

*Collaborative Research:*

Estimating Ecosystem Model Uncertainties in Pan-Regional Syntheses and  
Climate Change Impacts on Coastal Domains of the North Pacific Ocean

**FINAL VERSION** of 7 January 2008

## PROJECT SUMMARY

A sequence of Bayesian Hierarchical Models (BHM) will be developed to synthesize coastal ecosystem dynamics and responses to climate change across focus regions bounding the North Pacific Ocean.

BHM is a unified probabilistic modelling approach that updates uncertain distributional knowledge about process models and parameters in the presence of multi-platform observations (e.g. see Berliner et al. 2003; and references therein). Summary measures of the resulting “posterior” distributions provide realistic quantitative estimates of central tendencies and uncertainties. We will develop our process model distributions after the North Pacific Ecosystem Model for Understanding Regional Oceanography (NEMURO; Kishi et al., 2007). So, a significant outcome of the proposed research will be quantitative understanding and comparisons of the relative uncertainties of NEMURO state variables and parameters, region-by-region across the North Pacific.

A three-step BHM development plan will address pan-regional syntheses, climate change impacts, and ecosystem management tool concepts, over a three-year schedule. The initial BHM development will be a relocatable, time-dependent, one-dimensional (vertical) model intended to summarize ecosystem dynamics for different regimes (shelf, slope, upwelling loci, boundary current extensions, etc.) within the coastal regions of interest. Data and insights from multi-disciplinary observational programs and deterministic model implementations in coastal regions of the North Pacific will be fully exploited. In addition to emphasizing field observations, the BHM methodology will incorporate deterministic model output (e.g. the Regional Ocean Modelling System or ROMS; Haidvogel et al., 2000; Shchepetkin and McWilliams, 2005) as data, providing a rigorous and complete synthesis of the state of understanding for coastal ocean ecosystems of the North Pacific. We will focus on data and models in the Eastern Pacific from parts of the US GLOBEC program (i.e. California Current System, CCS; and Coastal Gulf of Alaska, CGOA) and in the Western Pacific (WPAC) from the North Pacific Marine Science Organization (PICES).

The 1D BHM will also be implemented in climate-scale calculations to document and compare climate change impacts within and across North Pacific coastal ocean ecosystems, and to quantify uncertainties in these comparisons. The ultimate BHM implementation will be in three dimensions, accounting for mesoscale ocean dynamical impacts on the coastal ecosystem regions, and demonstrating potential ecosystem management advantages of the BHM approach.

The intellectual merit of this proposal derives from a novel extension of probabilistic modelling methods (i.e. BHM) to synthesize disparate observations and deterministic model simulations from coastal regions on eastern and western boundaries of a major ocean basin. Application of BHM in Biological Oceanography represents a transformative research step and introduces a new paradigm. The research proposed here combines the strengths of deterministic and probabilistic models to obtain uncertainty estimates for state variables and parameters of a modern lower-trophic level ocean ecosystem model. A broader impact of the research will be the training of postdoctoral and graduate students (in statistics and oceanography) in this new synergy of ocean modelling approaches. As ecosystem managers and scientists learn to utilize state and parameter information in probability distributions, uncertain parts of the ecosystem model can be targeted for more intensive observations and/or more sophisticated parameterizations.

# PROJECT DESCRIPTION

## 1. Introduction

Three questions guide our approach to pan-regional synthesis and climate-change impacts studies for coastal ocean ecosystems bounding the North Pacific Ocean. They are:

- 1) What do observations and deterministic model results tell us about the states and dynamics of each of the coastal ocean ecosystem focus regions?;
- 2) Given observational data and deterministic model results, how confident can we be in the state estimates (including estimates for interaction parameters)?; and
- 3) How do these uncertainty estimates compare from region to region, and under the influence of a changing climate?

The goal of the proposed research is to quantitatively address these questions by means of Bayesian Hierarchical Models (BHM), implemented for coastal ocean ecosystems in the US GLOBEC study regions of the CCS and CGOA, and the PICES focus region in the WPAC. BHM is a probabilistic modeling methodology that combines probability distributions for observations (i.e. as in measurement error models) and physical-biological process models linking state variables and parameters for the ecosystems of interest. The posterior probability distribution output of BHM leads directly (e.g., via second moment estimates) to quantitative statements regarding uncertainty in state variables and parameters of the ecosystem. Also, BHM provides a formally rigorous framework within which deterministic model output can be treated as data, and synthesized with multi-platform observations, as well as with scientific understanding of key processes and state variables.

Objectives of the proposed research include:

- 1) implement 1-D (vertical) BHM for important regimes within each North Pacific coastal ocean ecosystem focus region to develop comparable posterior probability distributions (1D summary BHM);
- 2) couple 1-D BHM to climate-scale physical models for extended calculations to document climate-change impacts and uncertainties (1D climate BHM); and
- 3) implement 3-D regional BHM to emphasize mesoscale ocean dynamics impacts on coastal ocean ecosystem evolution, leading toward a probabilistic ecosystem research and management tool.

The pan-North Pacific coastal ocean ecosystem context for this proposal is described in section 2, and key results from PICES and US GLOBEC work to date are reviewed. The reviews serve to: a) introduce the NEMURO applications relevant to our proposal; and b) introduce ROMS implementations in the CCS and CGOA. A review of BHM implementations in geophysical and ecological systems is provided in section 3. Fundamental concepts of BHM are illustrated in a sample development of a reduced-complexity zero-dimensional (0-D) BHM based on an NPZ approximation for ecosystem dynamics. The proposed research is described in section 4, including details of the three-step development plan for increasingly sophisticated BHM implementations, and a new ROMS implementation in the WPAC to support BHM data stage inputs. Section 5 outlines a management plan and schedule. Results from prior NSF supported research are listed for all PI's in section 6.

## 2. Deterministic Physical-Biological Models for the North Pacific

A special issue of *Ecological Modelling* (volume 202, March 2007) documents climate- and synoptic-scale implementations of NEMURO in the WPAC <sup>1</sup>. In its basic form, NEMURO is an 11-compartment, lower-

---

<sup>1</sup>The NEMURO development is a product of the PICES, Climate Change and Carrying Capacity program (CCCC). See Batchelder and Kashiwai, 2007, and Werner et al., 2007, for overviews and summary articles.

trophic level ecosystem model that can be interfaced with physical circulation models of basin and/or regional scales (e.g. Aita et al., 2007; Hashioka and Yamanaka, 2007 a,b; Komatsu et al., 2007; Fiechter et al., 2008b). The NEMURO connections to ocean circulation models are one-way in that ecosystem responses to the physical environment (light intensity, temperature, currents, and stratification) do not feed back to the circulation. As part of US GLOBEC, Fiechter et al. (2008b) have augmented NEMURO to include iron cycling, and coupled this model with ROMS (i.e. ROMS-NEMURO) in the CGOA. Similarly, preliminary results from ROMS-NEMURO are being obtained now by the U.C. Berkeley group (Powell and co-workers) in the CCS. WPAC, CGOA and CCS coupled physical-biological coastal ocean ecosystem model developments are reviewed in this section.

## **2a. PICES Results with NEMURO in the WPAC**

Several papers in the special issue of *Ecological Modelling* document NEMURO applications in the Western Pacific boundary current separation region off Japan. These will provide essential guidance and comparisons for: a) our implementations of ROMS-NEMURO in the WPAC; and b) the 1D and 3D BHM to be developed in the region. WPAC 1D NEMURO implementations at ocean mooring stations include Yoshie et al. (2007) at station A7, and Fujii et al. (2007) at stations A7, KNOT and PAPA (Figure 1). Fujii et al. (2007) added process equations and parameters for carbon cycling in the ocean to NEMURO. They noted a need for high-frequency wind forcing to properly model the impacts of intermittent, large-amplitude atmospheric forcing events (e.g. associated with frontal passages), that can reset physical and biological ocean background states.

The WPAC 3D NEMURO implementations are coupled with coarse resolution Ocean General Circulation Models (OGCMs) to study seasonal and climate variations in phytoplankton abundance and responses to global warming (e.g. Hashioka and Yamanaka, 2007a,b; Aita et al., 2007). Our climate-change 1D BHM implementation plans are relevant to these studies. We will aggregate data by oceanic regime to provide data-rich data stage distributions for seasonal and longer timescale BHM calculations (see section 4).

Komatsu et al. (2007; Fig. 1) link a high-resolution, data-assimilating (sea-surface height, SSH) regional ocean circulation model with NEMURO on sub-daily timescales, for the period January through June, 1997. They connect regional synoptic-scale ocean circulation, including meanders and eddies of the Kuroshio, with ecosystem dynamics. The physical-biological interface is effective in demonstrating high-resolution properties of phytoplankton bloom and zooplankton population dynamical response. Useful inferences regarding Pacific Saury abundance and survival can be drawn from the high-resolution 3D NEMURO analyses. The Komatsu et al. (2007) study is relevant to our 3D BHM implementation plans. Their results will be used to validate our ROMS-NEMURO implementation in the WPAC that will precede BHM development there.

The WPAC NEMURO studies provide a consistent picture with respect to the control exerted on the coastal ocean ecosystem by the physical environment. Upper ocean ecosystem rates (i.e. grazing, egestion, etc.) increase with increasing upper ocean temperatures. The warming upper ocean also becomes more stratified, and mixed layer depths shallow. Upper ocean nutrients are depleted more rapidly, and more energy is required to mix nutrients up from below ocean mixed layers. Phytoplankton, and later zooplankton, blooms are smaller, and abundances are lower. There is evidence in Hashioka and Yamanaka (2007b) of a shift from large zooplankton (diatoms) dominance to smaller zooplankton that are more adaptable to changing ocean temperatures.

## **2b. US GLOBEC Results with ROMS-NEMURO in CGOA**

While a wide range of physical-biological models of varying complexity have been developed for the U.S. West Coast upwelling system (e.g. Wroblewski, 1977; Franks et al., 1986; Spitz et al., 2003; Powell et al., 2006) and the NE subarctic Pacific Ocean (Denman and Pena 1999, 2002), only two very recent studies have attempted to specifically address the role of iron limitation on phytoplankton growth in the CGOA with 3D, fully-coupled ocean circulation-ecosystem models (Fiechter et al., 2008a; Hinckley et al., 2008).

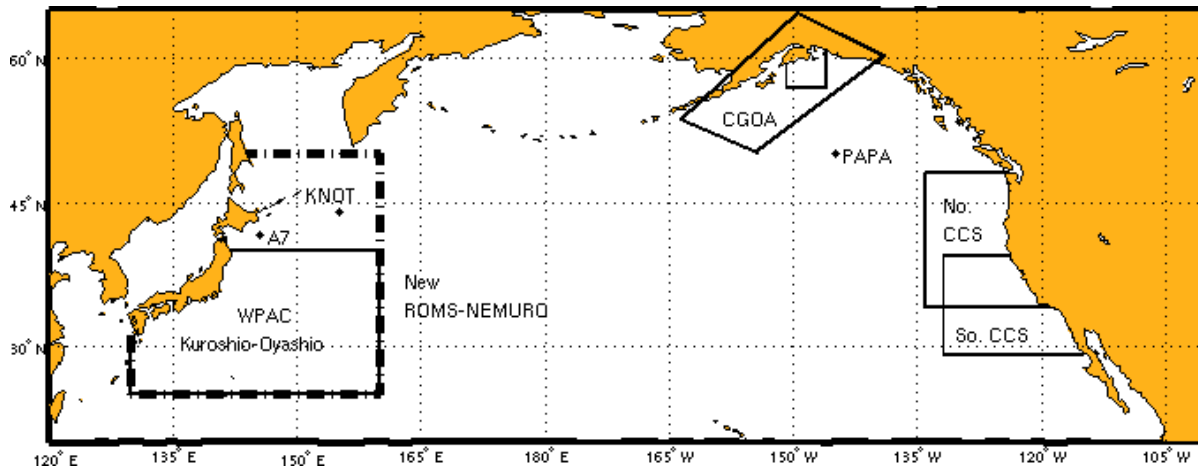


Figure 1: North Pacific Ocean coastal domain locator map. Long term observing stations, A7 and KNOT are located in the NW Pacific, and station PAPA is in the NE Pacific. US GLOBEC domains in the CGOA (Fiechter et al., 2008b) and CCS (North, as in Powell et al., 2006; South, as in Moore et al., 2007) are noted on the eastern boundary. The US GLOBEC Long Term Observations Campaign box in Prince William Sound is outlined inside the CGOA domain. A Kuroshio-Oyashio domain used by Komatsu et al. (2007) is noted on the western boundary (thin solid line). The proposed new ROMS-NEMURO domain for the WPAC extends this domain to the North (bold dash-dot line).

The ocean circulation model for the NW CGOA is an implementation of ROMS (Fig. 1). ROMS is a hydrostatic, primitive equation model that uses terrain-following  $\sigma$ -levels for the vertical coordinate, and a split-mode technique to solve efficiently for internal (i.e., depth-dependent) and external (i.e., depth-integrated) variables. The model grid for the CGOA has a horizontal resolution of 10 km and 42 non-uniform  $\sigma$ -levels, with clustering near the surface. The ROMS for the CGOA is driven on all open boundaries by sea surface height, velocity, temperature and salinity fields from the Northeast Pacific (NEP) ROMS simulations of Curchitser et al. (2005) at the same horizontal and vertical resolution. This approach allows the specification of realistic transport values and temperature and salinity profiles for the Alaskan Stream entering and exiting the CGOA domain (Figure 2). The surface forcing for the CGOA model is derived from the NEP model forcing (Common Ocean-Ice Reference Experiments, CORE; Large and Yeager, 2004) in the form of wind stress, freshwater ( $E - P$ ) flux, net heat flux, and short-wave radiation. Similar boundary and surface forcing strategies will be employed in the proposed ROMS development for the WPAC.

Ecosystem complexity in the NW CGOA is investigated by coupling NEMURO, with explicit iron limitation on nutrient uptake, to ROMS. Iron limitation is included in the ecosystem models by adding governing equations for micro-nutrient compartments representing dissolved iron and phytoplankton-associated iron (Fiechter et al., 2008b). Simulated nitrate and chlorophyll concentrations with NEMURO exhibit striking similarities with available remotely-sensed and *in situ* observations. In addition to reproducing the climatological seasonal variability, the coupled model provides important information on the different modes of variability affecting the physical and biological variables as a function of cross-shelf position.

The ecosystem complexity in NEMURO supports the reproduction of spatial and temporal features of the phytoplankton and zooplankton community structure in the CGOA. Figure 2 depicts a snapshot of the CGOA ROMS-NEMURO output for SSH and coastal currents (left panel) and surface chlorophyll (right panel) for 1 May 2001 (i.e., after 76 months of simulation). For our BHM developments in the CGOA, we will make use of the existing ROMS-NPZD (Fiechter et al., 2008a) and ROMS-NEMURO (Fiechter et al., 2008b) simulations.

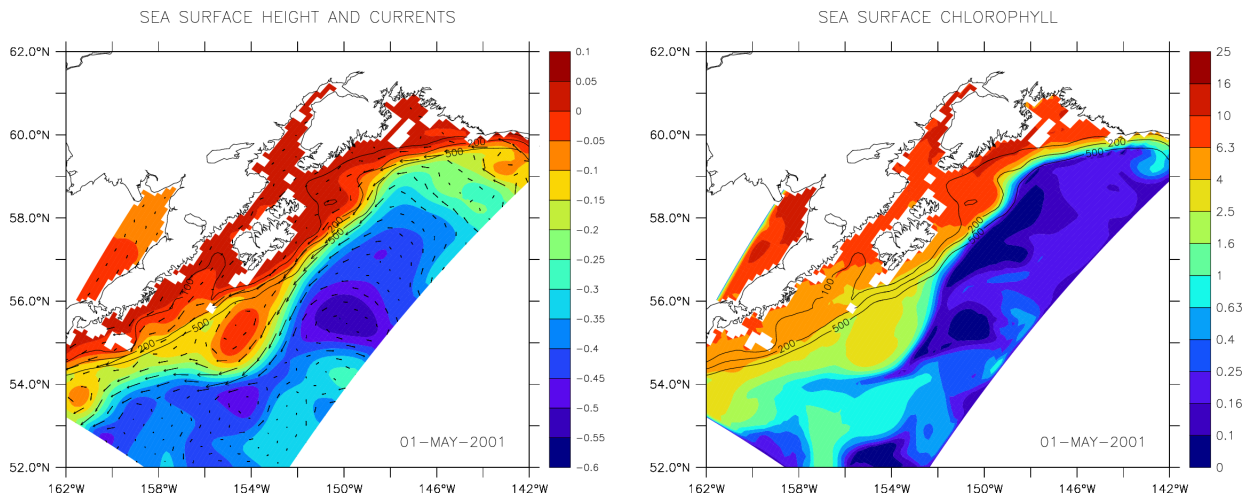


Figure 2: Sample output from the ROMS-NEMURO model in CGOA (Fiechter et al., 2008b), for 1 May 2001. Sea-surface height ( $m$ ), and surface currents are depicted in the lefthand panel, bathymetry and chlorophyll concentration are shown at right. The spring phytoplankton bloom concentrates chlorophyll nearshore, with the Alaskan stream posing a barrier to offshore phytoplankton abundance. Eddies and meanders of the Alaska stream are also evident in the chlorophyll concentration.

## 2c. Results with ROMS-NPZD in the CCS

The coupled physical-biological environment in the CCS has also been modeled using ROMS for a variety of model configurations and resolutions. Powell et al. (2006) considered the northern portion of the CCS centered on the the GLOBEC region (see No. CCS in Fig. 1). The model resolution was  $3\text{ km}$  in the horizontal with 30 levels in the vertical ranging in thickness from  $0.7\text{ m}$  at the coast to  $15\text{--}450\text{ m}$  in the deep ocean. The model was driven by NCEP daily surface forcing, and open boundary conditions derived from a suite of nested ROMS integrations for the NEP by Curchitser et al. (2005). The ecosystem component was an NPZD model similar to that of Spitz et al. (2003), but with modified saturation forms for phytoplankton photosynthetic response and zooplankton grazing.

Powell et al. (2006) compared the NPZD model simulations of phytoplankton to satellite measurements of surface chlorophyll from the Coastal Zone Color Scanner (CZCS). The model-simulated chlorophyll variability was remarkably consistent with spatial and temporal scales in CZCS observations within the GLOBEC study region.

Di Lorenzo et al. (2008) configured ROMS and the same NPZD model for the North American west coast  $25^{\circ}\text{N} - 62^{\circ}\text{N}$  and extending to  $180^{\circ}\text{W} - 110^{\circ}\text{W}$ , with  $15\text{ km}$  horizontal resolution and 30 levels in the vertical. The model was forced with NCEP reanalysis products for the period 1950-2006. Comparisons of the physical model with satellite and *in situ* observations show that the model captures much of the low frequency variability. In addition, the NPZD model also captures much of the low frequency variability in near surface nitrate and chlorophyll measured by the CalCOFI array. Di Lorenzo et al. show that this low frequency variability is associated with low frequency variations in the strength of the subtropical-subpolar gyre circulation in the NE Pacific.

Moore et al. (2007) have used ROMS and the same NPZD model in the southern portion of the CCS and Southern California Bight, (see So. CCS in Fig. 1) with  $20\text{ km}$  resolution and 20 levels in the vertical. The model was driven by NCEP climatological surface fluxes, and open boundary conditions taken from the Di Lorenzo et al. (2008) simulations. The focus of the Moore et al. study was the sensitivity of the physical environment to surface forcing using the adjoint of ROMS. It was found that various aspects of the

circulation (coastal SST and the potential for baroclinic instability) display pronounced seasonal sensitivities to variations in surface forcing (variations of almost an order of magnitude between years), and that local wind stress and wind stress curl can be important for controlling eddy kinetic energy. The second part of their study explores the sensitivity of the NPZD biological conditions to the physical environment and surface forcing (Moore et al., 2008).

The ROMS-NPZD models of Powell et al. (2006) and Moore et al. (2007) will also provide data stage inputs for 1D and 3D BHM developments to be described. The climate-scale 1D BHM development could be informed by summaries of the Di Lorenzo et al. (2008) longer-term ROMS-NPZD calculation.

NPZD models work well in upwelling-dominated environments such as the Southern CCS. In contrast, in the CGOA, NPZD and NEMURO require the addition of iron as a limiting nutrient in order to produce realistic simulations.

### 3. BHM Background

We review some fundamental concepts of BHM (section 3a) that lead to the development and interpretation of the sample, 0-D BHM based on an NPZ ecosystem model (section 3b). BHM is proving to be particularly appropriate for large degree of freedom (*dof*) geophysical and ecological systems as noted in section 3c below.

#### 3a. BHM Introduction

Berliner and co-workers (e.g. Berliner 1996; Royle et al., 1998; Berliner et al., 2000, 2003; Wikle et al., 2001; Wikle, 2003a,b; etc.) organize the development of BHM for large *dof* environmental systems in 3 stages of probability distribution hierarchies; i.e, a data stage, a process model stage, and a stage for parameter distributions. The 3 stages of distributions lead, via Bayes Theorem and Markov Chain Monte Carlo (MCMC) methods, to estimates of the posterior distribution of interest.

Symbolically, let the probability distribution of a random variable  $X$  be  $[X]$ . Let the dependence of  $X$  on another random variable  $Y$  be expressed as conditional distribution  $[X|Y]$ , read: “the distribution of  $X$  given  $Y$ ”. At the highest level of a probabilistic model hierarchy let  $X$  represent the coupled biological-physical ecosystem processes of interest. Let  $Y$  represent observations and deterministic model output relevant to  $X$ . Let  $[\theta_d]$  and  $[\theta_p]$  be parameter distributions that arise in the data stage and process model stage, respectively. Bayes Theorem states:

$$[X, \theta_d, \theta_p|Y] \propto [Y|X, \theta_d][X|\theta_p][\theta_d][\theta_p]. \quad (1)$$

Here, the first term on the righthand side is the data stage distribution  $[Y|X, \theta_d]$ . It identifies the observations conditional on the true (but unknown) state process and parameters. Thus, it accommodates measurement errors as well as representativeness errors between true process and data, and true process and deterministic model output. The next term in (1) is the process model stage distribution, given by  $[X|\theta_p]$ . It is a statement of our best scientific description of the ecosystem processes (i.e. we will base this on NEMURO in the proposed research). In statistics, this is called the *prior distribution* because the model of the process is posed prior to the consideration of the data in  $Y$ . To complete the righthand side of (1), the data and process model parameters are also given prior distributions,  $[\theta_d]$ ,  $[\theta_p]$ , to reflect understanding of their central tendencies and associated uncertainties. The formulation in (1) assumes that the data and process parameters are independent, *a priori*. This is a reasonable assumption, made for convenience here, but need not be the case in general. The lefthand side of (1) is the BHM output object of interest. This is the *posterior* distribution for the process  $X$  and the parameters  $\theta_d$  and  $\theta_p$ , given the data  $Y$  (i.e. observations and model output).

In the coastal ocean ecosystem context the analogous BHM includes the distributions of nutrients and plankton species, as well as the parameters linking their dependencies and expressing the representativeness

of observations and deterministic model solutions. There are more than 70 such parameters (i.e.  $\theta_p$ ) in the NEMURO model upon which we will base our process model distributions for the 1D and 3D BHM. Additionally, a significant part of our research effort will include developing estimates for data stage distribution parameters  $\theta_d$  representing error models for observational data (i.e. from US GLOBEC and PICES), and representativeness errors for ROMS-NEMURO and ROMS-NPZD output.

It is important to emphasize that BHM will yield posterior *distributions* on all the interaction parameters  $\theta_p$  that arise in the ecosystem model expression. In instances where the parameter distributions are peaked, there is evidence of Bayesian learning, i.e., the data has contributed significant information over the prior, leading to new scientific insight. Conversely, if the posterior parameter distributions are flat, little learning has occurred given the data and process model formulations of the BHM. Both cases are important for understanding and developing simplified models of complex real-world systems. This parameter-specific information is quantitatively available via BHM, and *distributions* for parameters and state variables can be compared across regions and from different climate states.

To express Bayes Theorem in (1) as an equality, a normalizing constant is needed in the denominator of the righthand side. The normalizer is an integral over all possible states of the process  $X$  and all possible parameter values; a term that is intractable in large *dof* systems with complicated hierarchies describing the data stage and process model distributions. Instead, realizations from the posterior distribution can be obtained by specialized Monte Carlo algorithms (e.g., MCMC and Importance Sampling Monte Carlo; see for example the overview in Robert and Casella, 2004). These are the numerical theory and numerical methods analogs in BHM, and their implementation is an area of active research within which members of our team are leading researchers (e.g. Berliner and Wikle, 2007).

### 3b. Sample 0-D BHM: NPZ model to find parameter distributions

A time-dependent BHM for 3 parameters of an NPZ model is developed here to demonstrate features of the methodology. The demonstration is a drastic simplification of the BHMs proposed below, but the concepts of random variables and posterior distributions are evident nonetheless.

Let the  $X(t)$  be a  $3 \times 1$  vector  $X(t) = [N(t), P(t), Z(t)]'$  (the prime indicates a vector transpose) for the state of a Nutrient-Phytoplankton-Zooplankton system at time  $t$ . The evolution of the system is modelled via 4th-order Runge-Kutta integration (with a timestep of 0.1 *day*) of the NPZ equations in the appendix of Wainwright et al. (2007), given known initial conditions. In a more realistic BHM implementation, the NPZ equations would be discretized, each term treated as random, to form a process model stage hierarchy (e.g., see Berliner et al. 2003). Also in realistic applications, the initial state need not be completely known.

In the present case, we assume the only uncertainty in the process model derives from 3 parameters we treat as random variables; i.e the zooplankton maximum grazing rate ( $GRmax$ ), the maximum phytoplankton growth rate ( $Vmax$ ), and the phytoplankton mortality rate ( $\mu_P$ ). The remaining 8 parameters of the NPZ model are fixed (i.e. not random) at the values in Wainwright et al. (2007), except the Ivlev constant which is reset to 0.84 (Fiechter et al., 2008a). Let  $\theta_p = [GRmax, Vmax, \mu_P]'$  be the vector of random parameters. Then the evolution of the state is symbolized by  $X(t; \theta_p)$ . Let  $X(\theta_p)$  denote the collection of  $X(t; \theta_p)$  for all times of interest. Note, we neglect an explicit model error term in this simple example.  $X(\theta_p)$  is a random process only because it is conditioned on the random parameters  $\theta_p$ .

Simulated data for the BHM is taken from 3-daily N, P and Z summaries of the CGOA ROMS-NPZD (with iron) implementation described in section 2, from a nearshore regime (i.e., inshore of the Alaska Stream). Let  $Y(t)$  be a  $3 \times 1$  vector of the data for N, P, Z at time  $t$ . The data stage distribution is:

$$Y(t)|\theta_p, \theta_d \sim N(X(t; \theta_p), R(\theta_d)) \quad (2)$$

which is a measurement error model that can be read: “the distribution of the observations, given random

parameters and variances, has a multivariate normal distribution with mean given by  $X(t; \theta_p)$ , and variance given by a matrix  $R(\theta_d)$ .<sup>2</sup> Here,  $R(\theta_d)$  is a  $3 \times 3$  diagonal matrix with the variance parameters,  $\theta_d = [\sigma_{eN}^2, \sigma_{eP}^2, \sigma_{eZ}^2]'$ , on the diagonal, and 0's at all off-diagonals, where  $\sigma_{eI}^2$  are the variances in the error models for  $I = N, P, Z$ . Clearly, more complicated error models can (and will) be considered in the proposed BHM.

We must specify the parameter distributions  $[\theta_p]$  and  $[\theta_d]$  to complete the Bayesian model. For this simple example, we let the  $\theta_p$  be uniformly distributed over ranges discerned from Fennel et al. (2005), as:

$$GRmax \sim Uniform(0.5, 1), \quad Vmax \sim Uniform(0.5, 3), \quad \mu_p \sim Uniform(0.05, 0.2), \quad (3)$$

where  $Uniform(a, b)$  denotes a uniform distribution over the range  $a$  to  $b$ . More complicated (informative) prior distributions could be specified if warranted. In the more complicated settings we will encounter in the proposed research, many of the model parameters have *a priori* dependence that will be included in the prior distribution formulation.

For ease of computation in the MCMC, we let each of the  $\theta_d$  variances have inverse gamma prior distributions with mean = 100, and variance = 100 in each case. The relatively large variances accommodate anticipated representativeness errors between the “data” (simulated with a more complicated model) and the simple 3-compartment NPZ model. Again, the formulation of data models is a key component of the proposed research and more complicated formulations are anticipated.

In this simple example Bayes theorem (i.e. as in (1)) takes the form:

$$[\theta_p, \theta_d | Y] \propto [Y | X(\theta_p), \theta_d][\theta_p][\theta_d] \quad (4)$$

where the process model stage distribution is absent since we are neglecting (additive or multiplicative error) uncertainty in the prior for simplicity of exposition. Thus, in this case, model uncertainty can only be accounted for by the three parameters in  $\theta_p$  and model representativeness is accounted for only by the parameters  $\theta_d$ .

An MCMC algorithm is designed to sample the  $\theta_p$  and  $\theta_d$  given the data. The state  $[X(\theta_p)]$  is obtained by integrating the model forward for each posterior sample of  $\theta_p$  (this is why  $X(\theta_p)$  does not appear explicitly in the posterior; i.e. the lefthand side of (4)). The preliminary results presented here are based on a 10,000 sample MCMC integration.

Figure 3 depicts the distributions for the state vector  $X$  (N in blue, P in red, and Z in green) over 200 day time series (year-day number on the x-axis). The solid color lines are the posterior mean values, the asterisks indicate the input data from the CGOA ROMS-NPZD summary, and the dashed, dash-dot, and dotted lines denote percentiles of the distributions. For example, 95% of the posterior distributions for N, P, and Z fall between the dotted and dash-dot lines in each panel. This *spread* of the distribution is a measure of the uncertainty in each state variable.

Competing influences between the simplified processes of the NPZ model and the data from a model that includes vertical resolution, ocean circulation, and detritus and nutrient remineralization are evident in Fig. 3. Centered on about year-day 125, the dominant Spring phytoplankton bloom occurs and nutrients are depleted. In response, a zooplankton bloom occurs at about year-day 140. Smaller, secondary phytoplankton and zooplankton blooms follow in sequence as nutrients resurge around year-day 150. Following these blooms, the nutrient supply resurges again after year-day 250 as nutrients are mixed vertically into the upper ocean, and phytoplankton and zooplankton die and recycle. Following the Spring bloom and throughout the Summer (days 150-250), the nitrate in the NPZ process model of the BHM remains high compared to the data because nitrogen is recycled immediately. The data on the other hand are from ROMS-NPZD

<sup>2</sup>Note that, different from the expression for the data stage in (1), the data  $Y$  are also conditioned on the process model parameters  $\theta_p$  here. This is because they appear explicitly in the expression for the mean value on the righthand side of (2).

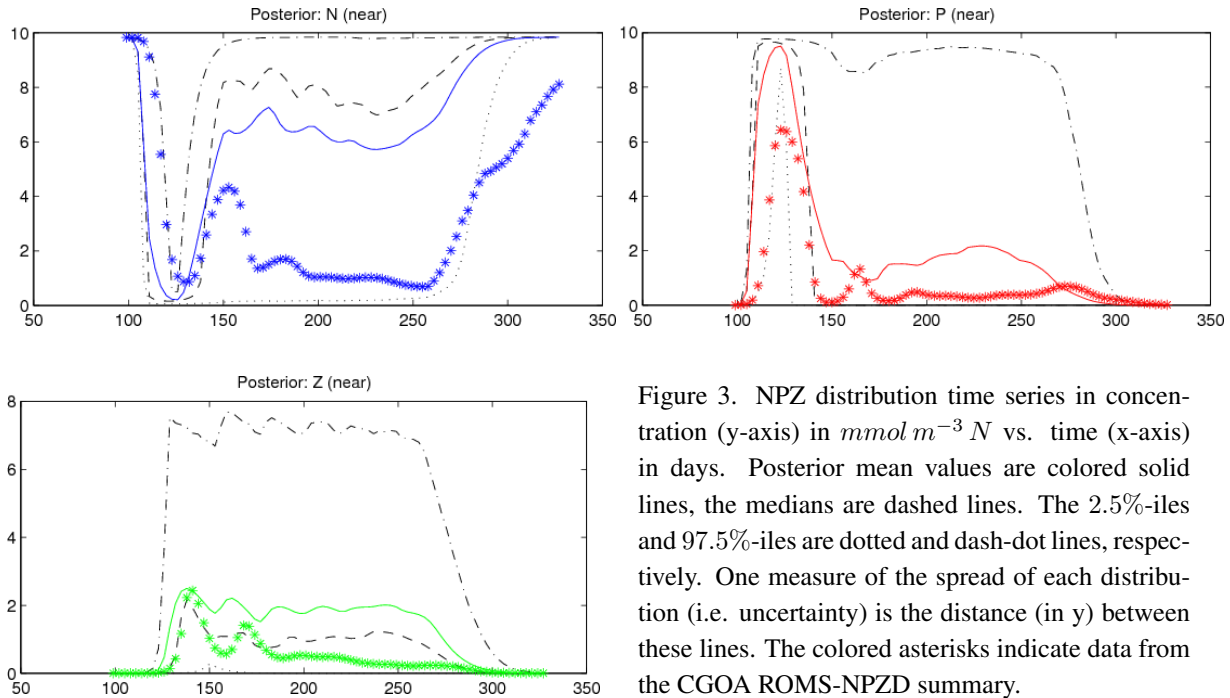


Figure 3. NPZ distribution time series in concentration (y-axis) in  $mmol\ m^{-3}\ N$  vs. time (x-axis) in days. Posterior mean values are colored solid lines, the medians are dashed lines. The 2.5%-iles and 97.5%-iles are dotted and dash-dot lines, respectively. One measure of the spread of each distribution (i.e. uncertainty) is the distance (in y) between these lines. The colored asterisks indicate data from the CGOA ROMS-NPZD summary.

where nitrogen is returned more slowly to the nitrate pool via detritus and remineralization. The data are influencing the posterior mean, but the competition between the (over-simplified) NPZ process model and data is adding uncertainty (note wider spreads after year-day 150).

Figure 4 depicts the initial and posterior distributions for the parameters  $\theta_p$ . Note that the posterior distribution for the zooplankton maximum grazing rate parameter ( $GR_{max}$ ) is not different from its uniform prior distribution. Thus, the data has not informed this parameter. However, the phytoplankton growth rate parameter ( $V_{max}$ ) does show a substantial difference from its uniform prior, with a strong preference for the lower levels of its range. Similarly, the phytoplankton mortality rate parameter ( $\mu_P$ ) also shows learning with a strong preference for values at the upper end of its range.

These posterior distributions for parameters  $\theta_p$  are consistent with the competing influences of an NPZ process model and data from a more complicated ecosystem representation (i.e. ROMS-NPZD). The tendency of the BHM to overestimate the amount of nitrate in Fig. 3 leads to phytoplankton and zooplankton concentrations that are higher than the data. So the NPZ model parameters adjust to accommodate the data by decreasing the phytoplankton growth rate, and increasing the phytoplankton mortality rate, so that phytoplankton abundance is more consistent with the data. There is no information in the data to move zooplankton grazing rates off the vague initial specification.

Taken together, the posterior distribution for parameters here,  $[\theta_p]$ , (Fig. 4) is indicating a model misfit. The parameters are either remaining vague (i.e.  $GR_{max}$ ) or trending toward values outside a reasonable range (i.e.  $V_{max}$ ,  $\mu_P$ ). Consequently, the uncertainty (i.e. spread) in the state variables is large (Fig. 3). In this case the BHM method is providing information more fundamental than uncertainty estimates. It is suggesting the model is not reasonably supported by the data.

*This simple example demonstrates the interplay between model formulation, data, and interpretation of posterior distributions that will characterize our research. These considerations will be much more complicated in the NEMURO process model context proposed here. Nonetheless, techniques of fixing vs. randomizing parameters and identifying limiting physical-biological cases for experimentation, will carry over to the developments proposed here (section 4).*

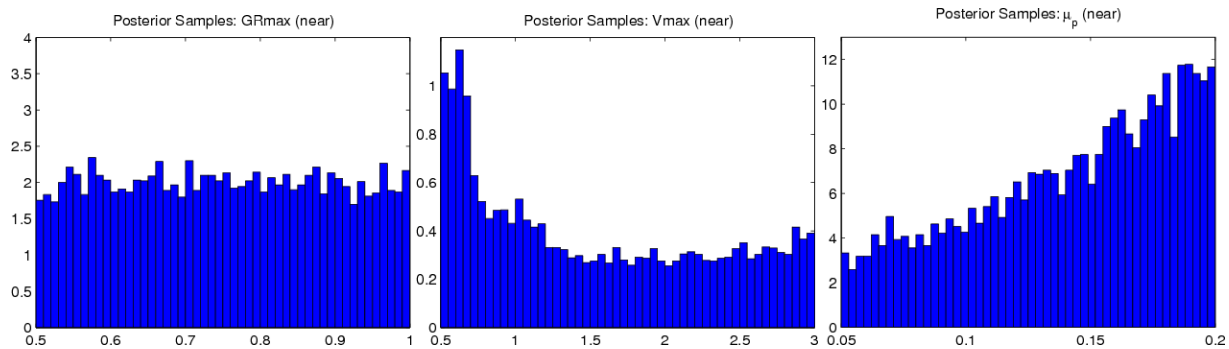


Figure 4: BHM posterior histograms for random parameters:  $GR_{max}$  (left),  $V_{max}$  (middle), and  $\mu_p$  (right). The impact of unmodelled vertical mixing and light limitation can be interpreted from these posterior distributions.

### 3c. BHM in Ecology and Geophysical Fluid Dynamics

Ecologists are increasingly interested in accounting accurately for various sources of uncertainty when modeling ecological processes. This is largely motivated by an interest in providing an accurate assessment of the state of knowledge of ecosystems and the need to quantify the uncertainty associated with predictions of local and global change (e.g., Clark et al. 2001). BHM provides a coherent probabilistic framework in which to quantify uncertainty for ecosystem processes (e.g., see the summaries related to environmental and ecological processes by Wikle, 2003a; Clark and Gelfand 2006; and Clark 2007). Relatively recent advances in computational efficiency and algorithmic development have allowed for the implementation of sophisticated BHMs in ecology (e.g., Calder et al. 2003; Clark 2003; Wikle 2003; Buckland et al. 2004; Su et al. 2004; Gelfand et al. 2005; Thomas et al. 2005; Wikle and Hooten 2006; Hooten et al. 2006; Hooten and Wikle 2007). It has been shown that BHMs of this class have much to offer, including more precise and less biased parameter estimation (Calder et al. 2003), accounting for multiple sources of uncertainty (Wikle 2003), and the ability to describe non-linear spatio-temporal dispersal and growth in relative abundance (Hooten and Wikle 2006), as well as dynamics of population demographics (Buckland et al. 2004; Hooten et al. 2006).

In a series of relatively recent papers, Wikle and colleagues utilize the BHM framework to model invasive species dynamics. Although not the subject of the current proposal, the problem is similar in accounting for uncertainties, and modeling systems with large *dof* in an efficient manner. In characterizing the house finch invasion of the Eastern U.S., Wikle (2003) was concerned with uncertainties in the Breeding Bird Survey (BBS) counts, uncertainty in the spatio-temporal dynamical (reaction-diffusion) process that characterized the growth and spread, and uncertainty in the associated parameters (e.g., a spatially dependent diffusion process). In the data stage of this model, the BBS bird count observations were modeled conditionally given the true (unobserved) bird counts.

Scientific information entered the invasive species BHM at the process model stage. A relatively simple reaction-diffusion equation (in two spatial dimensions) was considered. Model error terms, and the diffusion and growth parameters were also assumed random (i.e. endowed with probability distributions to be estimated in the posterior). The diffusion coefficients were assumed to be spatially heterogeneous, and thus spatially explicit models (that can easily be related to known covariates; e.g., human population in the case of the House Finch) were assigned to those coefficients at the parameter distribution stage of the BHM hierarchy. Given that the model was introduced on “landscape” scales, the dimension of the process that was being predicted (over space and time), as well as the number of parameters (diffusion and growth parameters for each grid box) were quite large. However, MCMC implementation was straightforward (see the example in Wikle and Hooten, 2006). The posterior analysis was able to provide some insight into the likely Allee<sup>3</sup> effect early in the invasion, as well as the importance of the heterogeneous spatial diffusion coefficients.

<sup>3</sup>A positive relationship between population density and the reproduction and survival of individuals.

In the research proposed here, we must also be able to account for uncertainties associated with combining physical and biological models, as well as being able to combine multiple sources of information (e.g., *in situ* data as well as deterministic model output “data”). There are recent examples which address these issues to various degrees. For example, Eknes and Evensen (2002) use an ensemble Kalman filter (EnKF) to assimilate observations in a one-dimensional three component ocean ecosystem model. The EnKF is a sequential approach that gives the approximate Bayesian posterior distribution of the state process in the presence of uncertain observations. Although such an approach cannot accommodate the complicated parameter model structures we will encounter in our BHM, it suggests that other types of approximate Bayesian procedures (e.g., approximate importance sampling) might be useful (Berliner and Wikle, 2007). Losa et al. (2003) implemented a sequential importance resampling filter to assimilate data into a nine-component ecosystem model in order to obtain parameter estimates. One difference between their approach and the research proposed here is that we explicitly account for random parameters in our model formulation, and allow those parameters to have complicated relationships (e.g., interdependence, spatial dependence, etc.) in a hierarchical framework. For example, Berliner et al. (2003) demonstrate that a relatively simple process model (quasi-geostrophic dynamics) with random parameters can adequately reproduce the dynamics from a more complicated (primitive equation) shallow-water system, in the presence of relatively few observations. Berliner et al. (2003) is important in that it also shows how one can couple systems (atmosphere and ocean) hierarchically, as well as accommodate random boundary conditions (see also Wikle et al. 2003). Wikle et al. (2001) show how one can accommodate model output as data, along with satellite observations in a hierarchical model for tropical winds. The Wikle et al. (2001) BHM encompassed a very large state space (order  $10^5$ ) and yet was still feasible in the MCMC environment (see also Hoar et al. 2003).

#### **4. Proposed Research: A Step-Wise BHM Development Plan**

Using BHM for pan-regional syntheses and climate-change impacts for coastal ocean ecosystems is a new application for this methodology that requires a step-wise approach. BHM is being used here to combine observations and deterministic model output as a synthesis tool such that common elements of posterior distributions can be compared across regions, and between regimes within regions. In comparing probability distributions, we will identify best estimates for coastal ocean ecosystem parameter values, and the uncertainties associated with these estimates.

##### **4a. Relocatable Summary BHM (1D)**

The initial BHM development will be a relocatable, time-dependent, 1-D (vertical) summary model to be used as a synthesis tool within and across the focus regions. The process model form will increase in complexity from the NPZ model in the example above (section 3b), to a process model based on the 1D NEMURO equations (Kishi et al., 2007), including more than 70 model parameters. The 1D Summary BHM will be implemented for each region (i.e. WPAC, CGOA, CCS), and in important sub-regions within each domain (i.e. shelf vs. slope, at upwelling loci, inshore and offshore of western and eastern boundary currents, etc.) Pan-regional comparison calculations will be run for several annual cycles. Implementing this first reduced-dimension BHM will also provide us with a BHM tool for testing process model complexities and developing efficient MCMC procedures.

As part of the 1D summary BHM development, we will augment the observational data summaries with output from ROMS-NEMURO deterministic models in each region. A ROMS-NEMURO implementation in the WPAC will begin immediately (Fig. 1). ROMS-NEMURO models can already be run in the CGOA and CCS. Characteristic physical process scenarios will be synthesized from regime subsets of the ROMS-NEMURO and ROMS-NPZD simulations to be used as data in the 1D BHM. For example, the offshore current effects, pycnocline depth and undulations characteristic of the continental slope regime north of Pt. Conception can be summarized from the CCS ROMS-NPZD simulations in that regime. Similarly,

summaries for shelf regions off Northern California and Oregon can provide data-stage inputs for 1D BHM implementations for the Eastern Boundary, Northern CCS, continental shelf regime. Effective summaries from ROMS and implementations in critical regimes are interdisciplinary decisions that will have critical impacts on the utility of the proposed work.

In the face of persistent data sparsity for ecosystem state variables and parameters, the likelihood that all of these parameters can “learn”<sup>4</sup> from the data is questionable. We will iterate the development of the 1D summary BHM, fixing some parameters while treating others as random coefficients (i.e. with probability distributions). Through the quantitative treatment of uncertainty in data stage distribution specifications (i.e.  $[Y|X, \theta_d]$  in (1)), the deterministic model output can be weighted appropriately with respect to field observations that are more scarce. The impact on parameter identifiability of including deterministic model output in data stage distributions is an innovation in BHM applications (e.g. Wikle et al., 2001) that is being extended to multi-variate considerations in this proposal.

The conditional dependencies in the hierarchical formulation facilitates the transfers of distributional information; from one parameter to another, and/or one time to another, and/or from one spatial region to another, in what is called “borrowing strength”. It also allows parameters whose prior distributions are more precisely known to influence those that are relatively unknown. The extent to which this occurs for say, a given parameter distribution, is quantifiable by comparing uncertainty in the posterior distribution for the parameter with the prior distribution (recall Fig. 4).

At the end of year one, we will have exercised the 1D Summary BHM to:

- a) identify comparable ecosystem component distributions (i.e. N,P,Z) across regions;
- b) compare ecosystem component distributions within regions (i.e. across physical oceanographic regimes);
- c) quantify uncertainty and identifiability in ecosystem model parameters;
- d) quantify the impacts of ROMS-NEMURO/ROMS-NPZD simulation output on uncertainty in model state variables and parameters; and
- e) propose the optimal complexity 1D probabilistic coastal ocean ecosystem model given US GLOBEC and PICES datasets, with and without ROMS-NEMURO/ROMS-NPZD output.

#### **4b. Summary 1D BHMs for Climate Change Response**

The 1-D BHM will be adapted to permit long-timescale climate integrations. Key stations bounding the North Pacific will be identified from the many stations explored in the summary BHM.

Summaries of the physical environment will be derived from a 50+ year integration of the NCAR Community Climate System Model (CCSM) for the North Pacific (see letter of support, Dr. Wm. Large). The NCAR OGCM hindcast calculation spans the period 1958-2004 in the  $\times 1^\circ$  configuration. Curchitser et al. (2005) describe the ROMS coupling to this OGCM integration. Current research at NCAR is directed toward coupling ROMS to OGCM integrations such that simulations of upwelling centers and coastal eastern boundary processes are better resolved. In the WPAC, we will explore direct driving of 1-D NEMURO models by the OGCM output (i.e. similar to Yoshie et al., 2007). In the CCS and CGOA we will explore using the physical model inputs of Curchitser et al. (2005) and/or Di Lorenzo et al. (2008). In the Di Lorenzo et al. case, we have the benefit of simulation inputs from the coupled NPZD model as well.

The focus of the 1D climate BHM will be ecosystem changes, and associated uncertainties, in responses to climate regime shifts and inter-decadal variability. A relatively large amplitude climate regime shift is associated with the late 1970’s in the North Pacific (i.e. a change in phase of the so-called Pacific Decadal Oscillation; Mantua et al., 1997). At the end of year 2 we will have exercised the 1-D climate BHM to:

---

<sup>4</sup>Henceforth we use the imperfect terminology “parameter identifiability” as a shorthand for the development of a central tendency in the posterior, from an initial vague distribution for a given parameter.

- a) compare ecosystem component sensitivities to climate regime shifts across North Pacific regions;
- b) compare ecosystem component variations in response to inter-decadal climate variability;
- c) quantify uncertainty and identifiability in ecosystem model parameters for climate scale calculations; and
- d) quantify the impacts of OGCM and ROMS-NPZD hindcast simulation output on identifiability and uncertainty in model state variables and parameters for climate scale calculations.

#### **4c. Domain-Scale 3D BHM**

ROMS-NEMURO regional simulation output will be critical in extending the 1D summary BHM to 3D. Examples of 3D-NEMURO interfaced with ocean circulation models include Aita et al. (2007) and Hashioka and Yamanaka (2007) for OGCMs, and Komatsu et al. (2007) and Fiechter et al. (2008b) for high-resolution regional models.

Our WPAC ROMS-NEMURO implementation will encompass the domain shown in Fig. 1, including the A7 and KNOT stations, and incorporating the Kuroshio and equatorward branch of the Oyashio. The model resolution will be  $1/12^\circ$ , with 40 unevenly-spaced terrain-following levels in the vertical. Seasonally varying open boundary conditions will be provided by the  $0.4^\circ$  ROMS North Pacific simulation of Curchitser et al. (2005), and the model will be forced at the surface by the NCEP-CORE products of Large and Yaeger (2004). In this way the WPAC ROMS simulations will be compatible with the model providing the boundary solution, which was also driven with NCEP-CORE derived forcing.

All of the ROMS-NPZD and ROMS-NEMURO model integrations required in this proposal will be run on computer clusters at UC Berkeley, UC Santa Cruz, and at NSF supported supercomputer centers. Separate proposals will be submitted for the latter.

The ROMS-NEMURO implementations in CCS, CGOA and WPAC, will resolve a range of regional eddies, coastal filaments, and velocity structures in boundary current regimes, as well as the biological responses to these physical environmental inputs (e.g. see Fig. 2). The full physical and biological impacts and associated uncertainties will be realizable in the 3D regional BHM, rather than the spatial summaries of impacts and uncertainties as required in the 1D summary BHM. This benefit comes at the cost of a much larger dimension problem, with associated computational overhead (i.e. in the MCMC). Dimension reduction strategies, that make physical-biological sense, will be explored from the beginning of our research effort. The larger dimension system will also require greater quantities of input data to more precisely determine modes and spreads in the posterior distribution. We will explore the utility of satellite data (e.g. CZCS, SeaWIFS) for this purpose (e.g. see Powell et al., 2006).

The 3D regional BHM will provide a probabilistic coastal ocean ecosystem research and management tool. For example, critical posterior distributions for ecosystem state variables and parameters can be used to guide future observations on timescales appropriate for ecosystem dynamics.

Research questions addressable given implementation of the 3D regional BHM include:

- a) What are the important spatio-temporal linkages between physical and biological processes at ocean synoptic-scale resolution in coastal ocean ecosystems of the N. Pacific?
- b) What are the uncertainties associated with the important ecosystem processes, region-by-region, across the N. Pacific?
- c) What is the relative importance of uncertainty in the physical circulation vs. uncertainty in parameters of the biological environment?

#### **5. Management and Schedule**

Our plans to develop BHMs to quantify pan-regional syntheses and climate change responses in North Pacific coastal ocean ecosystems require an interdisciplinary research team, and careful coordination.

Dr. Ralph F. Milliff will coordinate the project. The project management model follows that of an ONR-sponsored, international, interdisciplinary project entitled, “Bayesian Hierarchical Models to Augment the Mediterranean Forecast System”. Dr. Milliff coordinates this ongoing project as well, and participants include Profs. Wikle, Berliner (see below) and Moore. Prof. Christopher Wikle will lead the BHM development. Prof. Mevin Hooten (consultant) and Milliff will assist in this effort. Profs. Andrew M. Moore and Thomas (Zack) Powell, and Dr. Jerome Fiechter (Co-PI) will oversee ROMS-NEMURO implementation in the WPAC. Profs. Powell and Moore, and Dr. Fiechter are active in US GLOBEC efforts in the CCS and CGOA, and familiar with PICES programs in the coastal WPAC. Milliff, Fiechter, Powell and Moore will collaborate in summarizing field and model output data for the 1-D and 3-D BHM developments. Students in Statistics (U. Missouri) and Biological Oceanography (U.C. Berkeley) will be involved in the project in all three years.

All PIs and collaborators will convene in person every summer in Boulder, CO at “All-Hands Meetings” at the NWRA/CoRA offices. In addition to the project PIs and support personnel, these meetings will also include a multi-disciplinary advisory council to serve as cognizant reviewers while our research and development is in progress. Members of the advisory council include Prof. L. Mark Berliner (Statistics, Ohio State Univ.), Prof. Emanuele Di Lorenzo (Atmosphere-Ocean Sciences, Georgia Institute of Technology), Dr. William G. Large (Climate and Global Dynamics, National Center for Atmospheric Research), and Dr. Bernard A. Megrey (National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center). Advisory council members have provided letters of support for this proposal (supplementary documents) indicating their willingness to contribute in this capacity.

Prof. Hooten and the Advisory Council are listed as consultants in the NWRA/CoRA budget. NWRA policies minimize overhead charged on consulting resources to just 6%. Prof. Hooten will be compensated for his effort in BHM model development. The advisory council consulting fees are just sufficient to cover travel to Boulder for the annual All-Hands Meetings. Dr. Large is located in Boulder, his consulting fee covers meals and incidental *per diem* expenses for the annual meetings.

A project schedule is outlined in Table 1.

## **6. Results from Prior NSF Support**

Results from Prior NSF Support are listed for the 4 PIs (Milliff, Moore, Powell and Wikle) in this subsection. Papers stemming from this research (many of which have already been cited) are denoted by † in the References section of this proposal.

### **6a. Ralph F. Milliff**

*Rotationally Constrained Convections; Investigations of a New Class of Reduced Equations, 2/1/02 - 1/31/05, (NSF OCE-01-137166); Milliff Co-I.*

Systematic small parameter expansions derived for the equations of motion including: the horizontal component of rotation, relaxation to quasi-hydrostatics, and three aspect ratio regimes (Julien et al., 2006).

*SGER; Quantifying Upper Ocean Response to an Ensemble of Surface Wind Field Forcing during a Madden-Julian Oscillation, 11/15/03 - 4/30/04, (NSF ATM-035467); Milliff PI.*

Tropical surface winds from a Bayesian Hierarchical Model provide a feasible and appropriate means of generating OGCM simulation ensemble members.

### **6b. Andrew M. Moore**

*US-GLOBEC NEP Phase IIIb-CGOA: Synthesis of Biophysical Observations at Multiple Trophic Levels Using Spatially Nested, Data-assimilating Models of the Coastal Gulf of Alaska, 5/15/06 - 4/30/09 (NSF OCE-0624776); Powell and Moore Co-PIs.*

Develop a state-of-the-art coupled physical-biological model of the CGOA based on ROMS (Fiechter et al, 2008a,b).

*On the Role of Stochastic Wind Forcing in Shaping the Ocean Circulation, 8/15/00 - 8/14/05, (NSF OCE-0002370); Moore PI.*

Generalized stability analysis used to explore the impact of stochastic wind stress forcing on the subtropical gyre circulation of the North Atlantic, both at large scales and eddy scales (Chhak et al., 2006a,b, 2008, Chhak and Moore, 2007; Chhak, 2006; Weiss, 2003; Moore et al., 2002).

**6c. Thomas M. (Zack) Powell**

*GLOBEC Collaborative Research: Effects of Seasonal and Interannual Variability of Zooplankton Populations in the California Current System Using Coupled Biophysical Models, 07/01/2000 - 06/30/2006 (NSF OCE-0002893)*

A four-component trophic model is embedded within the nested system of circulation models (ROMS) for the NEP and CCS (Batchelder et al., 2002; Edwards et al. 2002a,b; Curchitser et al., 2005).

*Collaborative Research: US-GLOBEC NEP Phase IIIa-CCS: Effects of Meso- and Basin-scale Variability on Zooplankton Populations in the CCS using Data-Assimilative, Physical/Ecosystem Models, 04/15/2005 - 03/31/2008 (NSF OCE-0435574)*

Statistical estimates of the characteristic spatial and temporal scales of variability in the CCS as calculated from the nested model, agree well with those previously estimated from satellite images, for both surface temperature and pigment (Powell et al., 2006; Fiechter et al., 2008a).

**6d. Christopher K. Wikle**

*NSF MSPA-CSE: Statistical Methods for Precipitation Nowcasting and Verification, 10/1/2004 - 9/30/2008 (NSF ATM-0434213); Wikle PI.*

Development of spatio-temporal dynamical models for use in nowcasting weather radar reflectivities in a distributional context (Xu et al., 2005; Fox and Wikle 2005a,b; Song et al., 2007; Micheas et al., 2007a,b).

*NSF FRG: Statistical Analysis of Uncertainty in Climate Change, 8/1/2002 - 7/31/2006 (NSF DMS 0139903).*

Uncertainty in regional impacts of climate variability on meteorological and ecological processes (Hooten et al., 2003, 2006; Anderson et al., 2007; Berliner and Wikle, 2007; Hooten and Wikle, 2007a,b; Wikle and Berliner, 2005, 2007; Wikle and Royle, 2005; Royle and Wikle, 2005; Wikle 2003a,b; Wikle and Anderson, 2003).

*NSF CMG: Collaborative Research: Ocean Circulation Climatology and Dynamics Using Bayesian Hierarchical Methods, National Science Foundation, 9/1/2002 - 8/31/2006 (NSF ATM0222057).*

Development of efficient spatio-temporal dynamical models for eventual use with oceanic processes (Arab and Wikle, 2007a,b; Arab et al., 2007; Berliner and Wikle 2007; Wikle and Berliner, 2005, 2007; Xu and Wikle, 2007; Song et al., 2007; Cripps et al., 2005; Royle and Wikle, 2005; Xu et al., 2005; Hoar et al., 2003; Wikle 2002, 2003).

*NSF Conference on New Development of Statistical Analysis in Wildlife, Fisheries, and Ecological Research, 9/01/04 - 8/31/05 (NSF DMS 0439641)*

Workshop on modern statistical methods in Wildlife, Fisheries and Ecology that was held at the University of Missouri-Columbia in the fall of 2005.

## Schedule

Year 1	<i>1D Summary BHM Development and Calculations</i>
	<ul style="list-style-type: none"> <li>• derive and code 1D coupled bio-physical BHM process model</li> <li>• begin ROMS-NEMURO deterministic model implementation in WPAC</li> <li>• summarize regional observational datasets to be consistent with process model components (<i>in situ</i>, satellite)</li> <li>• summarize ROMS-NEMURO model output to be consistent with 1D Summary BHM</li> <li>• derive and code data stage expressions for 1D summary model</li> <li>• test 1D Summary BHM (probably in CCS first)</li> <li>• implement 1D Summary BHM in CCS, CGOA, WPAC</li> <li>• identify comparable distributions across regions</li> <li>• quantify identifiability of components and parameters</li> <li>• recode, recompare as necessary</li> <li>• implement 1D Summary BHM in different regimes within regions</li> <li>• compare distributions (components, parameters) as functions of regime</li> </ul>
Year 2	<i>1D Climate BHM Development and Calculations</i>
	<ul style="list-style-type: none"> <li>• modify coding of 1D Summary BHM to support climate timescale experiments</li> <li>• spin-up and climatology calculations in WPAC ROMS-NEMURO</li> <li>• summarize climate scale observational datasets (e.g. CalCOFI) to be consistent with 1D Climate BHM process model</li> <li>• summarize CCSM N. Pacific climate calculation to be consistent with 1D Climate BHM process model</li> <li>• derive and code data stage distributions for 1D Climate BHM</li> <li>• test 1D Climate BHM (probably in CCS first given CalCOFI)</li> <li>• implement 1D Climate BHM in CCS, CGOA, WPAC</li> <li>• compare distributions across regions</li> <li>• quantify identifiability of parameters, and compare these results with 1D Summary BHM results</li> <li>• implement 1D Climate BHM in different regimes within (at least one) region</li> <li>• compare distributions (components, parameters) as functions of regime</li> </ul>
Year 3	<i>3D Domain Scale BHM Development and Calculations</i>
	<ul style="list-style-type: none"> <li>• extend 1D BHM coding to 3D BHM (design and coding to begin in years 1 and 2)</li> <li>• modify observational data summaries to be consistent with BHM process model components in 3D</li> <li>• modify ROMS-NEMURO model output summaries to be consistent with BHM process model components in 3D</li> <li>• derive and code data stage distributions for 3D Domain Scale BHM</li> <li>• test 3D Domain Scale BHM (probably in CCS)</li> <li>• implement 3D Domain Scale BHM in at least one N. Pacific region</li> <li>• compare distributions within region</li> </ul>

## REFERENCES

- Aita, M.N., Y. Yamanaka and M.J. Kishi, 2007: "Interdecadal variation of the lower trophic ecosystem in the northern Pacific between 1948 and 2002, in a 3-D implementation of the NEMURO model", *Ecological Modelling*, **202**, 81-97.
- † Anderson, C.J., C.K. Wikle, Q. Zhou and J.A. Royle, 2007: "Population Influences on Tornado Reports in the United States", *Weather and Forecasting*, **22**, 571-579.
- † Arab, A. and C.K. Wikle, 2008: "Galerkin approaches for efficiently parameterizing spatio-temporal dynamical processes", In review.
- † Arab, A. and C.K. Wikle, 2008: "A hierarchical Bayesian finite-element approach for modeling tidal currents", In review.
- † Arab, A., M.B. Hooten and C.K. Wikle, 2007: "Hierarchical Spatial Models", in: *Encyclopedia of Geographical Information Science*, Springer, In Press.
- † Batchelder, H.P., C.A. Edwards, and T.M. Powell, 2002: "Individual-based models of zooplankton populations in coastal upwelling regions: Implications of diel vertical migration on demographic success and nearshore retention", *Progress in Oceanography*, **53**, 307-333.
- Batchelder, H.P. and M. Kashiwai, 2007: "Ecosystem modeling with NEMURO within the PICES Climate Change Carrying Capacity program", *Ecological Modelling*, **202**, 7-11.
- Berliner, L.M., 1996: "Hierarchical Bayesian time series models", In: Hanson, K. and Silver, R. (eds), *Maximum entropy and Bayesian Methods*, Kluwer Academic Publishers, pp 1522.
- Berliner, L.M., C.K. Wikle and N. Cressie, 2000: "Long-lead prediction of Pacific SSTs via Bayesian Dynamic Modeling", *Journal of Climate*, **13**, 3953-3968.
- Berliner, L.M., R.F. Milliff and C.K. Wikle, 2003: "Bayesian hierarchical modeling of air-sea interaction", *J. Geophys. Res., Oceans*, **108**(C4).
- † Berliner, L.M. and C.K. Wikle, 2007: "Approximate importance sampling Monte Carlo for data assimilation", *Physica D*, **230**, 37-49.
- Buckland, S.T., Newman, K.B., Thomas, L. and N.B. Koesters, 2004: "State-space models for the dynamics of wild animal populations", *Ecological Modelling*, **171**, 157-175.
- Calder, C., M. Lavine, P. Muller and J.S. Clark, 2003: "Incorporating multiple sources of stochasticity into dynamic population models", *Ecology*, **84**(6), 1395-1402.
- Caswell, H., 2001: *Matrix Population Models*. Sinauer Associates, Inc., Sunderland, MA.
- † Chhak, K.C., 2006: *Stochastic Forcing of the Wind-Driven Ocean Circulation*, PhD. Thesis, University of Colorado at Boulder, 254pp.
- † Chhak, K.C., A.M. Moore and R.F. Milliff, 2008: "Stochastically forcing of ocean variability by the North Atlantic Oscillation. *J. Phys. Oceanogr.*, submitted.
- † Chhak, K.C. and A.M. Moore, 2007: "The North Atlantic Oscillation as a source of stochastic forcing of the wind driven ocean circulation", *Dyn. Atmos. Oceans*, **43**, 151-170.
- † Chhak, K.C., A.M. Moore, R.F. Milliff, G. Branstator, W.R. Holland and M. Fisher, 2006: "Stochastic forcing of the North Atlantic wind-driven ocean circulation. Part I: A Diagnostic Analysis of the Ocean Response to Stochastic Forcing", *J. Phys. Oceanogr.*, **36**, 300-315.
- † Chhak, K.C., A.M. Moore, R.F. Milliff, G. Branstator, W.R. Holland and M. Fisher, 2006: "Stochastic forcing of the North Atlantic wind-driven ocean circulation. Part II: An Analysis of the Dynamical Ocean Response using Generalized Stability Theory", *J. Phys. Oceanogr.*, **36**, 316-334.
- Clark, J.S., S.R. Carpenter, M. Barber, S. Collins, A. Dodson, J. Foley, D. Lodge, M. Pascual, R. Pielke, Jr., W. Pizer, C. Pringle, W.V. Reid, K.A. Rose, O. Sala, W.H. Schlesinger, D. Wall and D. Wear 2001: "Ecological forecasts: An emerging imperative", *Science*, **293**, 657-660.

- Clark, J.S., 2003: “Uncertainty and variability in demography and population growth: A hierarchical approach”, *Ecology*, **84**(6), 1370-1381.
- Clark, J.S. and A. E. Gelfand (eds.), 2006: *Hierarchical Modelling for the Environmental Sciences, Statistical Methods and Applications*, Oxford University Press.
- Clark, J.S., 2007: *Models for Ecological Data, An Introduction*, Princeton University Press.
- † Cripps, E., D. Nott, W.T.M. Dunsmuir and C.K. Wikle, 2005: “Space-time Modelling of Sydney Harbour Winds”, *Australian and New Zealand Journal of Statistics*, **47**, 3-17.
- † Curchitser, E.N., D.B. Haidvogel, A.J. Hermann, E.L. Dobbins, T.M. Powell and A. Kaplan, 2005: “Multi-scale modeling on the North Pacific Ocean: Assessment and analysis of simulated basin-scale variability (1996-2003)”, *J. Geophysical Research, Oceans*, **110**, C11021.
- Denman, K.L. and M.A. Pena, 1999: “A coupled 1-D biological/physical model of the northeastern subarctic Pacific Ocean with iron limitation”, *Deep Sea Research II*, **46**, 2877-2908.
- Denman, K.L. and M.A. Pena, 2002: “The response of two coupled one-dimensional mixed layer/planktonic ecosystem models to climate change in the NE subarctic Pacific Ocean”, *Deep Sea Research II*, **49**, 5739-5757.
- Di Lorenzo, E.N. Schneider, P.J.S. Franks, S.J. Bograd, A.J. Miller, P. Riviere, H.G. Arango, E. Curchister, K.M. Cobb, T.M. Powell and K.C. Chhak, 2007: “North Pacific gyre-scale oscillation controls decadal variations of salinity, nutrients and chlorophyll-a”, *Geophys. Res. Letts.*, Submitted.
- † Edwards, C.A., T.M. Powell, and H.P. Batchelder, 2000a: “The stability of an NPZ model subject to realistic levels of vertical mixing”, *Journal of Marine Research*, **58**, 37-60.
- † Edwards, C.A., H.P. Batchelder, and T.M. Powell, 2000b: “Modeling microzooplankton and macrozooplankton dynamics within a coastal upwelling system”, *Journal of Plankton Research*, **22**(9), 1619-1648.
- Eknes, M. and G. Evensen, 2002: “An ensemble Kalman filter with a 1-D marine ecosystem model”, *Journal of Marine Systems*, **36**, 75-100.
- Fennel, K., J. Wilkin, J. Levin, J. Moisan, J. O’Reilly and D. Haidvogel, 2006: “Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget”, *Global Biogeochemical Cycles*, **20**, GB3007, doi:10.1029/2005GB002456.
- † Fiechter, J., C.A. Edwards, K.W. Bruland, A.M. Moore, E. Di Lorenzo, E.N. Curchitser, T.M. Powell and K. Hedstrom, 2008a: “A simple approach to model iron limitation on primary production in the Coastal Gulf of Alaska”, *Deep Sea Research*, Submitted.
- † Fiechter, J., A.M. Moore and H.G. Arango, 2008b: “Characterization of multi-nutrient limitation on phytoplankton growth in the coastal Gulf of Alaska”, In preparation.
- † Fox, N.I. and C.K. Wikle, 2005: “A Bayesian quantitative precipitation nowcast scheme”, *Weather and Forecasting*, **2**, 264-275.
- † Fox, N.I. and C.K. Wikle, 2005: “Providing distributed forecasts of precipitation using a Bayesian nowcast scheme”, *Atmospheric Science Letters*, **6**, 59-65.
- Franks, P.J., J.S. Wroblewski and G.R. Flierl, 1986: “Behavior of a simple plankton model with food-level acclimation by herbivores”, *Marine Biology*, **91**, 121-129.
- Fujii, M., Y. Yamanaka, Y. Nojiri, M.J. Kishi and F. Chai, 2007: “Comparison of seasonal characteristics in biogeochemistry among the subarctic North Pacific stations described with a NEMURO-based marine ecosystem model”, *Ecological Modelling*, **202**, 52-67.
- Gelfand, A.E., J.A. Silander, S. Wu, A. Latimer, P.O. Lewis, A.G. Rebelo and M. Holder, 2005: “Explaining species distribution patterns through hierarchical modeling”, *Bayesian Analysis*, **1**, 41-92.
- Haidvogel, D.B., H.G. Arango, K. Hedstrom, A. Beckmann, P. Malanotte-Rizzoli and A.F. Shchepetkin, 2000: “Model evaluation experiments in the North Atlantic Basin: Simulations in nonlinear terrain-following coordinates”, *Dynamics Atmosphere Oceans*, **32**, 239-281.

- Hashioka, T. and Y. Yamanaka, 2007a “Seasonal and regional variations of phytoplankton groups by top-down and bottom-up controls obtained by a 3-D ecosystem model”, *Ecological Modelling*, **202**, 68-80.
- Hashioka, T. and Y. Yamanaka, 2007b “Ecosystem change in the western North Pacific associated with global warming using 3D-NEMURO”, *Ecological Modelling*, **202**, 95-104.
- Hinckley, S., K.O. Coyle, G. Gibson, A.J. Hermann and E.L. Dobbins, 2008: “A biophysical NPZ model with iron for the Gulf of Alaska: Reproducing the differences between an oceanic HNLC ecosystem and a classic northern temperate shelf ecosystem”, *Deep Sea Research II*, submitted.
- † Hoar, T.J., R.F. Milliff, D. Nychka, C.K. Wikle and L.M. Berliner, 2003: “Winds from a Bayesian hierarchical model: Computation for atmosphere-ocean research”, *Journal of Computational and Graphical Statistics*, **4**, 781-807.
- † Hooten, M. B., D.R. Larsen and C.K. Wikle, 2003: “Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model”, *Landscape Ecology*, **18**, 487-502.
- † Hooten, M.B. and C.K. Wikle, 2007: “Shifts in the spatio-temporal growth dynamics of shortleaf pine”, *Environmental and Ecological Statistics*, **14**, 207-227.
- † Hooten, M. B. and C.K. Wikle, 2007: “A Hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian Collared-Dove”, *Environmental and Ecological Statistics*, DOI: 10.1007/s10651-007-0040-1.
- † Hooten, M.B., C.K. Wikle, R.M. Dorazio and J.A. Royle, 2007: “Hierarchical spatio-temporal matrix models for characterizing invasions”, *Biometrics*, **63**, 558-567. (Web Appendix)
- † Julien, K., E. Knobloch, R.F. Milliff and J.A. Werne, 2006: “Generalized quasi-geostrophy for spatially anisotropic rotationally constrained flows”, *J. Fluid. Mech.*, **555**, 233-274.
- Kishi, M., M. Kashiwai, D.M. Ware, B.A. Megrey, D.L. Eslinger, F.E. Werner, M.N. Aita, T. Azumaya, M. Fujii, S. Hashimoto, D. Huang, H. Iizumi, Y. Ishida, S. Kang, G.A. Kantakov, H-C. Kim, K. Komatsu, V.V. Navrotsky, S. Lan Smith, K. Tadokoro, A. Tsuda, O. Yamamura, Y. Yamanaka, K. Yokouchi, N. Yoshie, J. Zhang, Y.I. Zuenko, V.I. Zvalinsky, 2007: “NEMURO- a lower trophic level model for the North Pacific marine ecosystem”, *Ecological Modelling*, **202**, 12-25.
- Komatsu, K., Y. Matsukawa, K. Nakata, T. Ichikawa and K. Sasaki, 2007: “Effects of advective processes on planktonic distributions in the Kuroshio region using a 3-D lower trophic model and a data assimilative OGCM”, *Ecological Modelling*, **202**, 105-119.
- Large, W.G. and S.G. Yeager, 2004: “Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies”, *NCAR Technical Note 460*.
- Losa, S.N., G.A. Kivman, J. Schröter and M. Wenzel, 2003: “Sequential weak constraint parameter estimation in an ecosystem model”, *Journal of Marine Systems*, **43**, 31-49.
- Mantua, N., S. Hare, Y. Zhang, J. Wallace and R. Francis, 1997: “A Pacific interdecadal climate oscillation with impacts on salmon production”, *Bull. Amer. Meteorol. Soc.*, **78**, 1069-1079.
- † Micheas, A.C., N.I. Fox, S.A. Lack and C.K. Wikle, 2007: “Cell identification and verification of QPF ensembles using shape analysis techniques”, *Journal of Hydrology*, doi:10.1016/j.jhydrol.2007.05.036.
- † Micheas, A.C. and C.K. Wikle, 2008: “A Bayesian hierarchical non-overlapping random disc growth model”, submitted
- † Moore, A.M., C.L. Perez, and J. Zavala-Garay, 2002: “A non-normal view of the wind-driven ocean circulation”, *J. Phys. Oceanogr.*, **32**, 2681-2705.
- Moore, A.M., H.G. Arango, E. Di Lorenzo, A.J. Miller and B.D. Cornuelle, 2007: “An adjoint sensitivity analysis of the southern California Current circulation and ecosystem. Part I: The physical environment”, *J. Phys. Oceanogr.*, Under review.
- Moore, A.M., H.G. Arango, E. Di Lorenzo, C.V.W. Lewis, T.M. Powell, B.D. Cornuelle and A.J. Miller, 2008: “An adjoint sensitivity analysis of the southern California Current circulation and ecosystem. Part

- II: The biological environment”, *J. Phys. Oceanogr.*, Under review.
- † Powell, T.M., C.V.W. Lewis, E.N. Curchitser, D.B. Haidvogel, A.J. Hermann and E.L. Dobbins, 2006: “Results from a three-dimensional, nested, biological-physical model of the California Current System and comparisons with statistics from satellite imagery”, *Journal of Geophysical Research*, **111**(C0), 7018, doi:10.1029/2004JC002506.
- Robert, C.P. and Casella, G., 2004: *Monte Carlo Statistical Methods: Second Edition*, Springer.
- Royle, J.A., L.M. Berliner, C.K. Wikle and R.F. Milliff, 1998: ”A hierarchical spatial model for constructing wind fields from scatterometer data in the Labrador Sea.” in *Case Studies in Bayesian Statistics IV*, C. Gatsonis, R.E. Kass, B. Carlin, A. Cariquiry, A. Gelman, I. Verdinelli, and M. West (Eds.), Springer-Verlag, 367-381.
- † Royle, J.A. and C.K. Wikle, 2005: “Efficient statistical mapping of avian count data”, *Environmental and Ecological Statistics*, **12**, 225-243.
- Shchepetkin, A.F., and J.C. McWilliams, 2005: “The regional oceanic modeling system (ROMS): A split-explicit, free-surface, topography-following-coordinate oceanic model”, *Ocean Modelling*, **9**, 347-404.
- † Song, Y., C.K. Wikle, C.J. Anderson and S.A. Lack, 2007: “Bayesian estimation of stochastic parameterizations in a numerical weather forecasting model”, *Monthly Weather Review*, **135**, 4045-4059.
- Spitz, Y.H., P.A. Newberger and J.S. Allen, 2003: “Ecosystem response to upwelling off the Oregon coast: Behavior of three nitrogen-based models”, *Journal of Geophysical Research*, **108**(C3), 3062, doi:10.1029/2001JC001181.
- Su, Z., R.M. Peterman and S.L. Haeseker, 2004: “Spatial hierarchical Bayesian models for stock-recruitment analysis of pink salmon (*Oncorhynchus gorbuscha*)”, *Canadian Journal of Fisheries and Aquatic Sciences*, **61**, 2471-2486.
- Thomas, L., S.T. Buckland, K.B. Newman and J. Harwood, 2005: “A unified frame-work for modelling wildlife population dynamics”, *Australian and New Zealand Journal of Statistics*, **47**, 19-34.
- Wainwright, T.C., L.R. Feinberg, R.C. Hooff and W.T. Peterson, 2007: “A comparison of two lower trophic models for the California Current System”, *Ecological Modelling*, **202**, 120-131.
- † Weiss J.B., 2003: “Coordinate invariance in stochastic dynamical systems”, *Tellus*, **55A**, 208-218.
- Werner, F.E., S-I. Ito, B.A. Megrey and M.J. Kishi, 2007: “Synthesis of the NEMURO model studies and future directions of marine ecosystems modeling”, *Ecological Modelling*, **202**, 211-223.
- Wikle, C.K., R.F. Milliff, D. Nychka, L.M. Berliner, 2001: “Spatio-Temporal Hierarchical Bayesian Modeling: Tropical Ocean Surface Winds”, *J. Amer. Stat. Assoc.*, **96**, 383-397.
- † Wikle, C.K., 2002: “A kernel-based spectral model for non-Gaussian spatio-temporal processes”, *Statistical Modelling: An International Journal*, **2**, 299-314.
- Wikle, C.K., 2003a: “Hierarchical models in environmental science”, *International Statistical Review*, **71**, 181-199.
- † Wikle, C.K., 2003b: “Hierarchical Bayesian models for predicting the spread of ecological processes”, *Ecology*, **84**, 1382-1394.
- † Wikle, C.K., 2003c: “Modern statistical methods and weather modification research. Appendix in *Critical Issues in Weather Modification Research*, A National Research Council (NRC) report. The National Academies Press, Washington, D.C., 107-113.
- † Wikle, C.K., 2003d: “Bayesian statistics in the atmospheric sciences”, *ISBA Bulletin*, **10**, 5-6.
- † Wikle, C.K. and C.J. Anderson, 2003: “Climatological analysis of tornado report counts using a hierarchical Bayesian spatio-temporal model”, *Journal of Geophysical Research - Atmospheres*, **108**(D24), 9005, doi:10.1029/2002JD002806.
- Wikle, C.K., L.M. Berliner and R.F. Milliff, 2003: “Hierarchical Bayesian approach to boundary value problems with stochastic boundary conditions”, *Monthly Weather Review*, **131**, 1051-1062.

- † Wikle, C.K. and L.M. Berliner, 2005: “Combining information across spatial scales”, *Technometrics*, **47**, 80-91.
- † Wikle, C.K. and J.A. Royle, 2005: “Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data”, *Environmetrics*, **16**, 507-522.
- Wikle, C.K. and M.B. Hooten, 2006: “Hierarchical Bayesian Spatio-Temporal Models for Population Spread”, Clark, J.S. and A. Gelfand (eds). In: *Applications of Computational Statistics in the Environmental Sciences: Hierarchical Bayes and MCMC Methods*. Oxford University Press, 145-169.
- † Wikle, C.K. and L. M. Berliner, 2007: “A Bayesian tutorial for data assimilation”, *Physica D*, **230**, 1-16.
- Wroblewski, J.S., 1977: “A model of phytoplankton plume formation during variable Oregon upwelling”, *Journal of Marine Research*, **35**, 357-394.
- † Xu, K., C.K. Wikle and N.I. Fox, 2005: “A kernel-based spatio-temporal dynamical model for nowcasting radar precipitation”, *Journal of the American Statistical Association*, **100**, 1133-1144.
- † Xu, K. and C.K. Wikle, 2007: “E-M algorithm implementation of efficient spatio-temporal dynamical models”, *Journal of Statistical Planning and Inference*, **137**, 567-588.
- Yoshie, N., Y. Yamanaka, K.A. Rose, D.L. Eslinger, D.M. Ware and M.J. Kishi, 2007: “Parameter sensitivity study of the NEMURO lower trophic level marine ecosystem model”, *Ecological Modelling*, **202**, 26-37.