

## Climate–ocean variability and Pacific hake: A geostatistical modeling approach

V.N. Agostini<sup>a,\*</sup>, A.N. Hendrix<sup>b</sup>, A.B. Hollowed<sup>c</sup>, C.D. Wilson<sup>c</sup>,  
S.D. Pierce<sup>d</sup>, R.C. Francis<sup>a</sup>

<sup>a</sup> School of Aquatic and Fishery Science, University of Washington, Box 355020, Seattle WA 98149, USA

<sup>b</sup> R2 Resource Consultants, Inc., 15250 NE 95th Street, Redmond, WA 98052, USA

<sup>c</sup> National Marine Fisheries Service-AFSC, Sand Point Way, Seattle, WA 98143, USA

<sup>d</sup> Oregon State University-COAS, 104 COAS Admin Bldg, Corvallis, OR 97331, USA

Received 15 June 2006; received in revised form 18 December 2006; accepted 23 January 2007

Available online 13 November 2007

### Abstract

Climate forcing of the California Current has been known to impact the distribution and abundance of a number of local fish populations, but the mechanisms involved remain poorly understood. Climate metrics such as the Pacific Decadal Oscillation (PDO) and the El Niño Southern Oscillation (ENSO) are usually used to represent climate processes and direct links are made between climate forcing and production variability. This involves aggregation of impacts across large spatial scales and range of species. However, fluctuations in productivity are often the result of changes in physical habitat. In order to fully understand the relationship between climate and productivity, habitat changes should be addressed. In this study we use a geostatistical approach to quantify adult Pacific hake habitat during different climate regimes. Several authors have suggested that the distribution and intensity of the sub-surface poleward flow (the undercurrent) plays a key role in defining adult hake habitat along the west coast of North America. Here we build a model designed to predict hake habitat distribution in space based on sub-surface poleward flow distribution and bottom depth. Our results show that hake habitat expands in 1998 El Niño year compared to 1995. Given the important predatory role that hake plays in the CC, the amount and distribution of adult hake habitat has large implications for the Pacific Northwest food web and could thus serve as an ecosystem indicator representing important physical–biological interactions. Spatially based ecosystem indicators such as the one we develop here address two important yet neglected areas in the ‘Ecosystem Indicators debate’: the importance of developing metrics explicitly representing spatial and environmental processes shaping ecosystem structure. Without these, our power to fully describe ecosystems will be limited.

© 2007 Published by Elsevier B.V.

**Keywords:** Climate; Ecosystems; *Merluccius productus*; Habitat; Spatial distribution; USA; California Current

\* Corresponding author. Present address: Pew Institute for Ocean Science, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker way, Miami, FL 33133, USA. Tel.: +1 305 421 4165; fax: +1 305 421 4077.

E-mail addresses: [vagostini@rsmas.miami.edu](mailto:vagostini@rsmas.miami.edu) (V.N. Agostini), [nhendrix@r2usa.com](mailto:nhendrix@r2usa.com) (A.N. Hendrix), [Anne.hollowed@noaa.gov](mailto:Anne.hollowed@noaa.gov) (A.B. Hollowed), [Chris.Wilson@noaa.gov](mailto:Chris.Wilson@noaa.gov) (C.D. Wilson), [spierces@coas.oregonstate.edu](mailto:spierces@coas.oregonstate.edu) (S.D. Pierce), [bfrancis@u.washington.edu](mailto:bfrancis@u.washington.edu) (R.C. Francis).

## 1. Introduction

The production variability of a number of California Current (CC) fish species has been related to climate forcing (Ware and McFarlane, 1995; MacCall, 1996). Habitat variability is often invoked as the potential link between the two (Ware and McFarlane, 1995; MacCall, 1996; Benson et al., 2002), but the mechanisms involved remain poorly understood (Beamish et al., 2000). The use of climate metrics (i.e. Pacific Decadal Oscillation, El Niño Southern Oscillation) to represent potential environmental forcing on fish populations and ecosystems (e.g., Hare and Mantua, 2000) usually involves the aggregation of impacts across large spatial scales and range of species. Direct links are made between climate forcing and production variability. However, fluctuations in productivity are often the result of changes in physical habitat and food availability. In order to gain a complete understanding of the link between climate and production variability habitat changes should be addressed.

Habitat in terrestrial systems has long been interpreted as vegetation, sometimes with underlying gradients of moisture or soil chemistry (Rice, 2001). This definition seems to have been transferred to marine environments where classic definitions of marine habitat (e.g. rocky intertidal, kelp forests, coral reefs) usually involve vegetation or substrate types, which are all static features. While this is appropriate for benthic communities, it is not appropriate for organisms or animals that spend part of or their entire life-cycle in the pelagic zone. For these species habitat is often a dynamic entity, its boundaries changing according to time/space changes of the physical oceanographic variables defining it.

Understanding how pelagic habitats are distributed in space and how their characteristics vary will contribute to our understanding of ecosystem processes and the sustainability of fish populations (Kracker, 1999). Lately, researchers have placed more emphasis on recognizing the importance of spatial patterns in ecological processes (Petitgas, 1993; Horne and Schneider, 1995; Petitgas, 1998). Recent discussions regarding the long-term sustainability of ocean resources have focused on the need to understand the spatial distribution as well as the size of fish populations (Kracker, 1999). Quantifying these patterns has been recognized as an essential component of our research efforts to understand how harvest pressure and climate change impact the sustainability of fish stocks (Wiebe et al., 1996).

Individual fish often position themselves in response to a combination of features of the marine environment

(Maravelias et al., 1996) and in relation to other fish. Their distribution is not random, either in space or time, but rather organized in structures (schools, aggregations). Geostatistics is a branch of applied statistics that focuses on detecting, modeling and estimating spatial patterns (Rossi et al., 1992). This type of modeling approach assumes spatial dependence (the value at one location is conditioned by the values at neighboring locations) instead of spatial independence (values at one location are independent of values at neighboring locations). Spatial dependence is very important in ecology, “the scientific study of the relationships between organisms and their environments” (McNaughton and Wolf, 1973), yet traditional statistics typically fail to take spatial dependence into account (Rossi et al., 1992). During the last decade a number of studies have applied geostatistical techniques, which were initially developed for terrestrial applications, to marine systems (Petitgas, 2001). Their focus has been on a number of fishery problems, ranging from estimating abundances from survey data (Simard et al., 1992; Conan et al., 1994; Barange and Hampton, 1997; Fletcher and Summer, 1999; Barange et al., 2005) to quantifying relationships between environmental variables and fish distributions (Maravelias et al., 1996).

Pacific hake (*Merluccius productus*) is a commercially and ecologically important species in the CC System. It accounts for 61% of the pelagic biomass in the California Current system (Ware and McFarlane, 1995) affording it a key position as both predator and prey in the coastal food web (Livingston and Bailey, 1985). Although distinct populations of Pacific hake also exist in the Strait of Georgia (McFarlane and Beamish, 1985), Puget Sound (Pedersen, 1985) and inlets of the west coast of Vancouver Island (McFarlane and Beamish, 1985), the offshore population is of greatest economic importance, as it contributes large biomass for fisheries in both Canadian and United States (US) waters (Francis et al., 1989; Smith et al., 1990; Helser et al., 2004). Pacific hake spawn off the coast of California in the winter and migrate north to feed off the coast of Washington and British Columbia in the summer (Fig. 1). A great deal of controversy revolves around this fishery. The largest and most valuable fish migrate farther north (McFarlane and Beamish, 1985) and dramatic variability in the interannual distribution of biomass (Fig. 2) and therefore yield of hake between Canada and the US exists.

Several authors have suggested that Pacific hake distribution may be related to poleward flow, with stronger flow aiding the migration of hake, and weaker flow impeding it (Smith et al., 1990; Dorn, 1995; Benson et al., 2002; Agostini et al., 2006). Agostini et al. (2006)

discuss the variability of sub-surface poleward flow (the undercurrent) in the CC as well as the links with climate and hake abundance and distribution. By using acoustic data they examine sub-surface flow characteristics in 1995 and 1998 and their relationship with hake abundance and distribution. Their results show that the distribution and intensity of the undercurrent plays a key role in defining adult hake habitat along the west coast of North America; for example, habitat expands possibly as a result of changes in location and intensity of the poleward undercurrent in El Niño years. Here we test this hypothesis, and use a geostatistical approach to quantify adult Pacific hake habitat during different climate regimes. We model changes in amount and distribution of hake adult habitat as defined by the poleward undercurrent and examine the possibility of using the undercurrent as a predictive index of adult hake habitat abundance. Finally we discuss the potential role of hake habitat as a CC ecosystem indicator.

## 2. Methods

### 2.1. Study area

Data on hake abundance and distribution along the west coast of North America have been collected starting in 1977 by the National Marine Fishery Services (NMFS) and starting in 1992 in collaboration with the Department of Fisheries and Oceans Canada (DFO). Summer echo-integration trawl surveys have been conducted on a triennial basis along the continental shelf from California to the northern limit of hake aggregations in British Columbia. Details of the triennial surveys are available in [Wilson et al. \(2000\)](#). Surveys run from July 1 to September 1 off the west coast of North America. Transects are on average 52 km long and 18 km apart, generally running mid-shelf to mid-slope between the 50 and 1500 m isobaths. Our analysis focused on a sub-section of the survey area

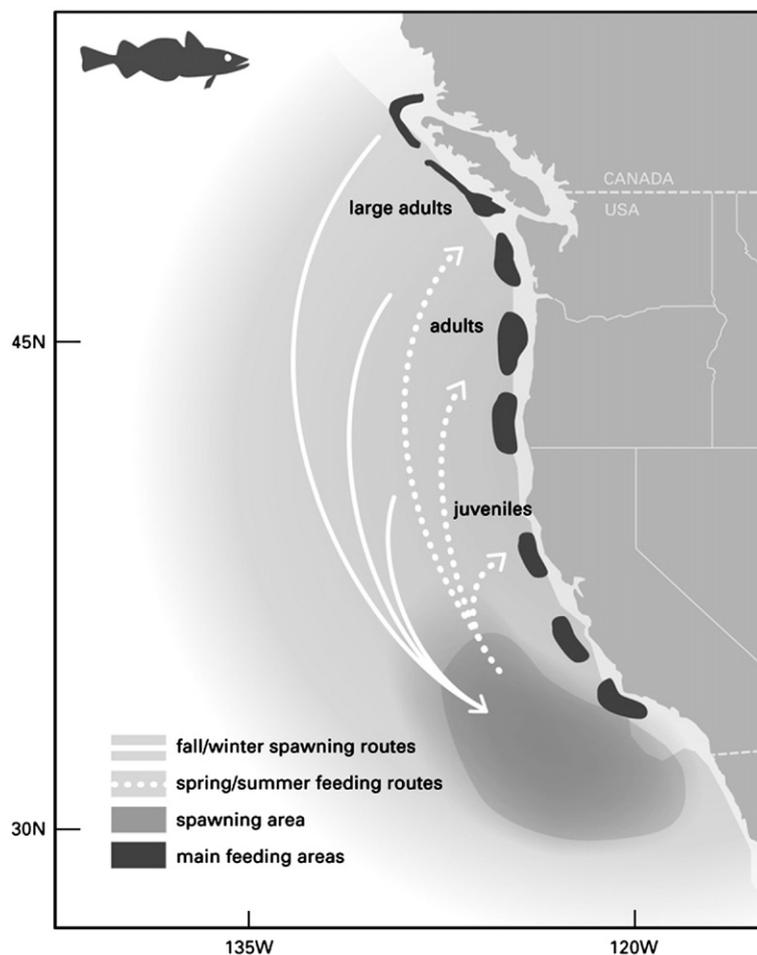


Fig. 1. Schematic representations of Pacific hake migrations in the California Current system.

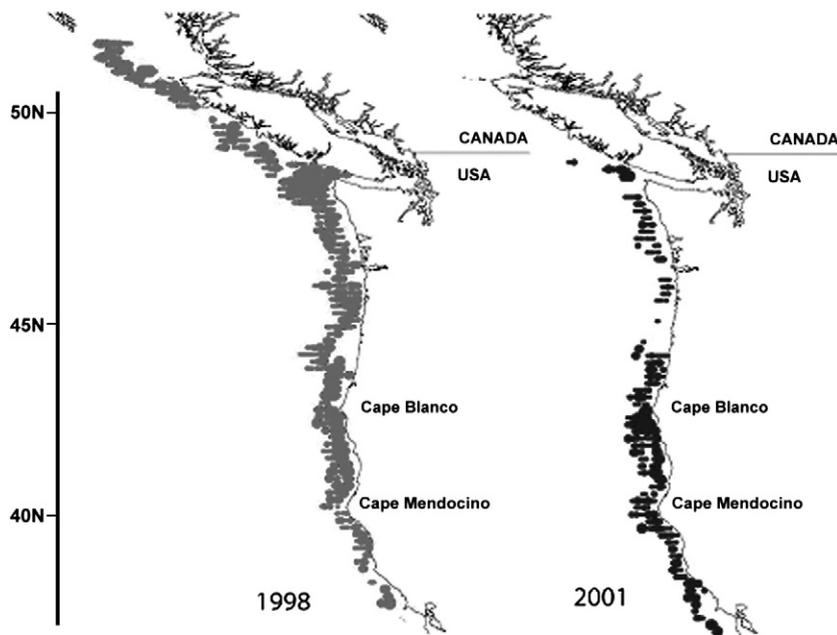


Fig. 2. Acoustic backscatter signal representative of hake abundance during 1998 (warm year) and 2001 (cold year). Data from the National Marine Fisheries Service-Alaska Fisheries Science Center (NMFS-AFSC).

(38°–43°N). In order to capture conditions during different climate regimes, we analyzed survey data from 1995 (a neutral year) and 1998 (an El Niño year).

## 2.2. Biological data

Abundance and distribution of adult hake were derived from acoustic data collected using a Simrad EK500 quantitative echosounding system (Simrad Inc., Lynwood Wash.) split-beam transducers (38 and 120 kHz; Simrad Inc.) mounted on the bottom of the vessel hull, 9 m below the surface (Wilson et al., 2000). Data collected with the 38 kHz transducer were used in this study. Standard target strength–length relationships were used to convert acoustic backscatter to fish density. The target strength relationship used was  $TS = 20 \log L - 68$ , where  $L$  represents fish length measured in centimeters. For detailed methods on data processing see Wilson and Guttormsen (1997) and Wilson et al. (2000). Fish values (numbers of fish) at each location  $i, j$ , represent an average of measurements taken at that location over 48 10 m depth bins (0–480 m).

## 2.3. Physical data

Distribution and intensity of alongshore (north/south) flow were derived from an RD instrument 153.6 kHz narrow-band, hull mounted shipboard acoustic Doppler

current profiler (ADCP). A vertical bin width of 8 m was used, pulse length of 8 m, and an ensemble averaging time of 2.5 min. The depth range of good data (good pings > 30%) was typically 22–362 m. The ADCP was slaved to the Simrad EK500 to avoid interference. Pre-survey tests confirmed no interference between the two instruments when the ADCP was in water-pulse mode. When the ADCP bottom-track feature was enabled, however, an artificial signal was detected on the EK500. For this reason, bottom tracking was never enabled during the survey. GPS P-code navigation was used for position and gyrocompass for heading, to determine absolute velocities. Tidal currents remain in the processed ADCP velocities. These are expected to be small (<0.05 m/s) offshore of the shelf break (Erofeeva et al., 2003). For detailed ADCP processing methods, see Pierce et al. (2000). The flow value at each location  $i, j$  represents flow values averaged over the 120–330 m depth bins, depths reported as typical of the California undercurrent (Agostini et al., 2006; Pierce et al., 2000).

## 2.4. Modeling approach

### 2.4.1. Structural analysis

The goal of the structural analysis is to evaluate which physical variables are associated with hake habitat. The focus therefore is on the regression between hake abundance and physical variables that define

habitat. There is spatial correlation among the observations, however, that must be incorporated into the structural analysis. We describe how particular habitat variables are related to hake abundance at a particular location (station) by means of models that assume spatial dependence (geostatistical analysis). We incorporated models that assume spatial independence as subsets of models that assumed spatial dependence.

The regression models are based on a hierarchical generalized linear modeling framework in which the errors from the regression are assumed to be spatially correlated (Diggle et al., 1998). The data are modeled as a function of predictor variables, associated coefficients, and error. The error is further modeled by a correlation function and a set of coefficients that define the degree of spatial correlation. We present the modeling framework first in its entirety and then describe the two levels of the hierarchy.

Assume the hake abundance  $Y_i$  (number of hake at location,  $\mathbf{x}_i=1, \dots, N$ ), is a function of habitat quality at that location  $\mathbf{x}_i$ , and by the hake in surrounding locations  $\mathbf{x}_j=1, \dots, i-1, \dots, N, j \neq i$ .

$$Y_i = \mu_i + S(\mathbf{x}_i) + \varepsilon_i \quad (1)$$

where  $\mu_i$  is a mean effect (see habitat discussion below),  $\mathbf{x}_i$  the observation location,  $S(\mathbf{x}_i)$  is a stationary Gaussian process with expected value  $E[S(x)]=0$  and  $\text{cov}[S(\mathbf{x}_i), S(\mathbf{x}_j)] = \sigma^2 \rho(\mathbf{x}_i - \mathbf{x}_j)$  ( $\sigma^2$ =variance;  $\rho$ =correlation coefficient) and  $\varepsilon$  are mutually independent Gaussian random variables with mean=0, and variance= $\tau^2$  (Diggle et al., 1998).

#### 2.4.2. Habitat

In biological terms,  $\mu_i$  is the effect of habitat, and can be modeled as a function of covariates (e.g.,  $\mu_i = \beta_0 + \beta_1 * C_i$ , where  $C_i$  is a covariate measured at location  $\mathbf{x}_i$ ). Current (average alongshore current velocity over 120 m–330 m sub-surface layer) and depth (bottom depth) were hypothesized to affect habitat using several competing mechanisms. Habitat quality was hypothesized to be affected by:

- depth only
- current only
- depth and current
- depth and current, an interaction term
- depth and a unimodal effect of current (intermediate levels of current produce higher habitat quality than minimum or maximum levels).

The working hypotheses were evaluated to determine which hypothesis was the most likely given the data.

Akaike's Information Criterion (AIC) was used to rank each of the models. The AIC statistic, estimated with the addition of each new parameter to the model, accounts for degrees of freedom used and the goodness of fit such that more parsimonious models have a lower AIC (Chambers and Hasties, 1992).

#### 2.4.3. The spatial process

The spatial process,  $S(\mathbf{x}_i)$ , is due to a self organization process such as the aggregation of schooling fish. The presence of a strong  $S(\mathbf{x}_i)$  process will result in fish abundance appearing "clumped" spatially. Spatial autocorrelation affects the regression by effectively reducing the degrees of freedom for the regression (Cressie, 1993). In spatially correlated observations, the distance between samples is informative about the level of abundance at another location (i.e., measures of abundance at short distances are more like the current observation than measures of abundance that are far away). The level of autocorrelation therefore provides useful information on the nature of the spatial structure and will depend on the distance or lag between two samples.

#### 2.4.4. Variogram estimation

The analysis of spatial structure involved two steps. The first step was to use an empirical variogram to describe the spatial structure of the abundance measurements. This allowed us to quantify the spatial dependency and partition it along the various distance classes. The empirical variogram represents the semi-variance between data points as a function of the spatial distance (lag) between them. The experimental variogram was calculated using:

$$\gamma^*(h) = \frac{1}{2n(h)} \sum_i^{n(h)} [f(\mathbf{x}_i) - f(\mathbf{x}_i + h)]^2 \quad (2)$$

Where  $\gamma^*(h)$  represents the empirical variogram for distance  $h$ ,  $n(h)$  is the number of points separated by lag  $h$ , and  $f(\mathbf{x}_i)$  is the value at data point  $\mathbf{x}_i$  (Petitgas, 1998). All the grid samples in the sub-section of the sampling area were included in the variogram calculation. Variogram behavior was assumed to be the same in all directions, thus the results presented correspond to omnidirectional variograms.

#### 2.4.5. Variogram model fitting

The spatial autocorrelation,  $S(\mathbf{x})$ , can be modeled with a covariance function, or with a semi-variance function. We choose to model it with semi-variance as this is the common approach in geostatistics (Cressie, 1993); therefore, we used an exponential variogram to

Table 1  
Model formulations

| Models without autocorrelation term   | AIC         |
|---|-------------|
| ${}^{\text{nac}}\beta_0 + {}^{\text{nac}}\beta_1 \text{depth}$  | 7092        |
| ${}^{\text{nac}}\beta_0 + {}^{\text{nac}}\beta_1 \text{current}$  | 7153        |
| ${}^{\text{nac}}\beta_0 + {}^{\text{nac}}\beta_1 \text{current} + {}^{\text{nac}}\beta_2 \text{depth}$  | 7056        |
| ${}^{\text{nac}}\beta_0 + {}^{\text{nac}}\beta_1 \text{current} + {}^{\text{nac}}\beta_2 \text{depth} + {}^{\text{nac}}\beta_3 \text{current} * \text{depth}$ | 7186        |
| ${}^{\text{nac}}\beta_0 + {}^{\text{nac}}\beta_1 \text{current}^2 + {}^{\text{nac}}\beta_2 \text{current} + {}^{\text{nac}}\beta_3 \text{depth}$              | <b>7057</b> |
| Models without autocorrelation term   | AIC         |
| $\beta_0 + \beta_1 \text{depth} + S(\mathbf{x})$  | 6016        |
| $\beta_0 + \beta_1 \text{current} + S(\mathbf{x})$  | 6018        |
| $\beta_0 + \beta_1 \text{current} + \beta_2 \text{depth} + S(\mathbf{x})$   | 6017        |
| $\beta_0 + \beta_1 \text{current} + \beta_2 \text{depth} + \beta_3 \text{current} * \text{depth} + S(\mathbf{x})$   | 6019        |
| $\beta_0 + \beta_1 \text{current}^2 + \beta_2 \text{current} + \beta_3 \text{depth} + S(\mathbf{x})$  | <b>6014</b> |

Value in bold indicates model chosen based on AIC value. Please note that differences in  $\text{AIC} \geq 7$  indicates strong support for the lower valued model. The model covariates are: depth=bottom depth and current=sub-surface alongshore flow velocity (120–330 m). The inclusion of a spatial process in the error component is represented by  $S(\mathbf{x})$ .

model the spatial autocorrelation in observations. In the initial stages of model development, a spherical variogram was also fitted to the data; however, the exponential variogram shape appeared to be more closely fit the residual spatial variation. The exponential variogram can be described by the equation

$$\gamma(\mathbf{x} - \mathbf{x}') = \tau^2 + \sigma^2 \left\{ 1 - \exp \left[ -\frac{(\mathbf{x} - \mathbf{x}')}{\phi} \right] \right\} \quad (3)$$

where  $\tau^2$  is the variability at scales smaller than the distance between samples and variability due to measurement error (the nugget effect),  $\phi$  is the distance over which samples are spatially autocorrelated (range), and  $\sigma^2$  is the background variability that occurs at distances greater than the range (sill).

We then estimated linear regression and variogram model parameters simultaneously using the geoR package (Ribiero and Diggle, 2001) in the R software (<http://www.r-project.org/>).

#### 2.4.6. Model selection and validation

The AIC statistic was used to rank models with spatial autocorrelation as well as models lacking spatial autocorrelation. Furthermore, to evaluate the fit between model and observations for the model with the lowest AIC, we applied a cross-validation method to a sub-set of the data. This method has been used for optimal choices of variogram models (Petitgas and Poulard, 1989; Simard et al., 1992). The procedure consists of deleting one datum and using the remaining data to predict the deleted value using the chosen model.

### 3. Results

Geostatistical regression models, which included explicit spatial autocorrelation, were superior to standard statistical regression for explaining trends in hake abundance. In general, models with autocorrelation fit the data substantially better than models without autocorrelation as indicated by AIC values (Table 1). Differences in AIC values of 7 or more are considered to be strong support for a lower AIC model (Burnham and Anderson, 2002), and the average decrease in AIC in the models with autocorrelation was approximately 900 U. This disparity in model fit is evident in the observed to predicted plots under the best fit model with autocorrelation and the same model without the spatial autocorrelation component (Fig. 3).

Within the regression models with autocorrelation, the best fitting model (AIC=6014) was:

$$y = \beta_0 + \beta_1 \text{current}^2 + \beta_2 \text{current} + \beta_3 \text{depth} + S(\mathbf{x}), \quad (4)$$

where  $y$ =number of fish, depth=bottom depth, and current=sub-surface alongshore flow velocity (120–330 m). This model was 2 AIC units lower than any of the other models, which suggested moderate support over the other models (Burnham and Anderson, 2002).

The model described in Eq. (4) has a spatial process error component  $S(\mathbf{x})$  that was modeled by an exponential variogram consisting of a nugget, sill, and

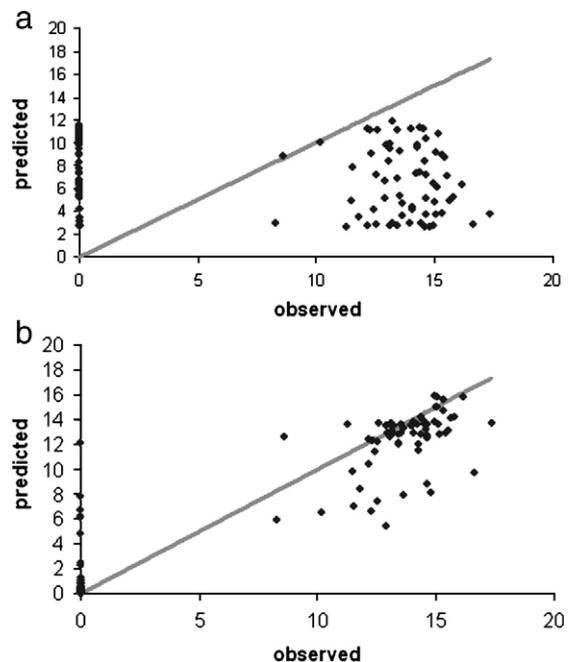


Fig. 3. Fits of model (a) without spatial autocorrelations predictions and (b) with spatial autocorrelation. Units: log number of fish.

Table 2

Model parameter estimates for the model structure with the lowest AIC value (autocorrelated model) and model parameter estimates for the same model without spatial autocorrelation

| Parameter   | Mean value | Standard deviation |
|---|------------|--------------------|
| <i>Model without autocorrelation term: <math>^{nac}\beta_0 + ^{nac}\beta_1\text{current}^2 + ^{nac}\beta_2\text{current} + ^{nac}\beta_3\text{depth}</math></i> |            |                    |
| $\beta_0^{nac}$   | 12.68      | 0.49               |
| $\beta_1^{nac}$   | -5.73      | 13.5               |
| $\beta_2^{nac}$   | 2.90       | 2.92               |
| $\beta_3^{nac}$   | -0.006     | 0.0004             |
| <i>Model with autocorrelation term <math>\beta_0 + \beta_1\text{current}^2 + \beta_2\text{current} + \beta_3\text{depth} + S(\mathbf{x})</math></i>             |            |                    |
| $\beta_0$   | 8.20       | 0.992              |
| $\beta_1$   | 20.28      | 12.04              |
| $\beta_2$   | -3.75      | 3.009              |
| $\beta_3$   | -0.0025    | 0.00087            |
| $\tau^2$ (nugget)   | 2.81       |                    |
| $\phi$ (range)  | 3.67 km    |                    |
| $\sigma^2$ (sill)   | 39.28      |                    |

range (Table 2, Fig. 4). The nugget (2.81) represents variability at distances less than the smallest distance between sample stations or variability due to measurement error. The range (2.67 km) is the average distance beyond which points are no longer spatially correlated. The sill (39.28) quantifies the maximum level of variability among points or the variability that occurs at large distances. The fitted variogram suggests that there is a substantial spatial component to the hake abundance after accounting for the effects of current and depth. In biological terms, hake abundance is a function of the physical attributes present in the environment (habitat) as well as a function of where other hake are located (schooling or aggregating).

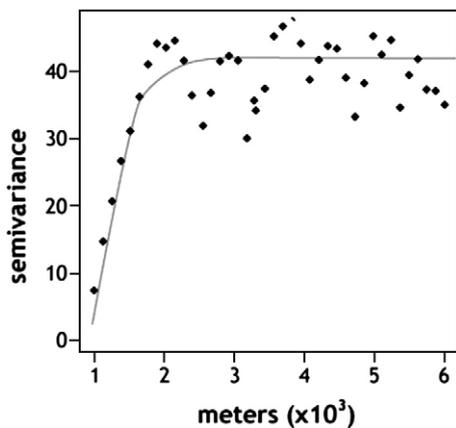


Fig. 4. Experimental variogram (semi-variance between data points as a function of distance) fitted to an exponential function (solid line).

### 3.1. Model inference

Inference on hake habitat relationships was based on inference from the model in Eq. (4). Namely, sub-surface flow and bottom depth were important determinants of hake habitat. We refer to the model represented in Eq. (4) as the hake habitat model with autocorrelation (HHMAC), and the same model structural form without autocorrelation (i.e.,  $y = \beta_0 + \beta_1\text{current}^2 + \beta_2\text{current} + \beta_3\text{depth}$ ) as (HMM) from this point forward. The HMM model was the second best model among models without autocorrelation (Table 1) and received almost equivalent support as the best fitting model (difference of 1 AIC unit). We comment on the results that would have been obtained if the HMM model had been used rather than the HHMAC model to illustrate the importance of accounting for spatial autocorrelation.

The relationship between habitat and hake abundance are different under the HMM and HHMAC models. Both models indicate curvi-linear relationships between current and hake abundance. The model without autocorrelation (HMM) indicates that the sub-surface flow hake relationship is dome-shaped with current velocities

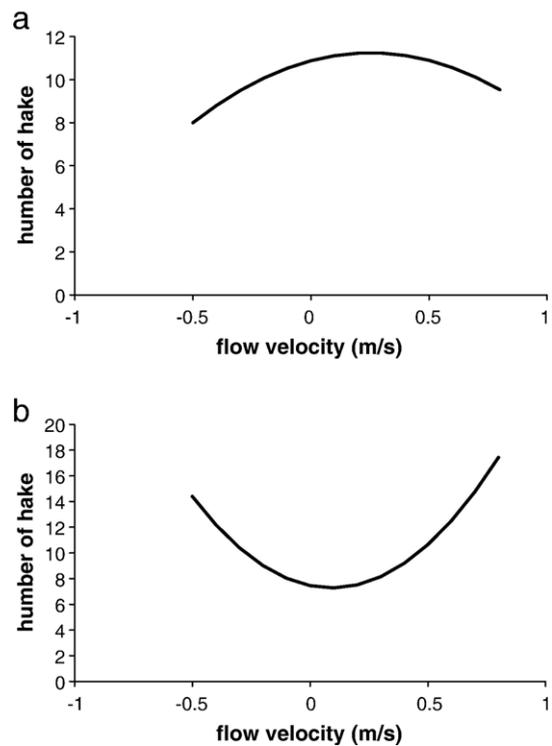


Fig. 5. Relationship between hake abundance and current velocity as predicted by the model without autocorrelation (upper panel) and the model with autocorrelation (lower panel).

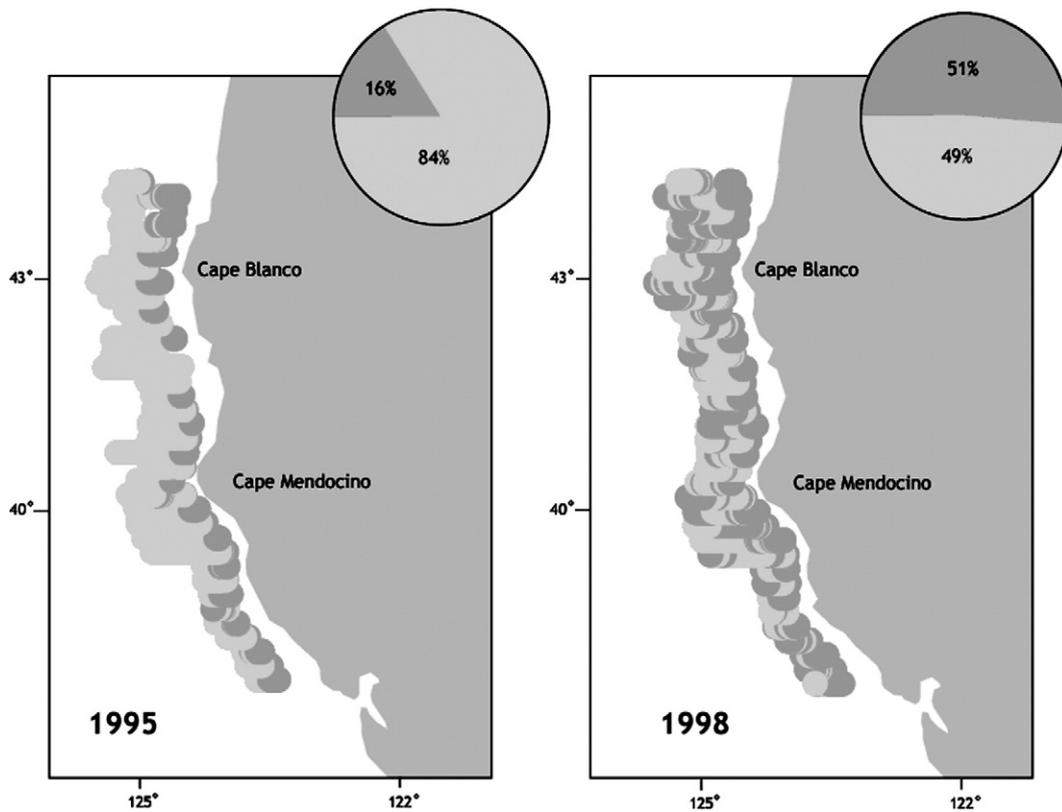


Fig. 6. Predicted hake habitat given bottom depth and undercurrent velocities. Light grey represents less suitable habitat (number of fish < 1100), dark grey represents more suitable habitat (number of fish > 1100). Pie charts represent overall habitat distribution for 1995 and 1998.

near zero leading to higher amount of favorable hake habitat (Fig. 5a). The best formulation for the model with autocorrelation (HMMAC) shows the opposite trend, with high and low flow velocities leading to higher amount of favorable hake habitat (Fig. 5b). The direction of the velocity did not appear to be important in the HMMAC model, however (Fig. 5b). In addition, the effect of depth is diminished (smaller absolute value of coefficient) in the HMMAC model (Table 2). To summarize, high quality hake habitat would have been expected at low current velocities under the HMM model, whereas high quality hake habitat would be expected at greater velocities in the HMMAC model.

Predicted hake habitat in 1995 was calculated by implementing the mean effect component of Eq. (4) using model coefficients from the HMMAC model and covariate values (depth and current velocity) from sampling stations in 1995 (Fig. 6). In addition, predictions of hake habitat under the physical conditions in 1998 were also made (Fig. 6). The fit of the autocorrelation model to 1998 data is inferior to the fit to the 1995 data (Fig. 7). More variability was explained by fitting a mean term to the observed 1998 data (residual

sums of squares of mean estimate=3549) than was explained by the 1995 model (residual sums of squares estimate=5630). The pattern in Fig. 7 is similar to Fig. 3a, suggesting that the spatial covariance structure was different between 1998 and 1995. In particular, the spatial aggregations in 1998 may have been tighter, thus the sill was probably higher and range shorter in 1998 than 1995.

We defined hake habitat as favorable at location where fish densities were higher than 1100 individuals and less favorable at locations where fish densities were less than 1100 individuals, based on a break in the histogram. Amount of favorable hake habitat was lower in 1995 (16% of area considered in modeled) compared to 1998 (51% of area considered in model) (Fig. 6).

#### 4. Discussion

We use a Geostatistical approach to uncover potential relationships between hake and its environment. This approach allows us to account for dependence between data points, a process we considered important in our analysis. Biologists have traditionally relied on methods

developed for independent data, even though independence is an unrealistic assumption (Legendre, 1993). The simplifying assumption that datum from one point in space is not influenced by another datum in the study at another point in space rarely holds true (Carroll and Pearson, 2000). The presence of an organism at a specific site is induced by a number of major forces such as ocean currents, winds and climate. It is also induced by the presence of another organism at a neighboring site. Conducting tests based on independence could lead to false identification of existing relationships.

Our results indicate that hake distribution is related to poleward flow (a dynamic variable) and bottom depth (a static variable). This relationship was identified by incorporating a polynomial model (the mean effect component) as well as a spatial autocorrelation component in our model. Including both of these components allowed us to identify and quantify important relationships. It is interesting to note that our results show that the shape of the relationship between hake distribution and poleward flow would have been different if an autocorrelation component had not been included. Identifying dynamic relationships such as the one between hake distribution and poleward flow is challenging and tools effective at capturing dynamic relationships as well as multiple processes are necessary.

Spatial approaches such as the one we used here allow us to effectively explore mechanisms that may explain the distribution of fish populations. Such an understanding could not have been possible by simply using general linear models. An understanding of the process driving these distributions is essential in our evaluation of the sustainability of fish stocks. Thus far fishery science has tended to evaluate management performance with indicators based on non-spatial population dynamics models (Pauly et al., 2003),

ignoring inherent spatial variability of a stock's distribution (Booth, 2000). For stocks such as Pacific hake where the biology of the fish has a spatial component, efforts should be made to incorporate spatial structure in indicators.

We have become increasingly aware of the distribution patterns of many species, yet what drives these distributions is often poorly understood, making prediction a difficult task (Verity et al., 2002). We too had difficulty applying our model when attempting to predict habitat quantity in 1998; however, there were several important differences between 1995 and 1998 that may have affected predictions of abundance in 1998. We applied the coefficients from the model fit to 1995 data to predict habitat quality in 1998; however, our 1998 model fit (Fig. 7) clearly points out the need for further refinement. Using a sub-set of the data probably introduced some measurement and process error. The range of habitat considered in the model we built is located at the southern most edge of the adult hake distribution (38–43°N). This area is occupied by a mixture of juveniles and adults. It is difficult to discriminate acoustic signal of smaller sizes (juveniles) from signal for other organisms (e.g. euphasiids). Measurement error is introduced here, as smaller sizes might not be fully reported. This could be one of the explanations for the inferior fit of the model to the 1998 data compared to the 1995 data, as in 1998 the population reportedly shifted north (Wilson et al., 2000; Helser et al., 2002) and fewer adults were observed in southern areas. As can be seen in Fig. 7 the model does not do well with predicting '0 values' (process error). Because of the shift north in the population described above, the 1998 data set had a higher number of locations with 0 fish. As a result, the 1998 model predictions were not as accurate as the 1995 model predictions. A model based on data from a section of the hake habitat located farther north (e.g. 43°–48°N), where the majority of the population sampled is older (fish size is bigger thus more accurately sampled) and the area consistently occupied in both 1995 and 1998, could have been more informative. The inaccuracy of 1998 model predictions also suggests that the spatial autocorrelation function for the 1998 data may be different than the autocorrelation function fit to the 1995 data. These results also suggest that a unique model should be fitted for El Niño years, as the HMMAC model was not robust to the two year types.

The three dimensional nature of the data set presented another challenge. Our model represents hake habitat in two dimensions (latitude and longitude); thus the three dimensional data had to be collapsed to

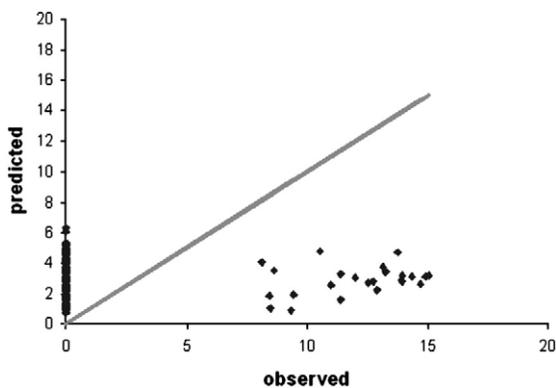


Fig. 7. Fits of model with spatial autocorrelations predictions to 1998 data. Units: log number of fish.

two dimensions. There are a number of ways the under-current could have been represented in the model; we chose an average of current values in the 130–220 m depth bins and measured the flow relative to the poleward direction (i.e., current could have a negative value). This choice might have affected our results by focusing on an aspect of the current that may not have been the most relevant to fish distribution. The resolution of the acoustic data set used in this study allowed us to examine hake habitat over a wide geographic range at fine detail. Although powerful, this also presented some challenges. The data set we used for this analysis was very large (information on flow characteristics and hake spanning 15° of latitude, ~300,000 grid points/year). Because of the large nature of the data set (thus the computational power necessary to build a model based on data from the entire hake distribution range), we used a sub-set of the data in our analysis as outlined above in the methods section. This allowed us to test approaches and methods that could in the future be applied to the entire data set and additional years.

The main aim of this study was to develop a quantitative measure of hake habitat. Our goal was to develop a metric that would quantify important physical–biological linkages in the California Current Ecosystem. Given the relationship between hake distribution and the physical structure of the CC described by a number of authors (Smith et al., 1990; Dorn, 1995; Benson et al., 2002; Agostini et al., 2006) and the importance of hake in the CC food web (Field, 2004) hake habitat was explored as a suitable metric. Our results indicate that climate forcing affects the spatial structure of hake habitat. Hake habitat in our study area expands during a warm year El Niño year (1998) and contracts during a cold year (1995) (Fig. 6). Agostini et al. (2006) suggest that this may be due to changes in the intensity and location of the poleward current. Our study supports this hypothesis, and point to the importance of accounting for physical processes in the study of fish distribution.

In this study we apply a spatial approach to outline hake habitat areas, and examine how climate forcing may affect these areas. This addresses both spatial and environmental processes of the CC system, facilitating the inclusion of broader ecosystem considerations and objectives in single species assessments and management plans. While detailed single species assessments still form the core of management advice in most cases, they are increasingly embedded in an ecosystem context, at least qualitatively (Mace, 2001). However, a number of important classes of ecosystem interactions are currently not being routinely evaluated (Sissenwinse

and Murawski, 2004). Amongst these are relationships between biological and physical components of ecosystems. The stability of biological communities is affected by the interaction between life history, environmental variation and fishing strategies (Sissenwinse and Murawski, 2004). Accounting for interactions between biological and physical components of the ecosystem such as the one we examine here will not only help us evaluate important ecosystem interactions, but it will also help determine appropriate spatial scales of data collection, science and management presently missing from conventional single species management (Hilborn, 2004).

The desire to represent key ecosystem interactions has led to the recent focus on ecosystem indices. A number of symposia and working groups have been convened on this topic (e.g.: “Ecosystem Considerations in Fishery Management”, Anchorage 1998; “Responsible Fishing in the Marine Environment”, Reykjavik 2001; IOC-SCOR working group 119, “Quantitative Indicators for Fisheries Management”, Paris 2004; “Advancing Scientific Advice for an Ecosystem Approach to Fisheries”, Dublin 2004). Most of the indices developed to date represent trophic interactions, while work on indices representing interactions between species and the physical environment is lagging behind. Climate forcing of ecosystems has mostly been described by large scale indices such as PDO and ENSO. Climate impacts are aggregated across large spatial scales and range of species and direct links are made between climate forcing and production variability. However, changes in ecosystem structure are often the result of changes in physical habitat with very distinct spatial structure. In order to fully describe ecosystems, spatially explicit indices directly representing physical–biological linkages should also be developed.

The focus of this study on pelagic habitat of a key trophic species addresses this issue. A quantitative measure of hake habitat such as the one we develop here could potentially serve as an ecosystem indicator. Hake is one of the major predators in the northern CC system. The amount and distribution of adult hake habitat has large implications for the Pacific Northwest food web. For example Field (2004) found that during warm years when hake are more abundant in northern CC waters (north of Cape Mendocino), there is an increase in predation (particularly on pandalid shrimp and small flatfish) and competition (for euphausiids, forage fish and other prey of resident groundfish). The absence or presence of hake in Pacific Northwest waters is likely related to habitat suitability along its range of

distribution. Changes in the amount of adult hake habitat could serve as an indicator of changes in the structure/energy flow of the northern CC ecosystem, as changes in hake distribution could imply changes in the productivity of other commercially and ecologically important species. Metrics able to detect these types of changes could be good ecosystem indicators, as they are related to both the physical and biological structure of the ecosystem. Examples of potential ecosystem indicators are: % of hake habitat distributed north of Latitude  $X$ , % of hake habitat distributed offshore of longitude  $X$ , % of overall habitat defined as suitable hake habitat, favorable/unfavorable habitat, (favorable habitat) $_{t+1}$  – (favorable habitat) $_t$ .

Spatial structure and environmental processes are discussed as essential to the development Ecosystem Based Fishery Management (EBFM). However, as we develop approaches to implement EBFM, habitat issues are not receiving the attention they warrant. Most of the work on habitat focuses on benthos, with knowledge of pelagic habitat lagging behind. Studies such as this one will help to increase the awareness of pelagic habitat and contribute to the development of effective EBFM strategies.

## Acknowledgments

We thank Mike Guttormsen for help with acoustic data processing; Elizabeth Babcock, Jodie Little, and Warren Wooster for helpful discussion and review of this manuscript; John Mickett for help with programming and Amity Femia for help with graphics. This study was supported by the JISAO/Climate Impacts Group (University of Washington) under NOAA Cooperative Agreement NAI17RJ123, contribution # XXX. S.D. Pierce was partially supported by National Oceanic and Atmospheric Administration (Grant AB133F05SE4205).

## References

- Agostini, V.N., Francis, R.C., Hollowed, A.B., Pierce, S., Wilson, C., Hendrix, A.N., 2006. The relationship between hake (*Merluccius productus*) distribution and poleward sub-surface flow in the California current system. *Canadian Journal of Fishery and Aquatic Sciences* 63, 2648–2659.
- Barange, M., Hampton, I., 1997. Spatial structure of co-occurring anchovy and sardine populations from acoustic data: implications for survey design. *Fisheries Oceanography* 6, 94–108.
- Barange, M., Coetzee, J.C., Twatwa, N.M., 2005. Strategies of space occupation by anchovy and sardine in the southern Benguela: the role of stock size and intra-species competition. *ICESAS Journal of Marine Science* 62, 645–654.
- Beamish, R.J., McFarlane, G.A., King, J.R., 2000. Fisheries climatology: understanding decadal scale processes that naturally regulate British Columbia fish populations. In: Harrison, P.J., Parsons, T.R. (Eds.), *Fisheries Oceanography: An Integrative Approach to Fisheries Ecology and Management*. Blackwell Science, Oxford, pp. 94–145.
- Benson, A.J., McFarlane, G.A., Allen, S.E., Dower, J.F., 2002. Changes in Pacific hake (*Merluccius productus*) migration patterns and juvenile growth related to the 1989 regime shift. *Canadian Journal of Fishery and Aquatic Sciences* 59, 1969–1979.
- Booth, A.J., 2000. Incorporating the spatial component of fisheries data into stock assessment models. *ICES Journal of Marine Science* 57, 858–865.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multi-model Inference: A Practical Information—Theoretic Approach*, 2nd Edition. Springer-Verlag, New York, New York, USA.
- Carroll, S., Pearson, D., 2000. Detecting and modeling spatial and temporal dependence in conservation biology. *Conservation Biology* 14 (6), 1893–1897.
- Chambers, J.M., Hasties, T.J., 1992. *Statistical Models in S*. Wadsworth and Brooks Cole Advanced Books, Pacific Grove, California.
- Conan, G.Y., Maynou, F., Stloyarenko, D., Mayer, L., 1994. Mapping and assessment of fisheries resources with coastal and depth constraints, the case study of snow crab in the Bay of Islands Fjord (Newfoundland). *ICES C.M. 1994/B+D+G+H: 4*. International Council for the Exploration of the Sea, Charlottenlund, Denmark.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*. John Wiley and Sons, Inc., New York, New York, USA.
- Diggle, P.J., Tawn, J.A., Moyeed, R.A., 1998. Model-based geostatistics. *Applied Statistics* 47 (3), 299–350.
- Dorn, M.W., 1995. The effects of age composition and oceanographic conditions on the annual migration of Pacific Whiting, *Merluccius productus*. *CalCOFI Reports* 36, 97–105.
- Erofeeva, S.Y., Egbert, G.D., Kosro, P.M., 2003. Tidal currents on the central Oregon shelf: models, data, and assimilation. *Journal of Geophysical Research* 108, 3148.
- Field, J.C., 2004. Application of ecosystem-based fishery management approaches in the Northern California Current. PhD Dissertation. University of Washington.
- Fletcher, W.J., Summer, N.R., 1999. Spatial distribution of sardine (*Sardinops sagax*) eggs and larvae: an application of geostatistics and resampling survey data. *Canadian Journal of Fisheries and Aquatic Sciences* 56, 907–914.
- Francis, R.C., Adlerstein, S.A., Hollowed, A.B., 1989. Importance of environmental fluctuations in the management of Pacific hake. In: Beamish, R.J., McFarlane, G.A. (Eds.), *Effects of Ocean Variability on Recruitment and an Evaluation of Parameters Used in Stock Assessment Models*. Special Publication of the Canadian Journal of Fisheries and Aquatic Sciences, 108.
- Hare, S.R., Mantua, N.J., 2000. Empirical evidence for North Pacific regime shifts in 1977 and in 1989. *Progress in Oceanography* 47 (103), 145.
- Helser, T.E., Methot, R.E., and Fleischer, G.W., 2002. Stock assessment of Pacific hake in U.S. and Canadian Waters in 2003. Northwest Fisheries Science Center, National Marine Fisheries Service.
- Helser, T.E., Methot, R.E. and Fleischer, G.W., 2004. Stock assessment of Pacific hake in U.S. and Canadian Waters in 2003. Northwest Fisheries Science Center, National Marine Fisheries Service.
- Hilborn, R., 2004. Ecosystem-based fisheries management: the carrot or the stick? *Marine Ecology, Progress Series* 274, 269–303.
- Horne, J.K., Schneider, D.C., 1995. Spatial variance in ecology. *Oikos* 74, 1–9.
- Kracker, L., 1999. The geography of fish: the use of remote sensing and spatial analysis tools in fisheries research. *Professional Geographer* 51 (3), 440–450.

- Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659–1673.
- Livingston, P.A., Bailey, K.M., 1985. Trophic role of the Pacific whiting, *Merluccius productus*. *Marine Fisheries Review* 47 (2), 16–22.
- MacCall, A.D., 1996. Patterns of low-frequency variability in fish populations of the California current. *CalCOFI Report* 37, 100–110.
- Mace, P.M., 2001. A new role for MSY in single-species and ecosystem approaches to fisheries stock assessment and management. *Fish and Fisheries* 2, 2–32.
- Maravelias, C.D., Reid, D.G., Simmonds, E.J., Haralabous, J., 1996. Spatial analysis and mapping of acoustic survey data in the presence of high local variability: geostatistical application to North Sea herring (*Clupea harengus*). *Canadian Journal of Fisheries and Aquatic Sciences* 53, 1497–1505.
- McFarlane, G.A., Beamish, R.J., 1985. Biology and fishery of Pacific Whiting, *Merluccius productus*, in the Strait of Georgia. *Marine Ecology, Progress Series* 47, 23–34.
- McNaughton, S.J., Wolf, L.L., 1973. *General Ecology*. Holt, Rinehart and Winston, New York, New York, USA.
- Pauly, D., Watson, R., Christensen, V., 2003. Ecological geography as a framework for transition toward responsible fishing. In: Sinclair, M., Valdimarsson, G. (Eds.), *Responsible Fisheries in the Marine Ecosystem*. FAO, Rome, pp. 87–101.
- Pedersen, M., 1985. Puget Sound Pacific Whiting, *Merluccius productus*, resource and industry: an overview. *Marine Fisheries Review* 47, 35–38.
- Petitgas, P., 1993. Geostatistics for stock assessment: a review and an acoustic application. *ICES Journal of Marine Sciences* 50, 285–298.
- Petitgas, P., 1998. Biomass-dependent dynamics of fish spatial distributions characterized by geostatistical aggregation curves. *ICES Journal of Marine Sciences* 55, 443–453.
- Petitgas, P., 2001. Geostatistics in fishery survey design and stock assessment: models, variances and applications. *Fish and Fisheries* 2, 231–249.
- Petitgas, P., Poulard, J.C., 1989. Applying stationary geostatistics to fisheries: a study on hake in the Bay of Biscay. *ICES C.M./G:62*. International Council for the Exploration of the Sea, Charlottenlund, Denmark.
- Pierce, S.D., Smith, R.L., Kosro, P.M., Barth, J.A., Wilson, C.D., 2000. Continuity of the poleward undercurrent along the eastern boundary of the mid-latitude north Pacific. *Deep Sea Research. (II)* 47, 811–829.
- Ribiero, Diggle, P.J., 2001. A package for geostatistical analysis. *R-NEWS* 1 (2), 15–18 ISSN 1609–3631, WWW Page, <http://cran.R-project.org/doc/Rnews>.
- Rice, J., 2001. Implications of variability on many time scales for scientific advice on sustainable management of living marine resources. *Progress in Oceanography* 49, 189–209.
- Rossi, R.E., Mulla, D.J., Journel, A.G., Franz, E.H., 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs* 62 (2), 277–314.
- Simard, Y., Legendre, P., Lavoie, G., Marcotte, D., 1992. Mapping, estimating biomass, and optimizing sampling programs for spatially autocorrelated data: case study of the northern shrimp (*Pandalus borealis*). *Canadian Journal of Fisheries and Aquatic Sciences* 49, 32–45.
- Sissenwine, M.P., Murawski, S., 2004. Moving beyond intelligent tinkering: advancing an ecosystem approach to fisheries. *Marine Ecology, Progress Series* 274, 291–295.
- Smith, B.D., McFarlane, G.A., Saunders, M.W., 1990. Variations in Pacific hake (*Merluccius productus*) summer length-at-age near southern Vancouver Island and its relationship to fishing and oceanography. *Canadian Journal of Fisheries and Aquatic Sciences* 47, 2195–2211.
- Verity, P.G., Smetacek, V., Smayda, R.J., 2002. Status, trends and the future of marine pelagic ecosystem. *Environmental Conservation* 29, 207–237.
- Ware, D.M., McFarlane, G.A., 1995. Climate-induced changes in Pacific hake abundance and pelagic community interactions in the Vancouver island upwelling system. In: Beamish, R.J. (Ed.), *Climate Change and Northern Fish Populations*. Special Publication of the Canadian Journal of Fisheries and Aquatic Sciences, 121, pp. 509–521.
- Wiebe, P.H., Beardsley, R.C., Mountain, D.G., Bucklin, A., 1996. Global ocean ecosystem dynamics: initial program in Northwest Atlantic. *Sea Technology* 37 (8), 67–76.
- Wilson, C.D., Guttormsen, M.A., 1997. Echo integration trawl survey of Pacific hake, *Merluccius productus*, off the Pacific coast of the United States and Canada during July–August 1998. NOAA Tech. Memo. NMFS AFSC-74, Seattle Wash.
- Wilson, C.D., Guttormsen, M.A., Cooke, K., Saunders, M.W., Kieser, R., 2000. Echo integration-trawl survey of Pacific hake, *Merluccius productus*, off the Pacific Coast of the United States and Canada during July–August, 1998. NOAA Technical Memorandum. NMFS-AFSC-118.