understanding of these relationships will be difficult to unravel from historical data alone, given that the trawl surveys are triennial, and the groundfish food habits database lacks good information on ATF feeding in autumn and winter, the likely seasons when they are feeding heavily on age-0 pollock. In summer, when most stomachs are collected for examination, age-0 pollock are mostly pelagic and have limited vulnerability to ATF predation. However, these results do stress the importance of understanding the complex mechanisms involved in the recruitment process.

Alternative explanations for the shift in recruitment conditions should be considered. One potential factor is a shift in the location of spawning or a density-dependent Allee effect, such that recruits originally spawned in Shelikof are themselves spawning elsewhere (Ciannelli et al. 2007). There are several other spawning aggregations in the GOA. When there is a low abundance of spawners in Shelikof, the new recruits to the spawning population may get entrained into another spawning aggregation. Another possible explanation is a change in physical environment (Hollowed et al. 2001). The role of these processes in recruitment also merits further consideration.

There are significant obstacles to making recruitment forecasts from statistical models incorporating environmental factors; over time, forecasts based on environmental factors fail (Walters & Collie 1988), and relationships with environmental factors change

in significance. In some cases, especially when a large number of indirect factors are included, the probability of a false positive (Type I error) increases, contributing to the forecast's downfall in future years. In other cases, a large-scale atmospheric effect may be implicated, such as the PDO, but the dominance of these factors over local conditions may shift over time. For example, in the GOA and Bering Sea, the dominant influence of the PDO on oceanic conditions has been supplanted by another climate pattern (Bond et al. 2003), now known as the Victoria pattern. Shifting environmental conditions have been implicated in complex dynamics of walleye pollock recruitment in the ocean off Japan as well (Shida et al. 2007) and a changing baseline makes forecasting from environmental factors very difficult. In such conditions, short term bases for forecasts might be a better approach rather than attempting to utilize a longer time series. We also believe that significant advances in utilizing data in our statistical models are possible, such as improved understanding of processes influencing larval mortality.

Better accounting for stock structure in the GOA is another area where forecasting improvements could be made. Currently, the recruitment level is estimated from the fishery catch-at-age data (VPA) for the whole western GOA, but forecasting schemes all assume that all recruits are derived from the Shelikof Strait spawning population. Beginning in the late 1990s, the Shumagin Islands population became a significant component of the western GOA population and the inclusion of these fish in the recruitment estimate confounds comparisons of recruitment and larval numbers (and environmental conditions) from Shelikof. Discrepancies in the pollock forecasts, especially in 2007, may indicate a further decoupling of larval processes in Shelikof Strait and recruitment, and possibly even changes in control mechanisms. One possibility is that the processes in the depleted population of Shelikof Strait no longer reflect the dynamics of the GOA metapopulation. Therefore, knowledge of stock structure relative to the population sampled by the larval census and that estimated from the age-structured fisheries models is critical for a more accurate forecast.

We have found obvious autocorrelation in the recruitment level of pollock. The autocorrelation likely arose from the interactions among various pollock cohorts at different ages. The underlying cause of the autocorrelation is presently unknown, whether it is competition, maternal contributions, cannibalism or a combination of factors. But a proper autocorrelation structure is necessary for a good recruitment

forecast model. Since the interaction among the pollock cohorts could be changed by external factors, such as environmental regime shift, identifying the autocorrelation structure is a challenging topic. From a statistical consideration, ignoring the autocorrelation in the data may result in invalid assessment of the variability of the parameter estimates and their statistical significance. Thus, if the residuals from a regression model appear to be autocorrelated, it is pertinent to account for the autocorrelation in the error terms, e.g. by modeling them as some autoregressive process, as done here.

Reasonable forecasts of recruitment are possible but should be tuned each season as new information becomes available. Starting with environmental conditions, and at various abundance checkpoints along the way, observed deviations may provide valuable information on changing conditions. If the larval rough count is continued as a recruitment forecasting parameter, we suggest a forecasting scheme based on the concept that forecasting accuracy will improve as new information becomes available during the development of a cohort, using a combined monitoring and modeling effort. Improvements in the models can come from improved measurements of environmental parameters and better sampling of life stages. Managers should determine whether a forecast at a stage earlier than age-1 is cost effective. Regardless of the management requirements, there is value in attempting to forecast recruitment from larval abundance as it provides feedback to our understanding of the recruitment process, and vice versa.

LITERATURE CITED

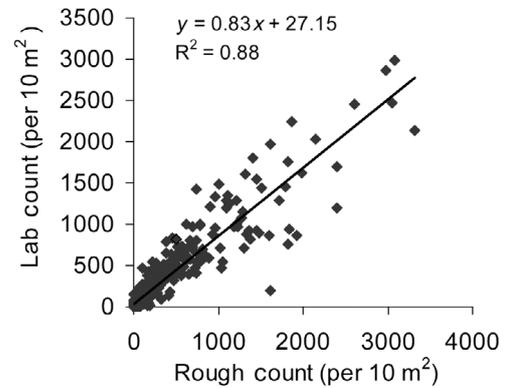
- Anderson P, Piatt JF (1999) Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Mar Ecol Prog Ser* 189:117–123
- Axenrot T, Hansson S (2003) Predicting herring recruitment from young-of-the-year densities, spawning stock biomass, and climate. *Limnol Oceanogr* 48:1716–1720
- Bailey KM (2000) Shifting control of recruitment of walleye pollock (*Theragra chalcogramma*) after a major climate and ecosystem change. *Mar Ecol Prog Ser* 198:215–224
- Bailey KM, Spring S (1992) Comparison of larval, age-0 juvenile and age-2 recruit abundance indices of walleye pollock *Theragra chalcogramma* in the western Gulf of Alaska. *ICES J Mar Sci* 49:297–304
- Bailey KM, Picquelle S, Spring S (1996) Mortality of larval walleye pollock in the western Gulf of Alaska, 1988–91. *Fish Oceanogr* 5:124–136
- Bailey KM, Ciannelli L, Bond N, Belgrano A, Stenseth NC (2005) Recruitment of walleye pollock in a complex physical and biological ecosystem. *Prog Oceanogr* 67:24–42
- Bond NA, Overland JE, Spillane M, Stabeno P (2003) Recent shifts in the state of the North Pacific. *Geophys Res Lett* 30(23):2183–2186

- Bradford MJ (1992) Precision of recruitment predictions from early life stages of marine fishes. *Fish Bull US* 90:439–453
- Ciannelli L, Bailey KM, Stenseth NC, Chan KS, Belgrano A (2005) Climate change causing phase transition of walleye pollock (*Theragra chalcogramma*) recruitment dynamics. *Proc Biol Sci* 272:1735–1743
- Ciannelli L, Bailey KM, Chan KS, Stenseth NC (2007) Phenological and geographical patterns of walleye pollock spawning in the western Gulf of Alaska. *Can J Fish Aquat Sci* 64:713–722
- Cryer DJ, Chan KS (2008) Time series analysis with applications in R. Springer Verlag, New York
- Dorn M, Aydin K, Barbeaux S, Guttormsen M, Megrey B, Spalinger K, Wilkins M (2009) Stock assessment and fishery evaluation report: Gulf of Alaska walleye pollock. North Pacific Fishery Management Council, Anchorage, AK. www.afsc.noaa.gov/REFM/docs/2009/GOApollock.pdf
- Dougherty AB, Bailey KM, Mier K (2007) Interannual differences in growth and hatch date distributions of age-0 walleye pollock (*Theragra chalcogramma*) from the Shumagin Islands region of the Gulf of Alaska, 1985–2001. *J Fish Biol* 71:763–780
- Helle K, Bogstad B, Marshall TT, Michalsen K, Ottersen G, Pennington M (2000) An evaluation of recruitment indices for Arcto-Norwegian cod (*Gadus morhua* L.). *Fish Res* 48:55–67
- Hjort J (1914) Fluctuations in the great fisheries of northern Europe, viewed in the light of biological research. *Rapp P-v Reun Cons Perm Int Explor Mer* 20:1–228
- Hollowed AB, Hare SR, Wooster WS (2001) Pacific basin climate variability and patterns of Northeast Pacific marine fish production. *Prog Oceanogr* 49:257–282
- Houde ED (2008) Emerging from Hjort's shadow. *J Northwest Atl Fish Sci* 41:53–70
- Hsieh C, Yamauchi A, Nakazawa T, Wang W (2010) Fishing effects on age and spatial structures undermine population stability of fishes. *Aquat Sci* 72:165–178
- Lin X, Zhang D (1999) Inference in generalized additive mixed models using smoothing splines. *J R Stat Soc B* 61: 381–400
- Meekan MG, Milicich MJ, Doherty PJ (1993) Larval production drives temporal patterns of larval supply and recruitment of a coral reef damselfish. *Mar Ecol Prog Ser* 93: 217–225
- Mukhina NV, Marshall CT, Yaragina NA (2003) Tracking the signal in year-class strength of northeast Arctic cod through multiple survey estimates of egg, larval and juvenile abundance. *J Sea Res* 50:57–75
- Porter S, Brown AB, Bailey K (2001) Estimating live standard length of net-caught walleye pollock (*Theragra chalcogramma*) larvae using measurements in addition to standard length. *Fish Bull* 99:691–696
- Schnute J (1981) A versatile growth model with statistically stable parameters. *Can J Fish Aquat Sci* 38:1128–1140
- Sette OE (1943) Biology of the Atlantic mackerel (*Scomber scombrus*) of North America. Part 1: early life history, including growth, drift, and mortality of the egg and larval populations. U.S. Fisheries Wildlife Service. *Fish Bull* 50:149–237
- Shida O, Hamatsu T, Nishimura A, Suzaki A, Yamamoto J, Miyashita K, Sakurai Y (2007) Interannual fluctuations in recruitment of walleye pollock in the Oyashio region related to environmental changes. *Deep-Sea Res II* 54: 2822–2831
- Smith TD (1994) Scaling fisheries. Cambridge University Press, Cambridge
- Speckman SG, Piatt JF, Minte-Vera CV, Parrish JK (2005) Parallel structure among environmental gradients and three trophic levels in a subarctic estuary. *Prog Oceanogr* 66:25–65
- Spencer PD (2008) Density-independent and density-dependent factors affecting temporal changes in spatial distributions of eastern Bering Sea flatfish. *Fish Oceanogr* 17:396–410
- Turnock BJ, Wilderbuer TK (2007) Stock assessment and fishery evaluation report: Gulf of Alaska arrowtooth flounder stock assessment. North Pacific Fishery Management Council, Anchorage, AK. www.afsc.noaa.gov/refm/docs/2007/GOAatf.pdf
- Van der Veer HW (1986) Immigration, settlement, and density-dependent mortality of a larval and early postlarval 0-group plaice (*Pleuronectes platessa*) population in the western Wadden Sea. *Mar Ecol Prog Ser* 29:223–236
- Walters CJ, Collie JS (1988) Is research on environmental factors useful to fisheries management? *Can J Fish Aquat Sci* 45:1848–1854
- Wood SN (2006) Generalized additive models: An introduction with R. Chapman & Hall/CRC, Boca Raton, FL
- Wu CFJ (1986) Jackknife, bootstrap and other resampling methods in regression analysis (with discussions). *Ann Stat* 14:1261–1350
- Zhang T, Bailey KM, Chan KS (2010) Recruitment forecast models of walleye pollock (*Theragra chalcogramma*) fine-tuned from juvenile survey data and predator community structure. *Mar Ecol Prog Ser* 417:237–248

Appendix 1. Additional data

Table A1. Schedule of instantaneous mortalities used for each length group. SL = standard length

Live SL (mm)	Age (d)	Mortality		
		Average	Warm year	Cold year
5	1	0.3	0.15	0.45
6	9	0.267	0.1335	0.4005
7	16	0.088	0.044	0.132
8	22	0.035	0.0175	0.0525
9	27	0.016	0.008	0.024
10	32	0.008	0.004	0.012
11	36	0.005	0.0025	0.0075
12	40	0.005	0.0025	0.0075
13	44	0.005	0.0025	0.0075
14	47	0.005	0.0025	0.0075
15	50	0.005	0.0025	0.0075

Fig. A1. *Theragra chalcogramma*. Accuracy of the at-sea (rough) counts versus the verified counts from the laboratory for walleye pollock in 9 cruises conducted from 1999 to 2005

Editorial responsibility: Alejandro Gallego,
Aberdeen, UK

Submitted: September 12, 2011; Accepted: January 18, 2012
Proofs received from author(s): April 13, 2012