Using Satellite Data to Improve Short-Term Recruitment Predictions for Marine Fish Stocks[1]

Jon Brodziak
NOAA National Marine Fisheries Service
Pacific Islands Fisheries Science Center
2570 Dole Street
Honolulu, HI 96822
Email: Jon.Brodziak@NOAA.GOV

Abstract

In the Northeast U.S. region, an age-structured projection model of fish population dynamics (AGEPRO) has been used for making short-term catch predictions under recruitment uncertainty. This model has been used to determine annual catch quotas under court-mandated stock rebuilding plans for eleven New England groundfish stocks, including the Georges Bank cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) stocks. Research has shown that recruitment strength of Georges Bank cod and haddock stocks may be influenced by environmental covariates such as the North Atlantic Oscillation Index. We used NASA and NOAA satellite data products to develop predictive models of how sea surface water temperature, as indexed by Pathfinder v5 5.5 km monthly SST and primary productivity, as indexed by SeaWiFS monthly ocean color measurements, influenced recruitment strength of the Georges Bank cod and haddock stocks. We found that Georges Bank cod recruitment strength was significantly negatively associated with average sea surface temperature during February-May. Haddock recruitment was positively associated with the strength of the autumn plankton bloom in the previous year and with the haddock age-0 research survey index in the current year. These results were used to develop several predictive models of cod and haddock recruitment using from 1985-2004. The resulting models were added to the NOAA Fisheries Toolbox AGEPRO projection module. These models were then compared to previous recruitment prediction models by predicting observed recruitments during 2005-2007. While some of the new recruitment prediction models performed poorly, several resulted in substantial reductions in the root mean-square error of predicted versus observed recruitment during 2005-2007. For cod, the best model was based on spring sea surface temperature and reduced the root mean-square error of predicted recruitment by about 70%. For haddock, five models that used combinations of spring sea surface temperature, primary productivity, and the haddock age-0 survey index variables reduced prediction error by 66% to 81%. In this case, the best predictive model was a model-averaged combination of two predictive models, one that used sea surface temperature and the haddock age-0 index, and one that used only sea surface temperature.

[1] PIFSC Internal Report IR-09-015
Issued 5 May 2009

The results of this project illustrate that it may be possible to improve short-term predictions of recruitment for some major fishery resources using readily available satellite data products. Further, the results suggest that it may be useful to consider multiple model inference techniques (e.g., Burnham and Anderson 2002) to better understand the factors affecting recruitment dynamics and to formulate better predictive models. In the future, this research approach may have practical application to Pacific marine fisheries where remotely-sensed oceanographic conditions may also be expected to influence recruitment strength, e.g., salmonids and tunas.

Introduction

Recruitment, the abundance of annual broods of fish, is an intrinsically important factor affecting fishery productivity and management. It has long been recognized that environmental conditions influence the recruitment strength of marine fishes by altering early life history survival. Spawner abundance and population age structure also affect recruitment strength through total egg production and egg size or quality. Predator and prey abundances of egg and larval stages also affect recruitment and such trophic interactions are often categorized as environmental conditions as well. Annual variability in environmental conditions has made it difficult to accurately predict recruitment on a short-term basis even with knowledge of spawner abundance. This uncertainty directly impacts fishery management decisions to set annual catch quotas, especially when adult fish abundance is low and quotas depend heavily on incoming recruitment. In this case, if recruitment is overestimated, catch quotas may be set too high compromising resource conservation. Alternatively, if recruitment is underestimated, quotas may be too low and short-term yields may be forgone. In general, improving the accuracy of recruitment predictions would help fishery managers to set appropriate catch quotas. This is particularly important when fishing effort and catch quotas have to be reduced to rebuild overfished stocks (e.g., Greene 2002).

In the Northeast US, an age-structured simulation model of fish population dynamics (AGEPRO) has been used by the Northeast Fisheries Science Center (NEFSC) to make short-term catch projections under recruitment uncertainty. This model has been used to estimate annual catch quotas under court-mandated stock rebuilding plans for eleven New England groundfish stocks (Brodziak et al. 1998, Overholtz et al. 1999, New England Fishery Management Council[NEFMC] 2004), including the commercially-important Georges Bank cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) stocks. The AGEPRO model has also been used to compute groundfish bycatch quotas for special access programs designed to harvest the abundant haddock stock in closed areas on Georges Bank (Figure 1). Overall, this model has served as an important decision support tool for managing New England groundfish stocks.

Recent research has shown that recruitment strength of Georges Bank cod and haddock stocks may be influenced by environmental covariates such as the North Atlantic oscillation index (Brodziak and O'Brien. 2005, Stige et al. 2006) which have localized effects on water temperatures and wind patterns. In 2003, the Georges Bank haddock stock produced an extremely abundant year class, the largest ever-recorded. Subsequent investigation of NOAA buoy data indicated a southwestward wind stress over haddock spawning areas during spring 2003 (Brodziak and Mountain, unpublished data). This southwesterly wind pattern likely enhanced the retention of haddock eggs and larvae in shoal waters of the continental shelf instead of advective losses off Georges Bank under a southeastward wind stress. To investigate whether oceanographic conditions that enhance larval retention on Georges Bank have a significant positive effect on cod and haddock recruitment (Page et al. 1999, Mountain et al. 2003, Brodziak 2005), a concurrent project was developed to assess the importance of wind stress on cod and haddock recruitment and this project was funded through NOAA's Fisheries and the Environment program (Mountain et al. 2006). As a result, the focus of this project was to investigate the importance of other environmental covariates on cod and haddock recruitment, such as water temperature and food availability, that are known to affect larval growth and early life history stage survival of cod and haddock (Laurence and Rogers 1976, Green et al. 2004, Rideout et al. 2005).

To address this research need, we extended the AGEPRO model to include environmental covariates or other biological variables for short-term recruitment predictions. We used NASA and NOAA satellite data products to develop predictive models of how sea surface water temperature and primary productivity influences recruitment strength of the Georges Bank haddock and cod stocks using data from 1985-2004. These models were then applied to make out-of-sample predictions of observed recruitments during 2005-2007. The accuracy of the predictions using environmental covariates was compared to that of existing recruitment prediction models for these stocks.

Materials and Methods

In this section, data and methods are described beginning with the AGEPRO model and the stock-recruitment and projection data for Georges Bank cod and haddock stocks. Satellite data and the associated environmental covariates are covered next followed by descriptions of the analyses to investigate recruitment strength and environmental covariates and the out-of-sample projections to compare the recruitment prediction models.

Age-Structured Projection Model

The AGEPRO model is an age-structured projection model for evaluating the outcomes of alternative harvest policies on an exploited fish stock accounting for random variation in initial stock size estimates, recruitment, and natural mortality (Brodziak et al. 1998, available at http://nft.nefsc.noaa.gov/). The AGEPRO model is coded in Fortran 95 with a graphical user interface (GUI) and standardized input and output files along with graphical output options. The inputs consist of biological data, fishery data and harvest scenario information to parameterize the age-structured projection model. The output consists of annual distributions of stock size, spawning biomass, landings, realized fishing mortality (if quotas are specified as the harvest control variable), recruitment, initial and final population size at age, landings by market category if applicable, probabilities that spawning biomass target is achieved in each year and overall, probabilities that fishing mortality threshold is exceeded if applicable and so on.

Two key uncertainties for the AGEPRO model are the fish population numbers at age at the beginning of the projection period and the stochastic recruitment generated during the projection time period. Typically, the uncertainty in the initial stock size is evaluated using resampling techniques such as the non-parametric bootstrap or using Markov Chain Monte Carlo samples generated from the posterior distribution of stock size estimates at the start of the projection. The submodels used to project future recruitment can be either dependent or independent of spawner abundance and include a random error term. For some fish stocks, such as Georges Bank haddock, there may be a pronounced maternal effect and adequate spawner abundance must be maintained to produce high recruitments, on average (Brodziak et al. 2001), while for other stocks, no apparent relation between spawner abundance and recruitment is apparent.

In this project, four new submodels for stochastic recruitment prediction were developed to incorporate relevant environmental data collected by NASA and NOAA satellites or other predictive biological variables. These new submodels were added to the fifteen existing recruitment submodels in the AGEPRO software for application to the Georges Bank cod and haddock data sets. The first new submodel was a linear recruits per spawner predictor with a normal error term. In this submodel, recruitment (R) per unit of spawner abundance (S) is predicted from a linear function of one or more covariates (X_j) and an additive normal error term (ε) with zero mean and constant variance (σ^2).

(1.1) $$\frac{R}{S} = \beta_0 + \sum_j \beta_j \cdot X_j + \varepsilon$$

The simulated value of recruits per spawner (R/S) is then multiplied by projected spawner abundance to generate recruitment.

The second submodel was a loglinear model to predict recruits per spawner as a function of environmental covariates using a multiplicative lognormal error in which the natural logarithm of the error is a normal random variable with $\varepsilon \sim N(-0.5\sigma^2, \sigma^2)$.

(1.2) $$\log\left(\frac{R}{S}\right) = \beta_0 + \sum_j \beta_j \cdot X_j + \varepsilon$$

In this case, the predicted value of R/S includes the bias-correction for back-transformation from the logarithmic scale under a lognormal error term. Recruitment is generated by multiplying the simulated R/S value by the projected spawner abundance.

The third recruitment submodel was a linear recruitment predictor with a normal error term similar to the first submodel.

(1.3) $$R = \beta_0 + \sum_j \beta_j \cdot X_j + \varepsilon$$

The fourth recruitment submodel was a loglinear recruitment predictor with a multiplicative lognormal error in which the natural logarithm of the error is a normal random variable with $\varepsilon \sim N(-0.5\sigma^2, \sigma^2)$.

(1.4)
$$\log(R) = \beta_0 + \sum_j \beta_j \cdot X_j + \varepsilon$$

For the two models that used an additive normal process error term (eqns 1.1 and 1.2), it was possible that an infeasible negative recruitment value could be simulated if the error variance was sufficiently large. In this case, the simulated value of recruitment was constrained to be positive by repeating the random sampling process until a positive value of R was obtained.

Stock-Recruitment and Projection Data

Stock-recruitment data Georges Bank cod and haddock were taken from the 2005 stock assessments (O'Brien et al. 2006, Brodziak et al. 2006). Recruitment data from the 1985-2004 year classes were used to fit recruitment submodels for cod (Figure 2a) and haddock (Figure 2b). This time period coincided with the available AVHRR sea surface temperature time series. Projection data were also taken from the 2005 stock assessments (O'Brien et al. 2006, Brodziak et al. 2006) as recommended in the 2005 Groundfish Assessment Review Meeting report (Mayo and Terceiro 2005); this included average spawning and landed weights at age, fraction mature at age, natural mortality at age, and fishery selectivity at age (Table 1). Projection data for the population size at age distributions of Georges Bank cod and haddock at the start of 2005 were the same as used in New England Fishery Management Council's Groundfish Plan Development Team's analyses (e.g., NEFMC 2007).

Satellite Data and Environmental Covariates

Satellite data were gathered from the NOAA Coastwatch West Coast node's online data sets (http://coastwatch.pfel.noaa.gov/). Satellite data were extracted using a Matlab-based program (xtracto_ts_bdap, see Appendix) developed by the Southwest Fisheries Science Center's Environmental Research Division (Dave Foley, SWFSC pers. comm.). This program allows the user to select a time series of available satellite data within a user-specified rectangular region. The program was applied to gather satellite data from three rectangular regions centered on the Georges Bank in the Northwest Atlantic. These were: the Georges Bank superregion R1 (Figure 3, dashed blue line) with upper-left and bottom-right coordinates of (43˚N, -70˚W) and (40˚N, -66˚W); the Georges Bank region R2 (Figure 3, dotted red line) with upper-left and bottom-right coordinates of (42˚N, -69˚W) and (40˚N, -66˚W); and the Georges Bank central subregion R3 (Figure 3, solid green line) with upper-left and bottom-right coordinates of (42˚N, -68˚W) and (41˚N, -67˚W). In this case, the primary region used for all of the recruitment analyses was R2, the Georges Bank region. The other regions were used to assess the spatial coherence of the satellite data and whether the satellite time series were correlated across spatial scales.

A total of 20 time series of satellite data in the three regions were collected from the NOAA Coastwatch West Coast node (http://coastwatch.pfel.noaa.gov/). These included two chlorophyll data sets (xtracto_ts_bdap indices 21 and 22), two primary productivity data sets (indices 41 and 42), one sea surface height data set (index 23), five sea surface temperature data sets (indices 10, 12, 15, 18, and 19), and ten wind speed or direction data sets (indices 25, 26, 27, 28, 30, 31, 32, 34, 35, and 36). Of these twenty data sets, two were selected for recruitment analyses based on the length of the available time series and on recent research results on the effects of environmental factors on cod and haddock early life history stage survival and growth.

The first data set selected was the SeaWiFS primary productivity monthly composite index (index 42), a derived satellite data product. The primary productivity data set was chosen because recent research by Friedland et al. (2008) has suggested that the magnitude of the autumn plankton bloom may have an important influence on haddock recruitment success the following spring. An autumn bloom time series for 1997 to 2006 was calculated as the three-month average of the primary productivity indices for September, October, and November in region R2; this variable was denoted as PP2.fall1 where the "1" denotes that it needs to be lagged forward to be compared to a recruitment response variable. In this case, higher autumn bloom values would be expected to improve adult foraging conditions and subsequent spawning capacity in spring under the parental condition hypothesis suggested by Friedland et al (2008). Spring bloom time series were calculated for cod and haddock spawning seasons. For cod, the spring bloom time series was calculated as the two-month average of primary productivity indices during February and March in region R2; this variable was denoted as PP2.spr.fm. For haddock, the spring bloom time series was calculated as the three-month average of primary productivity indices during March, April, and May in region R2; this variable was denoted as PP2.spr.mm. In this case, higher spring bloom values would be expected to be associated with higher larval survival rates and associated with higher recruitment in that year under the juvenile foraging condition hypothesis, which provides an alternative to the parental condition hypothesis (Payne et al. In press). As a result, there was one autumn and two spring primary productivity indices for use as environmental covariates in recruitment analyses (Figure 4a).

The second data set selected was the Pathfinder v5 sea surface temperature (SST) monthly composite index (index 19). This data set was chosen because growth rates of larval cod and haddock vary with temperature, there were 22 years of SST observations available from 1985-2006, and recent research has suggested an association between haddock growth increments and spring temperature experienced as young of the year (Brodziak and Link 2008). For cod, a spring SST time series was calculated as the average monthly SST indices for February and March in region R2 (ST2.fm) while for haddock, a spring SST series was calculated as the 4-month average SST indices for February through May in region 2 (ST2.mm). In this case, the hypothesis was that favorable temperatures for larval growth, generally on the order 6 ˚C for cod and 7 ˚C for haddock (Campana and Hurley 1989) would be associated with higher recruitment. As a result, there were two time series of SST indices for use as environmental covariates in recruitment analyses (Figure 4b).

An additional biological covariate was collected for Georges Bank haddock; this was the Northeast Fisheries Science Center autumn trawl survey haddock age-0 swept-area numbers index divided by 1000 (NEFSC 2008b and pers. comm. Michele Traver, NEFSC). This young-of-the-year index was included because it is positively correlated with VPA estimates of haddock recruitment and provided another potential within-year predictor of haddock recruitment strength.

Recruitment Analyses

The goal of the recruitment analyses was to determine a set of potential short-term recruitment prediction models for Georges Bank cod and haddock. In this context, developing a prediction for either recruitment (R) or recruits per spawner (R/S) as a measure of recruitment response was considered because either could be used to determine recruitment strength. Four sets of analyses were conducted. The first was a cross-check to assess the degree of spatial coherence in the satellite data. If these data were not positively correlated then this would indicate that problems may have occurred with the retrieval or summarization of the available satellite data. To assess this, Pearson correlations were calculated for each of the five environmental covariates across the pairs of regional scales.

The second analysis was to assess whether f R or R/S exhibited a nonlinear response to any of the predictors. To do this, generalized additive models (GAMs, Hastie and Tibshirani 1990) were used to predict recruitment, recruits per spawner, or log-transformed recruits per spawner for cod and haddock as a smoothed function one of the environmental covariates. In this case, the autumn primary productivity index was evaluated with a one-year lag to represent the parental condition hypothesis and was also evaluated as a contemporaneous predictor of age-0 growth conditions in autumn. Thus, six potential environmental covariates were evaluated for each of six recruitment response variables. The results of these nonparametric analyses were used to guide development of parametric models for recruitment prediction.

The third analysis was to investigate whether any of the environmental covariates were significantly correlated, and hence collinear, and also to assess whether any strong correlations between the environmental covariates and the recruitment response (R or R/S). The correlation analyses were also used to guide the development of parametric models for recruitment prediction, e.g., using pairs of highly correlated covariates was to be avoided.

The fourth set of analyses was to identify and fit a set of linear regression models to predict recruitment response given the findings in the previous analyses. In this case, linear models without an intercept were considered to allow for a direct proportional response.

Projection Analyses

Projection analyses were conducted to test whether any of the recruitment submodels identified in the recruitment analyses produced more accurate predictions than the existing models for Georges Bank cod and haddock. The existing (status quo) recruitment prediction models for cod and haddock were taken from the recommendations of the 2005 Groundfish Assessment Review Meeting (Mayo and Terceiro 2006). For Georges Bank cod, the recruitment prediction model was a Beverton-Holt curve with lognormal error (NEFSC 2002). For Georges Bank haddock, the recruitment prediction model was a two-stage cumulative distribution function for observed recruitments above and below the productivity threshold of 75,000 mt of spawning biomass (NEFSC 2002). Thus, for both status quo models, predicted recruitment was dependent on spawning biomass.

To compare the status quo and any new recruitment prediction models, estimates of recruitment for Georges Bank cod and haddock during 2005-2007 were gathered from the recently completed 2008 stock assessments (NEFSC 2008a, NEFSC 2008b). Observed values of sea surface temperatures were not available in 2007 and SST in 2007 was imputed using the average sea surface temperature during 1985-2006. Observed catch biomasses of Georges Bank cod and haddock during 2005 to 2007 were input to the AGEPRO model to compute annual fishing mortality during 2005-2007 for each projection. For cod, the catch biomasses in 2005-2007 were 4401, 4611, and 5957 mt while for haddock the catch biomasses in 2005-2007 were 21814, 15989, and 16815 mt. All of the projections were conducted using version 3.3 of the AGEPRO model software which includes the four additional recruitment prediction submodels developed for this project. For each model scenario, a total of 100 simulations were conducted for each of 1000 bootstrap initial population size vectors. Thus, each projection consisted of 100000 simulated population trajectories for summary analyses.

Because the 2008 stock assessments for Georges Bank cod and haddock were bench mark assessments, and not simple assessment updates, estimates of recruitment, spawning biomass, and other variables were expected to have a somewhat different scale than those from the 2005 assessments. In this case, comparing the projected recruitments during 2005-2007 with the observed values from the assessment could be misleading. To address this concern, the best-fitting linear model to predict observed from the 2008 assessment as a function of the 2005 assessment value during 1985-2004 was used to rescale predicted recruitments during 2005-2007 to be comparable to the values in the 2008 assessments of cod and haddock. Differences among recruitment predictions were quantified using the root mean-square error of the predicted recruitments during 2005-2007. In this way, any improvement in recruitment prediction was measured relative to using the status quo model.

Results

Each of the environmental covariates was highly positively correlated among the three regional scales (Table 2). This indicated that the satellite measurements were consistent across an order of magnitude of spatial extent. This also suggested that the choice of the spatial scale to represent the environmental covariates for Georges Bank was adequate.

The GAM analyses suggested that there were a few important nonlinear responses to the predictors. The GAM fits using the primary productivity suggested that there might be a relationship between cod recruitment ($P=0.05$) or recruits per spawner ($P=0.06$) and the spring primary productivity during February-March index (Figure 5). Similarly, the GAM fits of cod recruitment (P=0.01) and recruits per spawner (P=0.07) using sea surface temperature during February-May suggested that spring SST may have a curvilinear effect on cod recruitment response (Figure 6). Last, a GAM fit of log-transformed haddock recruits per spawner using the autumn primary productivity index forward-lagged one year suggested that there might be a positive influence (P=0.11) on haddock R/S for high values of primary productivity (Figure 7). Thus, a total of 5 out of 36 GAM fits suggested a potential nonlinear effect. However, as a practical modeling tool, the GAM analyses were rather limited by the amount of data.

Correlations among the environmental covariates were not substantial except for two cases. There was a significant positive correlation between spring primary productivity indices during February-March and during March-May (ρ=0.64, P=0.05). Similarly, there was a significant positive correlation between sea surface temperature indices during February-March and during February-May (ρ=0.66, P=0.00). As a result, these pairs of covariates were not used in the same model.

Correlations between environmental covariates and recruitment response variables suggested that there were a few important associations. In particular, sea surface temperate during February-May was negatively associated (Figure 8) with both cod recruitment (ρ= -0.50, P=0.02) and recruits per spawner (ρ= -0.40, P=0.08). This suggested that low SST values had a positive effect on cod recruitment. Similarly, the forward-lagged autumn primary productivity index was positively associated (Figure 9) with haddock recruitment (ρ= 0.65, P=0.11) and recruits per spawner (ρ= 0.69, P=0.09). This suggested that high autumn primary productivity had a positive effect on haddock recruitment the subsequent spring. Last, there was a significant positive correlation between cod and haddock recruits per spawner (ρ= 0.61, P=0.00). This indicated that the survival ratios of these two gadids fluctuated in a similar manner during 1985-2004.

Cod Recruitment Models

A total of three models to predict Georges Bank cod recruits per spawner were identified as being potentially useful. The first model ($M_{COD,RS1}$) was a no-intercept model using spring sea surface temperature during February-May. The estimated linear model had the form

$$\frac{R}{S} = 0.0463 \cdot ST2.spr.mm + \varepsilon \quad \text{where } \varepsilon \sim N\left(0,\ 0.0653\right) \tag{1.5}$$

The fitted model was highly significant (P<0.001) and explained a substantial amount of variation in R/S relative to the model R/S = 0 + ε (multiple R^2 = 0.61).

The second model for cod R/S ($M_{COD,RS2}$) was also a no intercept model but used log-scale R/S. In this case, the estimated model was

$$\log\left(\frac{R}{S}\right) = -0.1977 \cdot ST2.spr.mm + \varepsilon \quad \text{where } \varepsilon \sim N\left(-0.2058,\ 0.4117\right) \tag{1.6}$$

The fitted model was also highly significant (P<0.001) and explained a substantial amount of variation in the R/S data relative to the model log(R/S) = 0 + ε (multiple R^2 = 0.64).

The third model for cod R/S ($M_{COD,RS3}$) included an intercept and was estimated to be

$$\frac{R}{S} = 0.9988 - 0.0993 \cdot ST2.spr.mm + \varepsilon \quad (1.7)$$
$$where \;\; \varepsilon \sim N(0, 0.0489)$$

This model was not significant (P=0.08) and explained a low amount of variation in the R/S data relative to the model R/S = μ + ε (multiple R^2 = 0.16).

There were three models identified as potential predictors of Georges Bank cod recruitment strength (R, in units of millions of age-1 fish). For each, the predictor was spring SST during February-May. The first model ($M_{COD,R1}$) was a linear model with no intercept and was estimated to be

$$R = 1.5876 \cdot ST2.spr.mm + \varepsilon \quad (1.8)$$
$$where \;\; \varepsilon \sim N(0, 109.8)$$

The fitted model was highly significant (P<0.001) and explained a substantial amount of variation in R relative to the model R = 0 + ε (multiple R^2 = 0.78).

The second model for cod R ($M_{COD,R2}$) was a no intercept model fit to log-scale R and was estimated to be

$$\log(R) = 0.3082 \cdot ST2.spr.mm + \varepsilon \quad (1.9)$$
$$where \;\; \varepsilon \sim N(-0.4511, 0.9021)$$

The second model for cod R was highly significant (P<0.001) and explained much of the variation in the R data relative to the model log(R) = 0 + ε (multiple R^2 = 0.84).

The third model for cod R ($M_{COD,R3}$) included an intercept and was fitted to R. The estimated model was

$$R = 46.3885 - 5.1749 \cdot ST2.spr.mm + \varepsilon \quad (1.10)$$
$$where \;\; \varepsilon \sim N(0, 72.7)$$

This model was significant (P=0.02) and explained some of the variation relative to the model R = μ + ε (multiple R^2 = 0.25).

<u>Haddock Recruitment Models</u>

One model was identified to predict Georges Bank haddock recruits per spawner ($M_{HAD,RS1}$). This was a no-intercept linear model with sea surface temperature during February-May as the predictor. The estimated model had the form

$$\log\left(\frac{R}{S}\right) = -0.1746 \cdot ST2.spr.mm + \varepsilon \quad (1.11)$$
$$where \;\; \varepsilon \sim N(-1.065, 2.129)$$

The fitted model was highly significant (P=0.002) and explained a moderate amount of variation in haddock R/S relative to the model log(R/S) = 0 + ε (multiple R^2 = 0.41).

A total of five models were identified to predict Georges Bank haddock recruitment strength (R, in units of millions of age-1 fish). The first model ($M_{HAD,R1}$) was a linear model with no intercept fit to log-scale R as a function of sea surface temperature during February-May. The fitted model was

$$\log(R) = 0.3588 \cdot ST2.spr.mm + \varepsilon \quad \text{where } \varepsilon \sim N(-1.209, 2.418) \tag{1.12}$$

The fitted model was highly significant (P<0.001) and explained a good amount of variation in the R data relative to the model log(R) = 0 + ε (multiple $R^2 = 0.72$).

The second model for haddock R ($M_{HAD,R2}$) was a no intercept model fit to log-scale R using the lagged autumn primary productivity index. The estimated model was

$$\log(R) = 0.3588 \cdot PP2.fall1 + \varepsilon \quad \text{where } \varepsilon \sim N(-1.758, 3.516) \tag{1.13}$$

This model was also highly significant (P<0.001) and explained much of the variation in haddock R relative to the model log(R) = 0 + ε (multiple $R^2 = 0.78$).

The third model to predict haddock recruitment ($M_{HAD,R3}$) fitted log-scale recruitment using sea surface temperature during February-May and the haddock age-0 survey index. The estimated model was

$$\log(R) = 0.3195 \cdot ST2.spr.mm + 0.0101 \cdot age0.had + \varepsilon \quad \text{where } \varepsilon \sim N(-0.594, 1.188) \tag{1.14}$$

This model was highly significant (P<0.001) and explained a substantial amount of the variation in haddock R relative to the model log(R) = 0 + ε (multiple $R^2 = 0.87$).

The fourth model to predict haddock recruitment ($M_{HAD,R4}$) also used sea surface temperature during February-May and the haddock age-0 survey index but was fitted to untransformed haddock R. The estimated model was

$$R = 1.1362 \cdot ST2.spr.mm + 1.5567 \cdot age0.had + \varepsilon \quad \text{where } \varepsilon \sim N(0, 386.5) \tag{1.15}$$

This model was also highly significant (P<0.001) and explained much of the variation in haddock R relative to the model R = 0 + ε (multiple $R^2 = 0.99$).

The fifth model to predict haddock recruitment ($M_{HAD,R5}$) was a model-averaged combination of models $M_{HAD,R5}$) and $M_{HAD,R5}$). In the absence of a preference, the two model probabilities were set equal to 0.5. In this case, each model was randomly sampled with probability one-half to simulate recruitment in each year of the stochastic projections.

Regression analyses and associated Akaike information criteria values indicated that the best fitting linear model relating the new 2008 VPA estimates of Georges Bank cod recruitment to the old estimates from the 2005 assessment was $R_{NEW} = 0.9822 \cdot R_{OLD}$ while the best model for haddock R was $R_{NEW} = 6.076 + 0.6247 \cdot R_{OLD}$. These models were used to rescale the predicted recruitment values from the projections using both the status quo models and newly developed recruitment models using the environmental covariates.

Cod Projection Results

Projection results for Georges Bank cod indicated that the performance of the recruitment models using environmental covariates was disparate (Table 3). The three recruits per spawner models did not perform well in comparison to the status quo model at predicting cod year class strength during 2005-2007. These three models had average estimation errors on the order of 89%, 10%, and 80% higher than the status quo model (Table 3). By comparison, the three recruitment strength models performed much better with average estimation errors of 9%, -70%, and 1% relative to the status quo Georges Bank cod model. The best recruitment prediction model was $M_{COD,R2}$ which used sea surface temperature during February-May to predict cod recruitment strength (eqn 1.9). The best predictor had a root mean-square prediction error that was 3-fold lower than the status quo model (Figure 10).

Haddock Projection Results

Projection results for Georges Bank haddock indicated that the performance of the recruitment models developed using environmental covariates also differed across model types (Table 4). The only recruits per spawner prediction model for haddock did not perform well in comparison to the status quo recruitment model during 2005-2007 and had an average estimation error that was roughly 50% greater than the status quo. In contrast, the five haddock recruitment strength models performed much better than the status quo model (Table 4). These models had average estimation errors that were 73% to 81% lower than the status quo Georges Bank haddock model. Overall, the best-fitting model was $M_{HAD,R5}$ which was a model-averaged combination with equal model probabilities of ½ of models $M_{HAD,R1}$ (eqn 1.12) and $M_{HAD,R4}$ (eqn 1.15). This model-averaged combination had a root mean-square prediction error that was roughly 5-fold lower than the status quo model (Figure 11).

Discussion

While some of the recruitment prediction models performed poorly, several resulted in substantial improvements in the root mean-square error of predicted versus observed recruitment during 2005-2007. For cod, the best prediction model using spring sea surface temperature reduced the root mean-square error of predicted recruitment by about 70%. For haddock, the five models that used combinations of sea surface temperature, primary productivity, and the haddock age-0 survey index variables reduced prediction error by 66% to 81%. In this case, the best predictive model was a model-averaged combination of two predictive models, one that used sea surface temperature and the haddock age-0 index and one that used only sea surface temperature. The haddock example suggests that the use of multiple predictive models may be able to improve predictive accuracy in some cases. This might be expected because recruitment dynamics are generally influenced by multiple biotic and environmental processes operating at differing spatial and temporal scales. Regardless, this work does illustrate the potential utility of considering environmental covariates in making short-term predictions of recruitment for setting total allowable catches.

The predictive models for recruits per spawner developed in this study did not perform as well as the predictive models for recruitment strength. This is a not a general result and it is likely that attempts to predict recruitment from the recruits per spawning distribution may work well for other stocks that are not at unusually low or high abundances. In particular, the Georges Bank cod stock exhibited an abrupt decline in spawning biomass in the 1990s (Figure 2a) but has not yet exhibited a strong compensatory response of increased recruits per spawner. This pattern suggests that recruits per spawner dynamics of cod are being influenced by factors independent of spawner abundance. In contrast, recruits per spawner dynamics of haddock in recent years have been influenced by the maturation of the exceptional 2003 year class, which has, in turn, led to historic record spawning biomass levels of Georges Bank haddock. The unusually high abundance of adult haddock has likely altered the recruits per spawner dynamics in recent years making predictions based on historic R/S less reliable.

One positive result of this study was that off-the-shelf satellite data products can be readily used to investigate the influence of environmental covariates on recruitment dynamics. The ease of developing appropriate indices of relevant oceanographic conditions has been greatly facilitated by the development of online data distribution sites, such as the NOAA Coastwatch programs. This is a positive development for future research on the influence of changing climate conditions on fisheries productivity and stock-recruitment dynamics.

Overall, the results of this project show that it may be possible to improve short-term predictions of recruitment for some major fishery resources using ocean remote sensing data. Further, the results suggest that it may be useful to consider multiple model inference techniques (e.g., Burnham and Anderson 2002) to better understand the factors affecting recruitment dynamics and to formulate better predictive models. In the future, this research approach may have practical application to Pacific marine fisheries where remotely-sensed oceanographic conditions may also be expected to influence recruitment strength, e.g., salmonids and tunas. Nonetheless, this is by no means a synoptic study of the potential factors influencing the recruitment strength of Georges Bank cod and haddock and further work to integrate existing research on this topic is clearly warranted.

Acknowledgments
Many thanks to Dave Foley for providing programs to access satellite data products and to Lyn Wagatsuma for help with the graphics.

References

Brodziak, J., Rago, P., and Conser, R.. 1998. A general approach for making short-term stochastic projections from an age-structured fisheries assessment model. In F. Funk, T. Quinn II, J. Heifetz, J. Ianelli, J. Powers, J. Schweigert, P. Sullivan, and C.-I. Zhang (Eds.), *Proceedings of the International Symposium on Fishery Stock Assessment Models for the 21st Century*. Alaska Sea Grant College Program, Univ. of Alaska, Fairbanks.

Brodziak, J. K. T., Overholtz, W. J., and Rago, P. J.. 2001. Does spawning stock affect recruitment of New England groundfish? Can. J. Fish. Aquat. Sci. 58:306-318.

Brodziak, J. 2005. Essential fish habitat source document: Haddock, *Melanogrammus aeglefinus*, life history and habitat characteristics, 2nd edition. NOAA Tech Mem. NMFS-NE-196, 74 pp. Available at: http://www.nefsc.noaa.gov/nefsc/publications/tm/tm196/

Brodziak, J., and O'Brien, L. 2005. Do environmental factors affect recruits per spawner anomalies of New England groundfish? ICES J. Mar. Sci. 62:1394-1407.

Brodziak, J., Traver, M., Col, L., and Sutherland, S. 2006. Stock assessment of Georges Bank haddock, 1931-2004. NEFSC Ref. Doc. 06-11. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0611/

Brodziak, J., and Link, J. 2008. The incredible shrinking Georges Bank haddock (Melanogrammus aeglefinus). Proceedings of the International Symposium on Resiliency of Gadid Stocks to Fishing and Climate Change. Alaska Sea Grant College Program, Univ. of Alaska, Fairbanks. In G.Kruse, K. Drinkwater, J. Ianelli, J. Link, D. Stram, V. Wespestad, and D. Woodby (Eds.), Resiliency of gadid stocks to fishing and climate change. Alaska Sea Grant College Program, AK-SG-08-01, Univ. of Alaska, Fairbanks.

Burnham, K., and Anderson, D. 2002. Model selection and multimodel inference, 2nd ed. Springer-Verlag, New York, 488 p.

Campana, S., and Hurley, P. 1989. An age- and temperature-mediated growth model for cod and haddock in the Gulf of Maine. Can. J. Fish. Aquat. Sci. 46:603-613

Friedland, K., Hare, J., Wood, G., Col, L., Buckley, L., Mountain, D., Kane, D., Brodziak, J., Lough, G., and Pilskaln, C. 2008. Does the fall phytoplankton bloom control recruitment of Georges Bank haddock, Melanogrammus aeglefinus, through parental condition? Can. J. Fish. Aquat. Sci. 65:1076-1086.

Green, J., Jones, R., and Brownwell, S. 2004. Age and growth of larval cod and haddock on Georges Bank during 1995 and 1996. Mar Ecol. Prog. Ser. 283: 255-268.

Greene, K. 2002. Bigger populations needed for sustainable harvest. Science (Wash. D.C.) 296: 1229-1230.).

Hastie, T., and Tibshirani, R. 1990. Generalized additive models. Chapman and Hall, NY, 335 p.

Laurence, G.C., and Rogers, C. A. 1976. Effects of temperature and salinity on comparative embryo development and mortality of Atlantic cod and haddock. J. Cons. Int. Explor. Mer. 36:220-228.

Mayo, R.K. and Terceiro, M., editors. 2005. Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, 15-19 August 2005. U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 05-13; 499 p. Available at: http://www.nefsc.noaa.gov/publications/crd/crd0513/index.htm

Mountain, D., Berrien, P., and Sibunka, J. 2003. Distribution, abundance and mortality of cod and haddock eggs and larvae of Georges Bank in 1995 and 1996. Mar. Ecol. Prog. Ser. 263:247-260.

Mountain, D., Manning, J., O' Brien, L., and Brodziak, J. 2006. Wind-driven transport indices for cod and haddock recruitment on Georges Bank. NOAA Fisheries and the Environment Program, 2006 Final Report, 2 p. Available at: http://fate.nmfs.noaa.gov/

New England Fishery Management Council [NEFMC]. 2004. Final amendment 13 to the Northeast Multispecies Fishery Management Plan including a final supplemental environmental impact statement and an initial regulatory flexibility analysis. New England Fishery Management Council, 50 Water Street, Mill 2 Newburyport, MA 01950. Available at: http://www.nefmc.org/nemulti/index.html

NEFMC. 2007. Framework adjustment 42 to the Northeast Multispecies Fishery Management Plan and framework adjustment 3 to the Monkfish Fishery Management Plan. New England Fishery Management Council, 50 Water Street, Mill 2 Newburyport, MA 01950. Available at: http://www.nefmc.org/nemulti/index.html

Northeast Fisheries Science Center [NEFSC]. 2002. Final Report of the Working Group on Re Evaluation of Biological Reference Points for New England Groundfish. NEFSC Ref. Doc. 02 04, p. 254. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0204/

NEFSC. 2008a. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0815/

NEFSC. 2008b. Appendix to the Report of the 3rd Groundfish Assessment Review Meeting (GARM III): Assessment of 19 Northeast Groundfish Stocks through 2007, Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-16; 1056 p. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0816/

O'Brien, L., Shepherd, N., and Col., L. 2006. Assessment of the Georges Bank Atlantic cod stock for 2005. NEFSC Ref. Doc. 06-10. Available at: http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0610/

Overholtz, W., Murawski, S., Rago, P., Gabriel, W., Terceiro, M., and Brodziak, J. 1999. Ten-year projections of landings, spawning stock biomass, and recruitment for five New England groundfish stocks. NEFSC Ref. Doc. 99-05. Woods Hole, MA, 74 p.

Page, F., and Frank, K. 1989. Spawning time and egg stage duration in northwest Atlantic haddock stocks with emphasis on Georges and Browns Bank. Can. J. Fish. Aquat. Sci. 46(Suppl. 1):68-81.

Page, F., Sinclair, M., Naimie, C., Loder, J. Losier, R., Berrien, P. and Lough, G. 1999. Cod and haddock spawning on Georges Bank in relation to water residence times. Fish. Oceangr. 8(3):212-226.

Payne, M., Dickey-Collas, M. and Nash, R. In press. Comment on "Does the fall phytoplankton bloom control recruitment of Georges Bank haddock, Melanogrammus aeglefinus, through parental condition?". Can. J. Fish. Aquat. Sci.

Rideout, R., Trippel, E., and Litvak, M. 2005. Effects of egg size, food supply and spawning time on early life history success of haddock *Melanogrammus aeglefinus*. Mar. Ecol. Prog. Ser. 285:169-180.

Stige, L. C., Ottersen, G., Brander, K., Chan, K., and Stenseth, N. 2006. Cod and climate: effect of the North Atlantic Oscillation on recruitment in the North Atlantic. Mar. Ecol. Prog. Ser. 325:227-241.

Table 1. Projection data for Georges Bank cod and haddock.

Georges Bank cod projection data

Age	Average Spawning Weight at Age(kg)	Average Landed Weight at Age (kg)	Fraction Mature at Age	Natural Mortality at Age	Fishery Selectivity at Age
1	0.416	0.593	0.07	0.2	0
2	1.060	1.886	0.34	0.2	0.06
3	2.100	2.401	0.79	0.2	0.57
4	2.699	3.098	0.96	0.2	0.8
5	3.593	4.137	1	0.2	1
6	4.663	5.110	1	0.2	1
7	5.747	6.294	1	0.2	1
8	7.244	8.063	1	0.2	1
9	8.742	9.408	1	0.2	1
10+	11.629	11.697	1	0.2	1

Georges Bank haddock projection data

Age	Average Spawning Weight at Age(kg)	Average Landed Weight at Age (kg)	Fraction Mature at Age	Natural Mortality at Age	Fishery Selectivity at Age
1	0.186	0.258	0.01	0.20	0.01
2	0.445	0.650	0.60	0.20	0.09
3	1.056	1.232	0.95	0.20	0.21
4	1.495	1.615	1	0.20	0.62
5	1.761	1.823	1	0.20	1.00
6	2.070	2.158	1	0.20	1
7	2.455	2.572	1	0.20	1
8	2.814	2.876	1	0.20	1
9+	3.139	3.139	1	0.20	1

Table 2. Pair wise correlations of environmental covariates across three regional scales (R1, R2, and R3) including the autumn primary productivity index (PP.fall), spring primary productivity indices during the cod (PP.spr.fm) and haddock (PP.spr.mm) spawning seasons, sea surface temperature indices during the cod (ST.spr.fm) and haddock (ST.spr.mm) spawning seasons.

Environmental covariate	Correlation for (R1, R2)	Correlation for (R1, R3)	Correlation for (R2, R3)
PP.fall	0.91	0.86	0.97
PP.spr.fm	0.96	0.91	0.95
PP.spr.mm	0.95	0.86	0.87
ST.spr.fm	0.93	0.94	0.96
ST.spr.mm	0.72	0.74	0.84

Table 3. Projection results for Georges Bank cod.

Georges Bank cod out-of-sample recruitment predictions

		Average Annual Recruitment Estimate by Model						
Year Class	GB Cod 2008 VPA Recruitment Estimates (000000s)	$M_{COD,SQ}$	$M_{COD,RS1}$	$M_{COD,RS2}$	$M_{COD,RS3}$	$M_{COD,R1}$	$M_{COD,R2}$	$M_{COD,R3}$
2005	6.490	9.031	10.176	8.002	10.948	13.096	7.176	13.936
2006	7.037	12.563	18.001	11.676	15.964	13.705	8.644	11.533
2007	4.875	14.799	23.597	16.746	23.283	13.379	7.835	12.774

Year Class	Average Recruitment Prediction Error by Model						
2005	2.541	3.686	1.512	4.458	6.606	0.686	7.446
2006	5.526	10.964	4.639	8.927	6.668	1.607	4.496
2007	9.924	18.722	11.871	18.408	8.504	2.960	7.899
Sum of Squared APE	135.481	484.315	164.719	438.412	160.403	11.816	138.050
Mean-Square Error	45.160	161.438	54.906	146.137	53.468	3.939	46.017
Root Mean-Square Error	6.720	12.706	7.410	12.089	7.312	1.985	6.784
Percent Change in RMSE	**0.0%**	**89.1%**	**10.3%**	**79.9%**	**8.8%**	**-70.5%**	**0.9%**

Table 4. Projection results for Georges Bank haddock.

Georges Bank haddock out-of-sample recruitment predictions

Year Class	GB Haddock 2008 VPA Recruitment Estimates (000000s)	Average Annual Recruitment Estimate by Model $M_{HAD,SQ}$	$M_{HAD,RS1}$	$M_{HAD,R1}$	$M_{HAD,R2}$	$M_{HAD,R3}$	$M_{HAD,R4}$	$M_{HAD,R5}$
2005	26.45	52.373	45.197	12.399	15.251	11.722	25.908	19.106
2006	7.421	52.556	68.483	13.850	18.389	12.206	19.370	16.655
2007	16.376	52.604	84.096	13.049	23.685	11.603	19.132	16.106

Year Class	Average Recruitment Prediction Error by Model						
2005	25.923	18.747	-14.051	-11.199	-14.728	-0.542	-7.344
2006	45.135	61.062	6.429	10.968	4.785	11.949	9.234
2007	36.228	67.720	-3.327	7.309	-4.773	2.756	-0.270
Sum of Squared APE	4021.682	8665.994	249.823	299.148	262.589	150.672	139.279
Mean-Square Error	1340.561	2888.665	83.274	99.716	87.530	50.224	46.426
Root Mean-Square Error	36.614	53.746	9.125	9.986	9.356	7.087	6.814
Percent Change in RMSE	**0%**	**46.8%**	**-75.1%**	**-72.7%**	**-74%**	**-80.6%**	**-81.4%**

Figure 1. Northeast Fisheries Science Center commercial fishery statistical areas for western Georges Bank along with U.S. and Canada shared management area on eastern Georges Bank. Large-scale closed areas are: Closed Area I (CAI), Closed Area II (CAII), western Gulf of Maine Closed Area (WGOM CA) and Nantucket Lightship Closed Area (Nantucket Lightship CA).

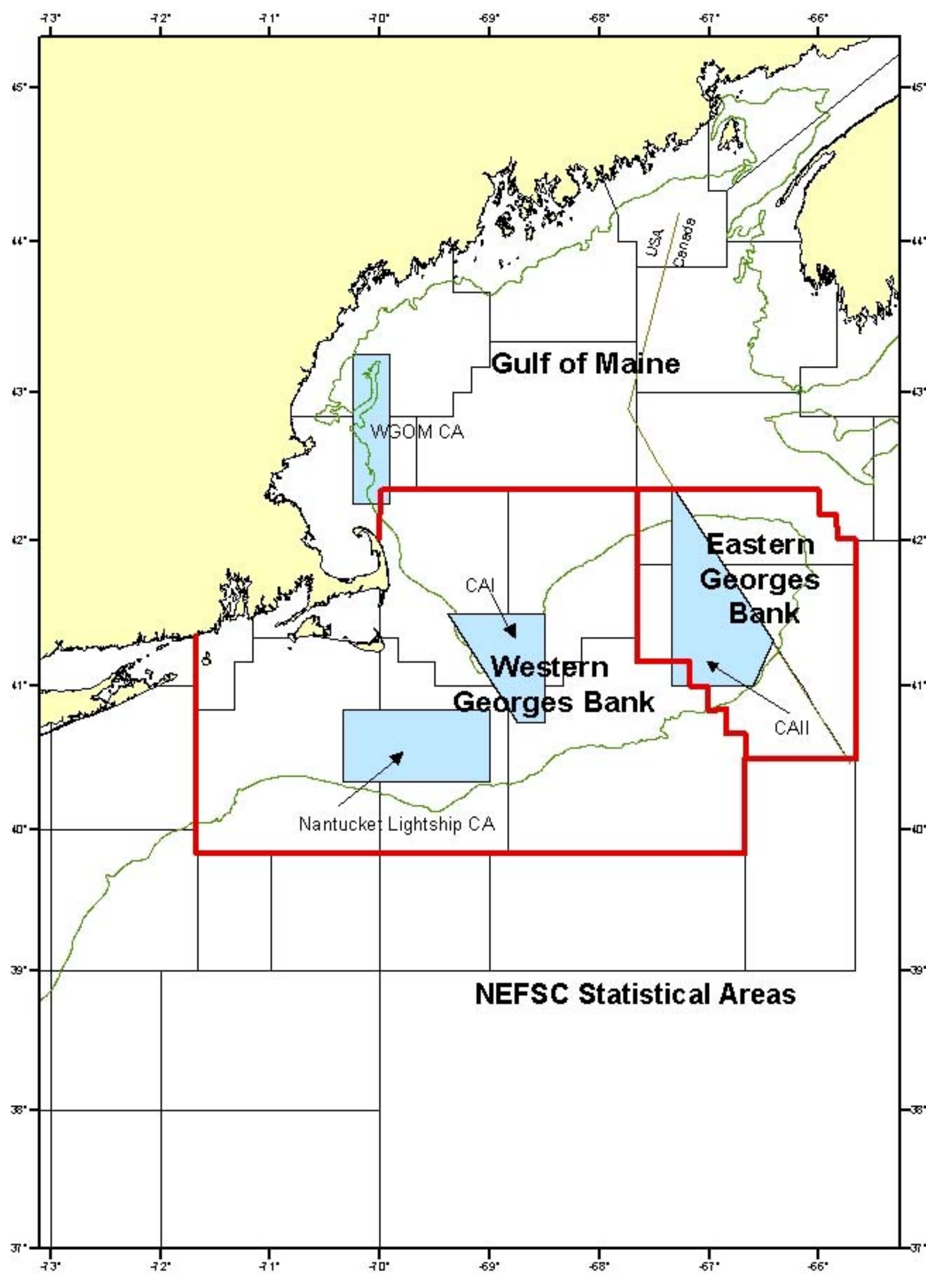


Figure 2. Recruitment (open circle) and recruits per spawner (solid) as a function of spawning biomass during 1985-2004 for Georges Bank cod (a) and haddock(b) taken from O'Brien et al. (2006) and Brodziak et al. (2006).

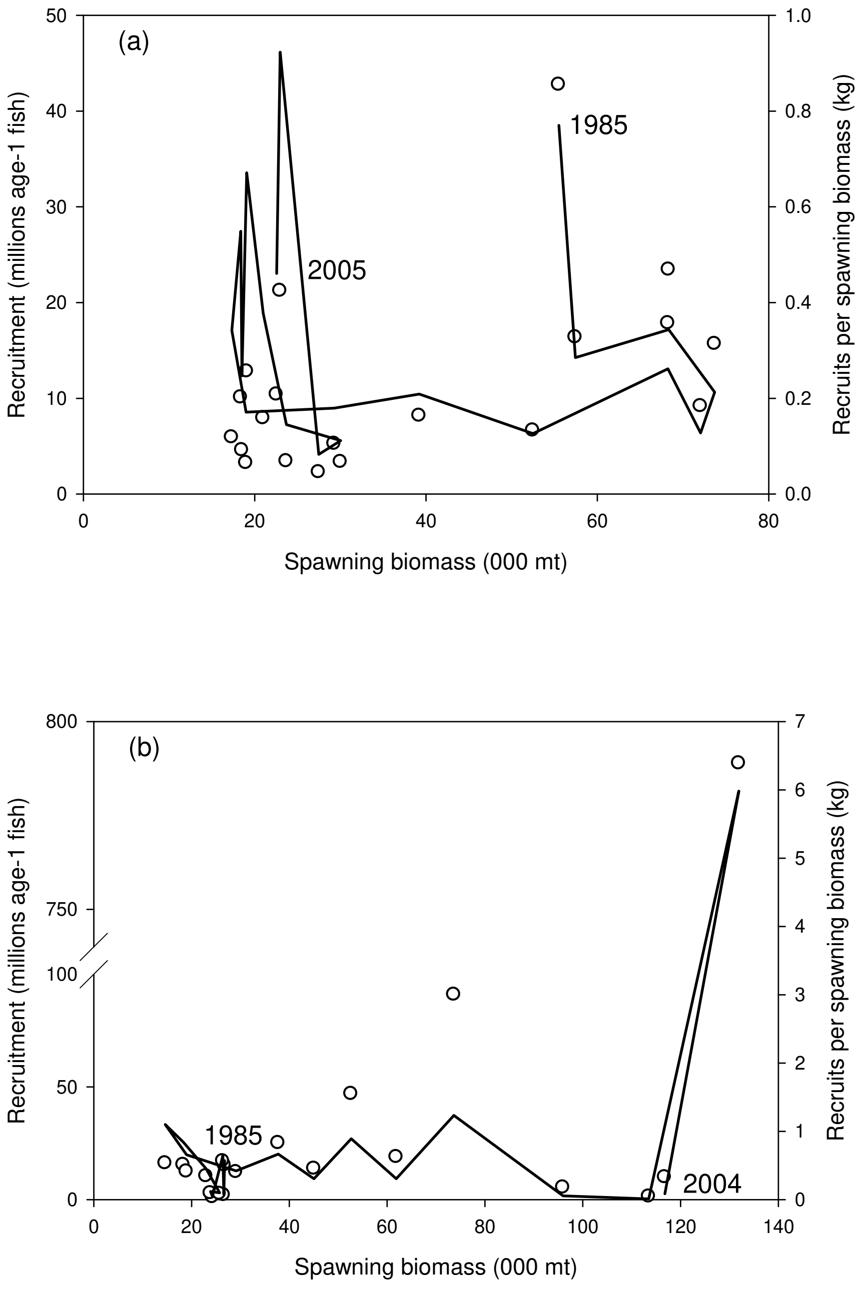


Figure 3. Location of three regions used to summarize satellite-derived environmental covariates: the Georges Bank superregion R1 (dashed blue line) with upper-left and bottom-right coordinates of (43˚N, -70˚W) and (40˚N, -66˚W); the primary Georges Bank region R2 (dotted red line) with upper-left and bottom-right coordinates of (42˚N, -69˚W) and (40˚N, -66˚W); and the Georges Bank central subregion R3 (solid green line) with upper-left and bottom-right coordinates of (42˚N, -68˚W) and (41˚N, -67˚W).

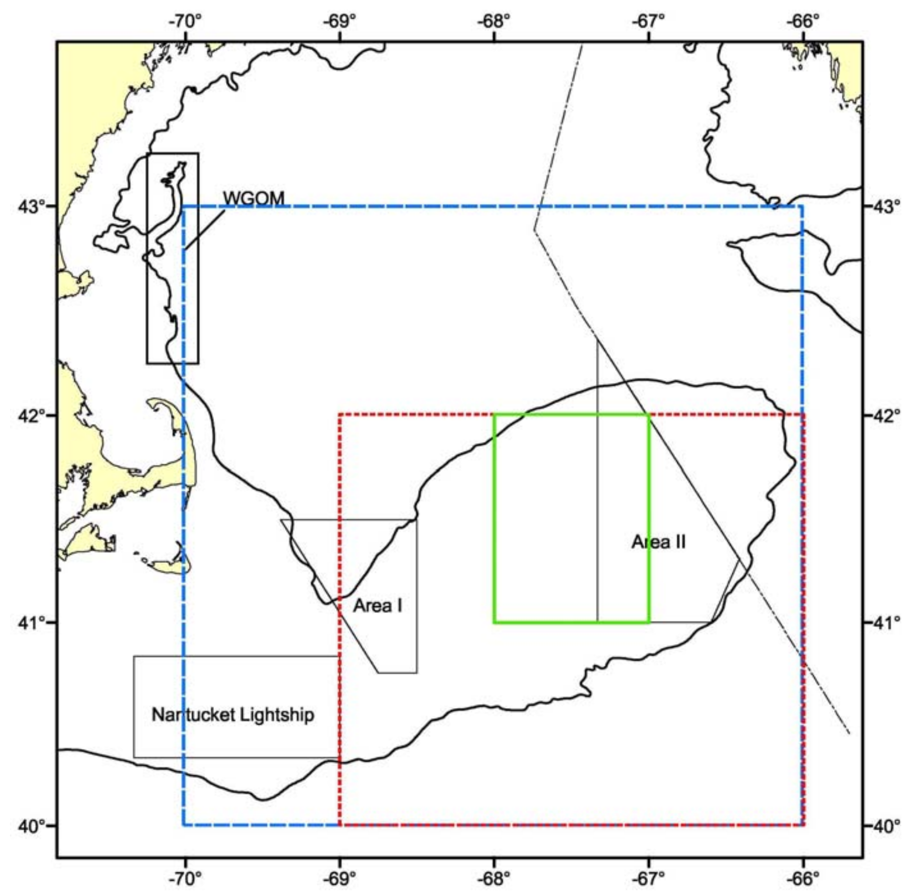


Figure 4. Primary productivity indices (a) and sea surface temperature indices (b) for use as environmental covariates in recruitment analyses for Georges Bank cod and haddock.

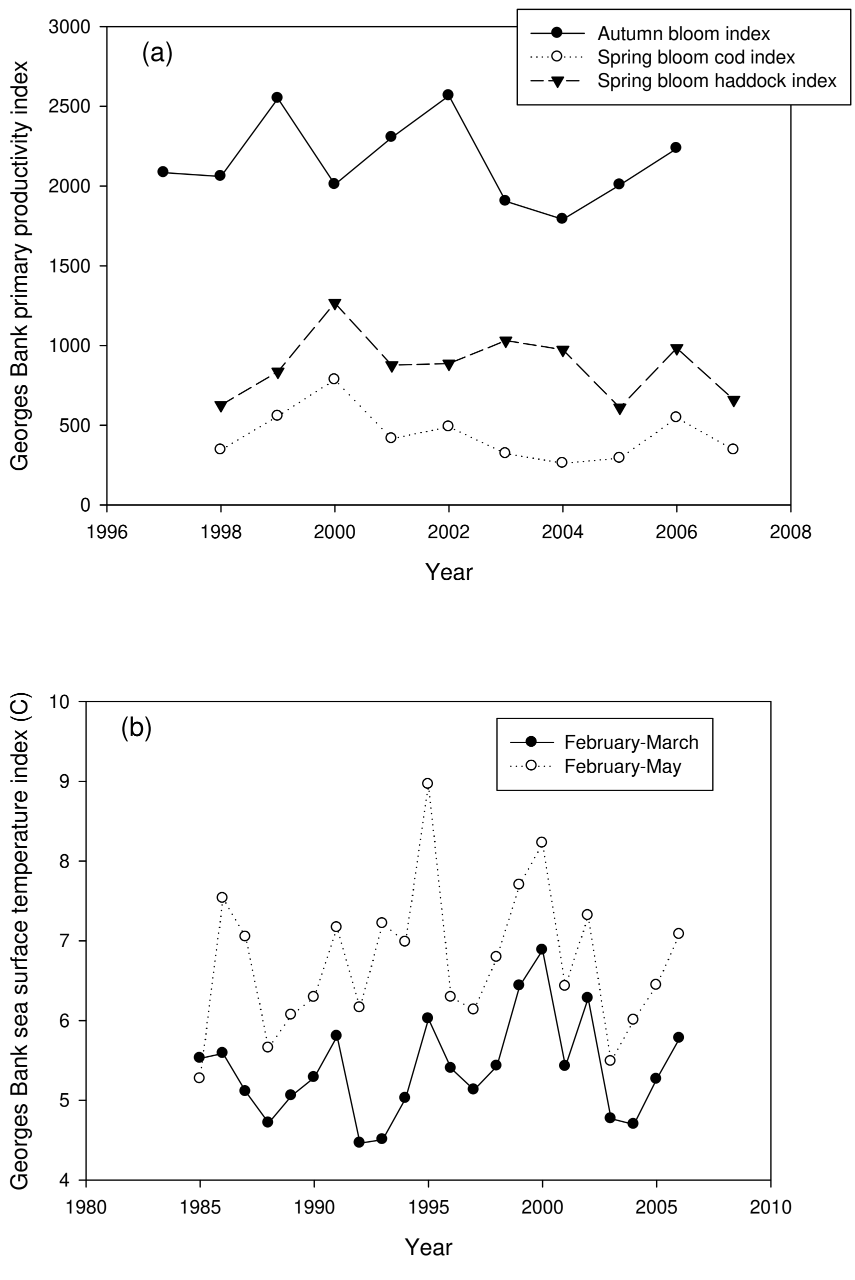
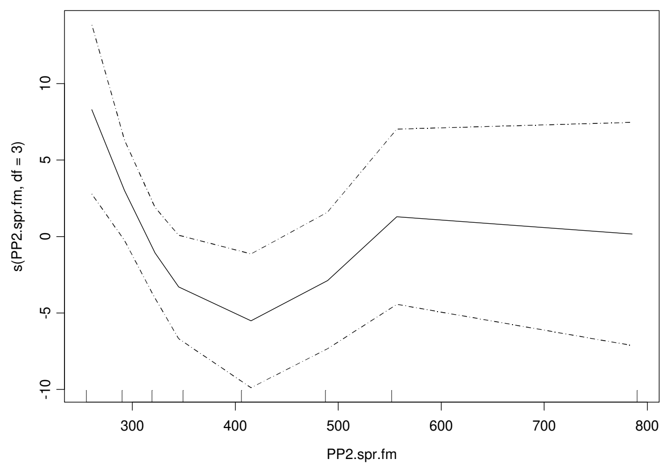


Figure 5. Effect of spring primary productivity on cod recruitment (a) and recruits per spawning as represented by a smoothed GAM function (solid line) with approximate 95% confidence intervals (dashed lines).

(a) Cod recruitment



(b) Cod recruits per spawner

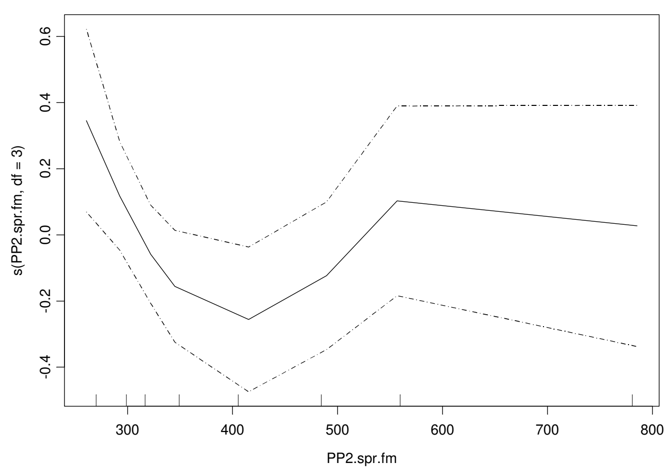
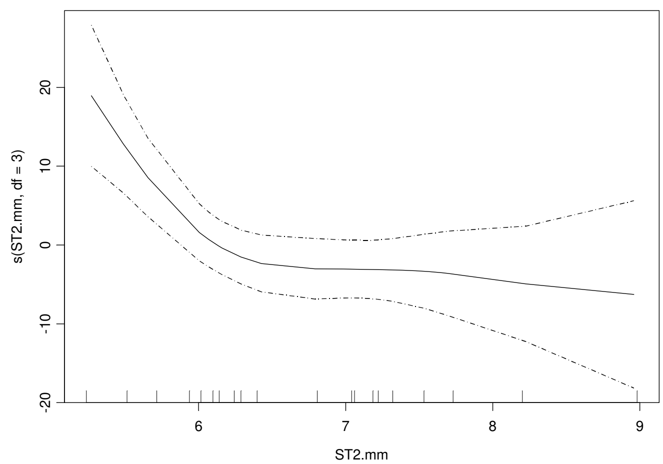


Figure 6. Effect of spring sea surface temperature on cod recruitment (a) and recruits per spawning as represented by a smoothed GAM function (solid line) with approximate 95% confidence intervals (dashed lines).

(a) Cod recruitment



(b) Cod recruits per spawner

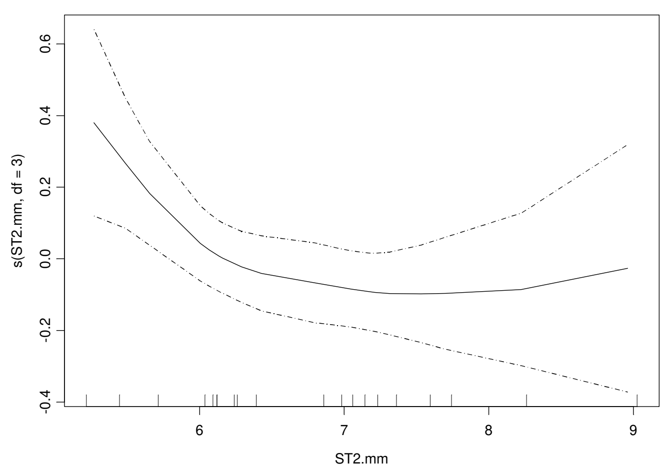


Figure 7. Effect of autumn primary productivity on log-scale haddock recruits per spawner lagged one year as represented by a smoothed GAM function (solid line) with approximate 95% confidence intervals (dashed lines).

Haddock log-scale recruits per spawner

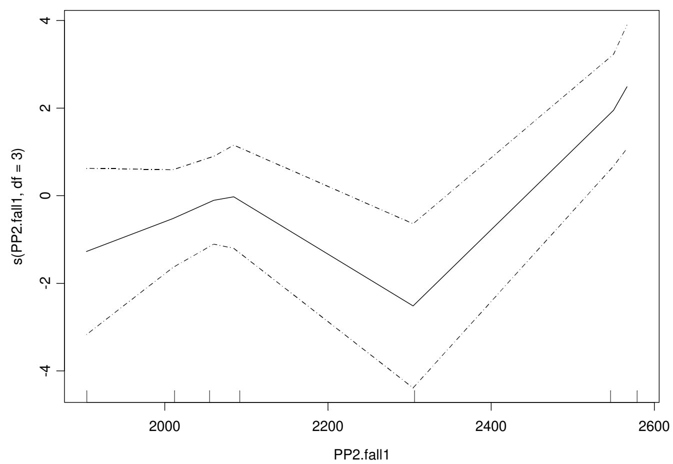


Figure 8. Correlation of Georges Bank cod recruitment (solid circle, solid line) and recruits per spawner (open triangle and dashed line) during 1985-2004 from O'Brien et al. (2006) with average sea surface temperature on Georges Bank during February-May taken from Pathfinder v5 composite monthly data.

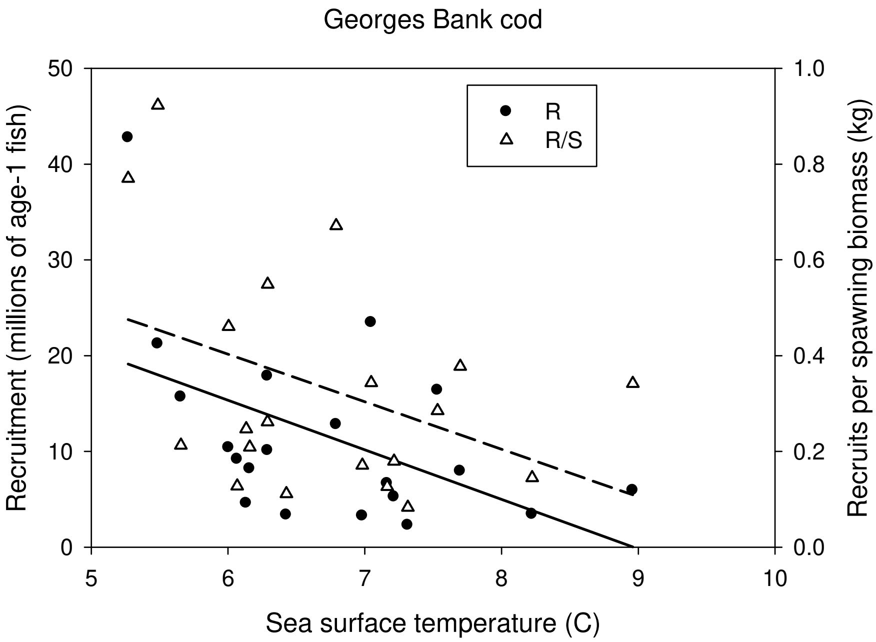


Figure 9. Correlation of Georges Bank haddock recruitment (solid circle, solid line) and recruits per spawner (open triangle and dashed line) during 1985-2004 from Brodziak et al. (2006) with average primary productivity on Georges Bank during September-November the previous year derived from Seawifs composite monthly data (http://coastwatch.pfel.noaa.gov/).

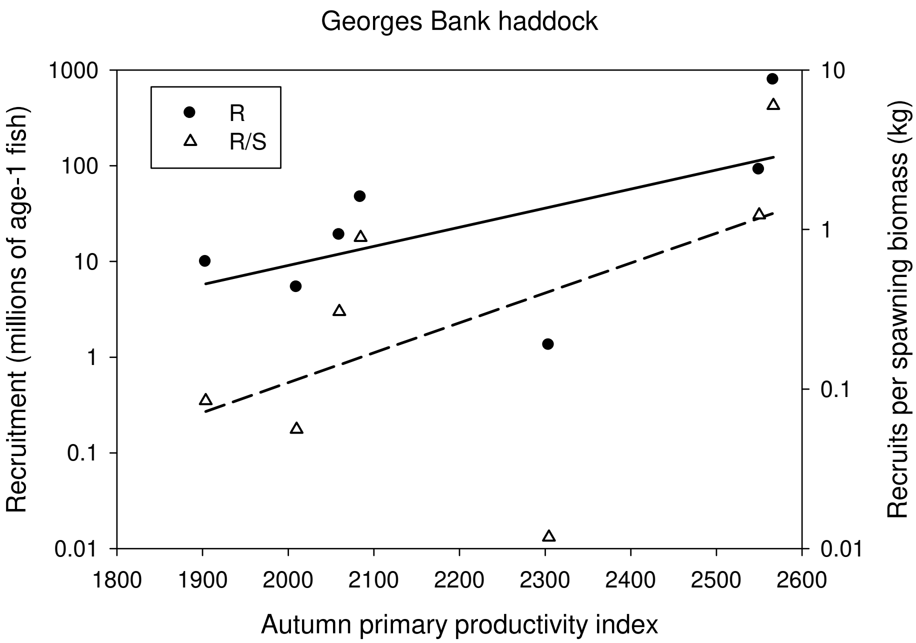


Figure 10. Comparison of Georges Bank observed recruitment (solid circle) during 2005-2007 (NEFSC 2008a) and rescaled recruitment predictions from the best predictive model using average sea surface temperature (SST) during February-May (open circle) and the status quo model (solid triangle) from Mayo and Terceiro (2006) along with 80% confidence intervals for the SST prediction.

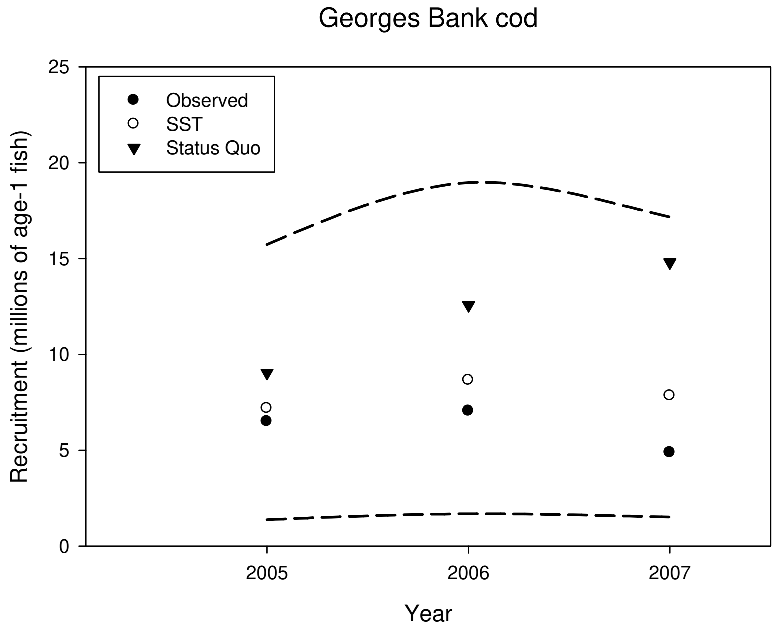
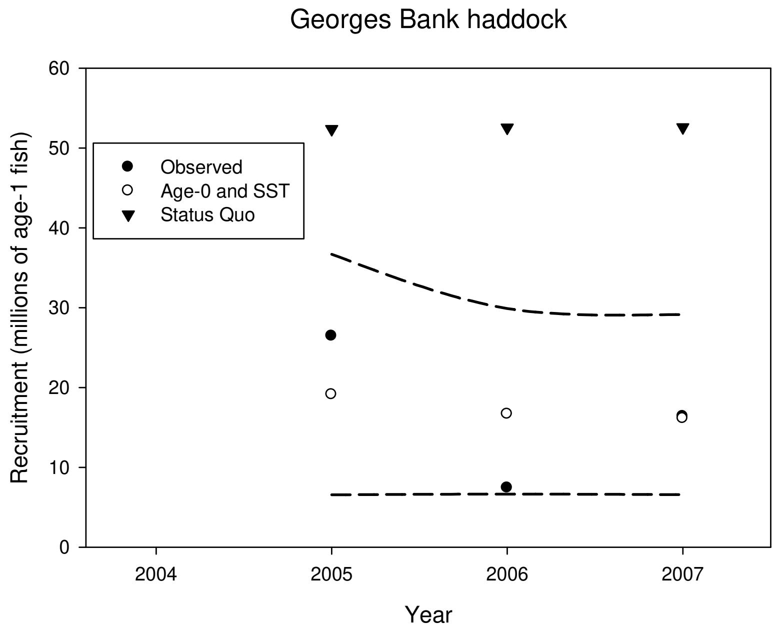


Figure 11. Comparison of Georges Bank haddock observed recruitment (solid circle) during 2005-2007 (NEFSC 2008a) and rescaled recruitment predictions from the best predictive model (open circle), a model-averaged combination of predictors using the haddock age-0 survey index and average sea surface temperature (SST) during February-May, and the status quo model (solid triangle) from Mayo and Terceiro (2006) along with 80% confidence intervals for the Age-0 index and SST-based prediction.



Appendix. Satellite data extraction program “xtracto_ts_bdap” developed by SWFSC’s Environmental Research Division. Basic requirements for running the program are a fully-licensed version of the Matlab software and a high-speed internet connection.

```
function [extract] = xtracto_ts_bdap(xpos,ypos,tpos,dtype);
% Example script to get the last two days of data from set
% help can be gotten from the command "help loaddods"
%
% INPUTS:  xpos = longitude (in decimal degrees East, either 0-360 or -180 to 180)
%          ypos = latitude (in decimal degrees N; -90 to 90)
%          tpos = time (preferably in matlab [Julien] days)
%          dtype = data ID Code (data types listed below)
%
% OUTPUT:
%
%  Extract = 4 column array
%            column 1 = mean of data within search radius
%            column 2 = standard deviation of data within search radius
%            column 3 = number of points found withn search radius
%            column 4 = actual time associated with satellite data set.
%
%
% Sample Calls:
%
% to extract Seawifs 8-day Primary Productivity for a given box
%   surrounding each available time point
% [extract] = xtractomatic([xmin xmax],[ymin ymax],[tmin tmax],'41');
%
%  to extract pathfinder SST 8-day mean data
%  [extract] = xtractomatic([xmin xmax],[ymin ymax],[tmin tmax],'18',.1,.1);
%
% Note: xtracto_ts does not interpolate onto requested time points - it merely finds all of those
% available within the specified bounds.  To capture the entire data set, use something like
% [extract] = xtracto_bdap_ts([xmin xmax],[ymin ymax],[datenum(1950,1,1) datenum(2050,1,1)],'18');
%
% %
% see the following link to get data codes and full data set information
% http://coastwatch.pfel.noaa.gov/coastwatch/CWBrowserWW360.jsp?get=griddata
%
% V0.1  17 Aug 2006.
% CoastWatch/DGF

% add my own path where I keep my auxiliary m-files
```

```
% path(path,'/home/cwatch/mfiles');

% make sure data type input is a string and not a number
if ~isstr(dtype)
  datatype = num2str(dtype);
else
  datatype = dtype;
end

switch lower(dtype)
     % AVHRR HRPT 1.4 km nighttime SST data for West coast
     case {'1','atsstnhday'}
       dataset = 'TATsstdhday';

     % AVHRR HRPT 1.4 km daytime SST data for West coast
     case {'2','atssdhday'}
       dataset = 'TATsstnhday';

     % AVHRR HRPT 1,4 km night and day SST 1-day composite
     case {'3','atssta1day','hrpt','avhrr hrpt'}
       dataset = 'TATssta1day';

     % AVHRR HRPT 1.4 km night and day SST 3-day composite
     case{'4','atssta3day'}
       dataset = 'TATssta3day';

     % AVHRR HRPT 1.4 km night and day SST 8-day composite
     case{'5','atssta8day'};
       dataset = 'TATssta8day';

     % AVHRR HRPT 1.4 km night and day SST 14-day composite
     case{'6','atssta14day'}
       dataset = 'TATssta14day';

     % AVHRR HRPT 1.4km night and day SST monthly composite
     case{'7','atsstamday'}
        dataset = 'TATsstamday';

     % AVHRR GAC SST 11km 1-day composite
     case{'8','agssta1day'}
       dataset = 'TAGssta1day';

     % AVHRR GAC SST 11km 3-day composite
```

```
case{'9','agssta3day'}
  dataset = 'TAGssta3day';

% AVHRR GAC SST 11km 8-day composite
case{'10','agssta8day'}
  dataset = 'TAGssta8day';

% AVHRR GAC SST 11km 14-day composite
case{'11','agssta14day'}
  dataset = 'TAGssta14day';

% AVHRR GAC SST 11km monthly composite
case{'12','agsstamday'}
  dataset = 'TAGsstamday';

% GOES SST 5.5 km 1-day composite
case{'13','gassta1day','goes sst','goes','geostationary'}
  dataset = 'TGAssta1day';

% GOES SST 5.5 km 3-day composite
case{'14','gassta3day'};
  dataset = 'TGAssta3day';

% GOES SST 5.5 km 8-day composite
case{'15','gassta8day'}
  dataset = 'TGAssta8day';

% GOES SST 5.5 km 14-day composite
case{'16','gassta14day'}
  dataset = 'TGAssta14day';

% Pathfinder v5 5.5km SST 1-day composite
case{'17','phssta1day'}
  dataset = 'TPHssta1day';

% Pathfinder v5 5.5km SST 8-day composite
case{'18','phssta8day','pathfinder','sst','avhrr','sea surface temperature'}
 dataset = 'TPHssta8day';

% Pathfinder v5 5.5km SST monthly composite
case{'19','phsstmday','monthly sst'}
 dataset = 'TPHsstamday';
```

```
% MODIS Aqua 2.5 km chla 1-day composite
case{'20','mbchla1day'}
 dataset = 'TMBchla1day';

% MODIS Aqua 2.5 km chla 8-day composite
case{'21','mbchla8day','chla','chlorophyll','modis chl','modis aqua'};
 dataset = 'TMBchla8day';

% MODIS Aqua 2.5 km chla 14-day composite
case{'22','mbchla14day'}
 dataset = 'TMBchla14day';

% Jason-1 25km SSH deviation, 10-day composite
case{'23','j1sshd10day','ssh','ssha','jason','sea surface height'}
  dataset = 'TJ1sshd10day';

% Quikscat 25 km zonal wind, 1-day composite
case{'24','qnux101day'}
  dataset = 'TQNux101day';

% Quikscat 25 km meridional wind, 1-day composite
case{'25','qnuy101day'}
 dataset = 'TQNux101day';

% Quikscat 25 km zonal wind, 3-day composite
case{'26','qnux103day','zonal wind','ux10'}
  dataset = 'TQNux103day';
case{'27','qnuy103day','meridional wind','uy10'}
  dataset = 'TQNuy103day';
case{'28','qnumod3day','wind speed','wind modulus'}
  dataset = 'TQNumod3day';
case{'29','qncurl3day','curl','wind stress curl','curl of wind stress'}
  dataset = 'TQNcurl3day';
case{'30','qnux108day'}
  dataset = 'TQNux108day';
case{'31','qnuy108day'}
  dataset = 'TQNuy108day';
case{'32','qnumod8day'}
  dataset = 'TQNumod8day';
case{'33','qncurl8day'}
  dataset = 'TQNcurl8day';
case{'34','qnux1014day'}
  dataset = 'TQNux1014day';
```

```
case{'35','qnuy1014day'}
  dataset = 'TQNuy1014day';
case{'36','qnumod14day'}
  dataset = 'TQNumod14day';
case{'37','qncurl14day'}
  dataset = 'TQNcurl14day';
case{'38','qncurlmday'}
  dataset = 'TQNcurlmday';
case{'39','qnux10mday'}
  dataset = 'TQNux10mday';
case{'40','qnuy10mday'}
  dataset = 'TQNux10mday';

% Primary productivity, 8-day, seawifs chl.
case{'41','ppbfp18day','primary productivity','seawifs productivity'}
  dataset = 'TPPbfp18day';

% Primary productivity, monthly, seawifs chl
case{'42','Tppbfp1mday','monthly productivity'}
  dataset = 'TPPbfp1mday';

% GOES frontal index 14-day
case{'43','gatfnt14day','GOES fronts','frontal index','frontal probability'}
 dataset = 'TGAtfnt14day';

% GOES frontal index 14-day
case{'44','gatfntmday'}
  dataset = 'TGAtfntmday';

case{'45','qncur4day'}
  dataset = 'TQNcurl4day';
case{'46','qnux104day'}
  dataset = 'TQNux104day';
case{'47','qnuy104day'}
  dataset = 'TQNuy104day';
case{'48','tasshd1day'}
  dataset = 'TTAsshd1day';
case{'49','bassta5day'}
  dataset = 'TBAssta5day';
case{'50','mhchla8day'}
  dataset = 'TMHchla8day';
case{'51','mhk4908day'}
  dataset = 'TMHk4908day';
```

```
      case{'52','mhsstd8day'}
        dataset = 'TMHsstd8day';
      case{'53','mhcflh8day'}
        dataset = 'TMHcflh8day';
      case{'54','qscurl3day'};
        dataset = 'TQScurl3day';
      case{'55','qsumod3day'}
        dataset = 'TQSumod3day';
       case{'56','qsux103day'};
        dataset = 'TQSux103day';
      case{'57','qsuy103day'}
        dataset = 'TQSuy103day';
  end

% breakup string into components
satid = dataset(1:3);
param = dataset(4:7);
duration = dataset(8:end);

% correct for Bob/Dave discontinuity
if strcmp(duration,'hday')
    duration = '1observation';
end

if strcmp(duration,'mday')
    duration = '1month'
end

% default URL for NMFS/SWFSC/ERD  THREDDS server
urlbase='http://coastwatch.pfel.noaa.gov/coastwatch/CWBrowserWW360.jsp';

% get list of available time periods
% First, make bad call to CW page,
bobcallbad = strcat(urlbase,'?get=griddata&dataset=',strcat(satid,param),...
               '&timeperiod=',duration,...
               '&centeredTime=')
string=urlread(bobcallbad);

% sift through the crap for ISO dates
stind = regexp(string,'<nobr>','start');
endind = regexp(string,'</nobr>','end');
icnt = 1;
for i = 1:length(stind),
```

```
    str2 = string(stind(i):endind(i));
    stind2 = regexp(str2,'\d\d\d\d\-\d\d-\d\d','start');
    endind2 =  regexp(str2,'\d\d\d\d\-\d\d-\d\d','end');
    if ~isempty(stind2)
      for j = 1:length(stind2),
        alldates(icnt,1:10) = str2(stind2(j):endind2(j));
        icnt = icnt + 1;
      end
    end
end
btime = unique(alldates,'rows')

% convert to matlab days
year=str2num(btime(:,1:4));
month=str2num(btime(:,6:7));
day=str2num(btime(:,9:10));
sattime = datenum(year,month,day);

% handle case of -180 to 180 longitude
ind=find(xpos<0);
xpos(ind)=xpos(ind)+360;

% find all time points within tmin, tmax
  tmin = min(tpos);
tmax = max(tpos);
tind = find(sattime>=tmin & sattime<=tmax);

% define bounding box
xmax = max(xpos)
xmin = min(xpos)
ymax = max(ypos)
ymin = min(ypos)

% loop on available time slots

if ~isempty(tind)
  for i = 1:length(tind),

    % get time in year month day
    yrstr = datestr(sattime(tind(i)),'YYYY');
    monstr = datestr(sattime(tind(i)),'mm');
    daystr = datestr(sattime(tind(i)),'dd');
```

```
    % text string for data retrieval call
    bobcall = strcat(urlbase,'?get=griddata&dataset=',strcat(satid,param),...
               '&minlon=',num2str(xmin),'&maxlon=',num2str(xmax),...
               '&minlat=',num2str(ymin),'&maxlat=',num2str(ymax),...
               '&timeperiod=',duration,...
               '&centeredTime=~',yrstr,'-',monstr,'-',daystr,...
               '&filetype=.mat')

    % extract data array and import to Matlab depending on structure
    varname = strcat(satid(2:3),param);
    fileout='tmp.mat';
    urlwrite(bobcall,fileout);
    try
      load('-MAT',fileout);
      eval(strcat('sstd=',varname,';'));

      mean(sstd(find(~isnan(sstd))))
      % get array dimensions - note that the order of data returned is not the same

      extract(i,1) = mean(sstd(find(~isnan(sstd))));
      extract(i,2) = std(sstd(find(~isnan(sstd))));
      extract(i,3) = length(sstd(find(~isnan(sstd))));
      extract(i,4) = sattime(tind(i));
    catch
      extract(i,1:4)=nan;
      sprintf('Caught a THREDDS access error. Pausing for 2 minutes...')
      pause(120)
    end
  end
else
  extract(1,1:4)=nan;
  sprintf('no valid points found within target parameters');
end

% fin
```