A Bayesian Approach to the Ecosystem Inverse Problem with Application to a Shellfish Growth Model

Michael Dowd *

Ocean Sciences Division, Fisheries and Oceans Canada

Biological Station, St. Andrews, N.B., Canada

Renate Meyer

Department of Statistics, University of Auckland

Auckland, New Zealand

KEYWORDS: Bayesian statistics, inverse methods, marine ecosystem model, shellfish growth, data assimilation

* *Corresponding Author*: Michael Dowd, Fisheries & Oceans Canada, Biological Station, St. Andrews, N.B., CANADA, E5B 2L9. *Phone*: 506-529-5848. *Fax*: 506-529-5847. *email*: dowdm@mar.dfo-mpo.gc.ca

Abstract

This study investigates a probabilistic approach for the inverse problem associated with blending time-dependent dynamic models of marine ecosystems with observations. The goal is to combine prior information, in the form of model dynamics and substantive knowledge about uncertain parameters, with available measurements in order to produce posterior estimates of the time-varying ecological state variables, along with their uncertainty. Ecological models of interacting populations are considered in the context of the nonlinear, non-Gaussian state space model. This comprises a nonlinear stochastic difference equation for the model dynamics, and an observation equation relating the model state to the measurements. Complex error processes are readily incorporated. The posterior probability density function provides a complete solution to the inverse problem. Bayes' theorem allows one to obtain this posterior density through synthesis of the prior information and the observations. To illustrate this Bayesian inverse method, these ideas are applied to a simple ecosystem box model concerned with predicting the seasonal co-evolution of a population of grazing shellfish and its two food sources, plankton and detritus. Observations of shellfish growth over time are available. Lognormal system noise was incorporated into the ecosystem equations at all time steps. Ingestion and respiration parameters for shellfish growth are considered as uncertain quantities described by beta distributions. Stochastic simulation was carried out and provided predictions of the model state with uncertainty estimates. The Bayesian inverse method was then used to assimilate the additional information contained in the observations. Posterior probability density functions for the parameters and time-varying state were computed using Markov Chain Monte Carlo methods. The ecological dynamics spread the measurement information to all state variables, even those not directly observed. Probabilistic state estimates are refined in comparison to those from the stochastic simulation. It is concluded that this Bayesian approach appears promising as a framework for ecosystem inverse problems, but requires careful control of the dimensionality for practical applications.

1 Introduction

Inverse methods, or data assimilation, are concerned with combining dynamical models and measurements in order to produce optimal estimates for the state of the system. They have a well established role in fields such as geophysics and physical oceanography (Tarantola 1987, Bennett 1992, Wunsch 1996). Marine ecological modelling and its integration with ocean circulation models is advancing rapidly (Hofmann and Lascara 1998). These developments are driven, in part, by the revolution in observing technologies. For example, the foundational elements of the aquatic food chain - the plankton - are now routinely surveyed by satellite remote sensing (Kahru and Mitchell 2001). They are also being measured using new *in situ* sensors based on acoustics (Batchelder et al. 1995), optics (Cullen et al. 1997) and image analysis (Sieracki et al. 1998). The development of methods for the fusion of these and other data sources with marine ecosystem models is underway (Lawson et al. 1995, Vallino 2000, Natvik et al. 2001). To date these studies have mainly relied on an optimization approach to data assimilation. The purpose of this study is to address the ecosystem inverse problem using a general nonlinear state space model and Bayesian statistics in order to jointly estimate the time-varying ecological state and parameters using available observations.

Probabilistic approaches provide the conceptual foundation for inverse methods (Bryson and Ho 1969, Jazwinski 1970). The building blocks are a time-dependent dynamical model and an observation equation, which together comprise the state space representation (Kitagawa and Gersch 1996). If these are assumed linear with Gaussian errors, the probabilistic approach reduces to state estimation algorithms such as the Kalman filter, and its associated fixed interval smoother (Jazwinski 1970). Ignoring model errors allows the use of efficient parameter estimation techniques based on optimization approaches, which may be viewed as nonlinear regression (Thompson et al. 2000). Unfortunately, marine ecological models often do not satisfy the criteria necessary for these simplications and the effectiveness of such state estimation techniques is unclear (Vallino 2000). Ecological systems are inherently, and often strongly, nonlinear with complex dynamical behaviour (Popova et al. 1997, Edwards 2001). There is also significant uncertainty as to the appropriate model structure (Reichart and Omlin 1997), implying that that model errors and parameterizations can require complicated stochastic descriptions. Available measurements may have distributions which are far from Gaussian (Soundant et al. 1997, Dowd et al. 2002). As a result of these factors, it has been argued that estimation of the ecological state using models and data should rely on a Bayesian statistics (Dilks et al. 1992, Aldenberg et al. 1995, Steinberg et al. 1997, Harmon and Challenor 1997). These approaches are based on a probabilistic specification of prior information pertaining to the ecological dynamics and model parameters, as well as specification of distributional assumptions of the observations. Bayesian application synthesizes this information leading to posterior probability density functions describing the state of the ecological system.

This study investigates Bayesian inverse methods for blending time dependent marine ecological models and measurements. The goal is to determine the joint probability density of the ecological state and parameters, using both prior information and observations. The approach taken here relies on a nonlinear state space framework and Bayes' theorem, providing a flexible framework for incorporating nonlinear models and measurement operators, as well as allowing for complex error processes. A simple, but not unrealistic, shellfish ecosystem model is used to illustrate how posterior state and parameter estimates can be determined from prior information and observations under a variety of distributional assumptions. The paper is structured as follows. Section 2 introduces the theory and background of our Bayesian analysis of the nonlinear, non-Gaussian system governing the ecosystem inverse problem. Section 3 applies these ideas to a simple shellfish ecosystem model, including both stochastic simulation and data assimilation. A summary and conclusions follow in Section 4.

2 Model

Models of interacting populations in marine ecology frequently take the form of a nonlinear system of coupled ordinary differential equations, i.e.

$$\dot{x} = f(x; \theta). \tag{1}$$

These include, for example, the prototypical phytoplankton - zooplankton - nutrient model (Steele and Henderson 1982, Franks et al. 1986, Edwards 2001), as well as the shellfish ecosystem models considered later in this study (Raillard and Ménesguen 1994, Dowd 1997, Chapelle et al. 2000). In the above, x is a vector that represents the state of the system and $\dot{x} = dx/dt$, where t is time. The components of x are each identified with the populations comprising the ecological system. These are generally measured in terms of the numbers, or density, of individual organisms in a finite volume of water. The function $f(x;\theta)$ describes the interactions between the populations in terms of mathematical descriptions of the flow of matter or energy. These fluxes are dependent on the population levels, x, as well as an underlying set of parameters, θ. The parameters contain information on biological rates and environmental forcing. Given initial conditions, $x(0)$, and the parameters, θ, integration of (1) provides the time sequence $x(t)$. This procedure is usually carried out numerically as the complexity and nonlinearity of most ecological models generally precludes analytic approaches. Numerical integration methods are based on the discretization of the governing equations, i.e. converting the differential equations to difference equations. This operation means that the state x is now defined at a finite number of instances in time. This discrete time state is designated x_t for $t = 0, \ldots, T$.

Inverse problems are concerned with estimating the state of a system by combining mathematical models with measurements. They may be viewed as fitting models to data (Thacker and Long 1988), or as data interpretation in the context of a dynamical model. In the absence of inverse methods, a typical model exercise for estimating the state of the ecological system takes the following form. The forward model (1) is integrated to produce predicted values for the state. These predictions are retrospectively compared to available observations. Parameter values, or the state variables themselves, are then adjusted so that the discrepancy between observations and predictions is minimized. (If this 'tuning' exercise proves unsuccessful, the model structure itself may be modified.) One view of inverse methods is as a systematic means of carrying out the above procedure. Such an approach relies on optimization methods and has provided the foundation for current applications in marine ecology (Lawson et al. 1995, Vallino 2000). An alternative view of the inverse problem is based on Bayesian perspective. That is, we seek a (posterior) description of the ecological state based on the synthesis of all available prior knowledge of the system, and information from the measurements. The perspective is one of a conjunction of different sources of information, rather than as an optimization problem. Bayesian methods in ecological modelling have considered mainly parameter estimation using data (Dilks et al. 1992, Aldenberg et al. 1995 Harmon and Challenor 1997). Here, we investigate a Bayesian approach to the ecosystem inverse problem wherein the joint posterior probability density function of the time-varying ecological state and parameters are determined under very general assumptions about the underlying form of error processes.

Prior information on the ecological system takes the form of model dynamics and substantive knowledge about parameters. Each of these are subject to some level of uncertainty and therefore naturally described in a probabilistic manner using the joint probability density function (pdf) of the state and parameters. As observations are subject to measurement error, the information stemming from the data is given in the form of a sampling distribution, or the so-called *likelihood* function. The overall goal is to update the *prior* pdf, in the light of new information obtained from the observations, to obtain a joint *posterior* pdf of the state and parameters. The only mathematically rigorous way to perform this update is through the Bayesian paradigm. It provides a way of formalising the process of learning from data so as to update prior beliefs in accord with recent notions of knowledge synthesis. The primary advantage of this approach is a complete characterization of the ecological state as a random variable, i.e. through obtaining its full probability density function. State estimates and measures of uncertainty can then be derived from these quantities. These do not depend on the assumptions of asymptotic normality underlying classical estimation methods. The full distributional profile of the state is provided, and any deviations from normality can be easily detected. The main difficulty is that for nonlinear and non-Gaussian systems, a characteristic of most ecological models, computationally intensive Monte Carlo based methods must be used. Adopting a sampling-based approach to posterior computation, however, has various advantages. In particular, it is easy to derive any *marginal* posterior pdfs, and posterior pdfs of any transformed parameter of interest.

The basic framework for our probabilistic approach to inverse modelling is the nonlinear state space representation (Kitagawa and Gersch 1996). Define $z_t = (x_t,\ \theta)'$ as the augmented state vector, which includes both the ecological state variables as well as the parameters, some of which may be time dependent (Kitagawa 1998). Available observations are given as y_t. A nonlinear state space representation of this system takes the form

$$z_t|z_{t-1} = g(z_{t-1}, n_t) \tag{2}$$

$$y_t|z_t = h(z_t, v_t) \tag{3}$$

for $t = 1, \ldots, T$. Here, the notation $A|B$ denotes the conditional distribution of A given B. The functions $g(z_{t-1}, n_t)$ and $h(z_t, v_t)$ are possibly nonlinear functions of the state and the noise inputs. The first equation represents the ecological dynamics. The function g includes the discretized dynamics of (1) modified to account for the augmented state. It is assumed that a Markovian form is sufficient, wherein the state depends only on values at the previous time step. The system noise term, n_t, represents model errors and parameter uncertainty. The observation equation is given by (3) and relates the observations to the state through the nonlinear operator h. This serves the dual purpose of interpolating the measurements to the computational grid, and converting measured quantities to modelled variables. Observation noise v_t has also been included. The state z_t and the observations y_t are random vectors and the representation (2)-(3) describes a nonlinear and non-Gaussian system. It is an extension of the well-known linear state-space model

$$x_t|x_{t-1} = Gx_{t-1} + n_t \tag{4}$$

$$y_t|x_t = Hx_{t-1} + v_t \tag{5}$$

where G and H are matrices and additive Gaussian white noise sequences $n_t \sim Normal(0, R_t)$ and $v_t \sim Normal(0, S_t)$ are used.

For notational convenience, define the following quantities: $Y = (y_1, \ldots, y_T)'$ and $Z = (z_0, \ldots, z_T)$. These contain the full state and complete observations sets for the period of interest. The quantity $P(Z|Y)$ represents the posterior joint probability density function of the state Z conditioned on the observation set Y, i.e. after taking the data into account. Knowledge of $P(Z|Y)$ provides a complete solution to the inverse problem. It embodies all statistical information about the state coming from the two sources: the observations, and the prior knowledge on the dynamics and parameters.

Further insight into the inverse problem can be gained by expanding $P(Z|Y)$ using Bayes' theorem, i.e.

$$P(Z|Y) = \frac{P(Y|Z)P(Z)}{P(Y)} \tag{6}$$

$$\propto \quad P(Y|Z)P(Z). \tag{7}$$

Now, consider each of the terms on the right-hand side of (6):

1. $P(Z)$ represents the prior pdf of the state. It is based on prior knowledge of the dynamics and parameters. The Markov property of the dynamics (2) allows us to write
$$P(Z) = P(z_0)\prod_{t=1}^{T} P(z_t|z_{t-1}). \tag{8}$$
This implies that $P(Z)$ is comprised of a product of the pdf of the initial conditions, $P(z_0)$, and all the transition pdfs, $P(z_t|z_{t-1})$. These transition densities are identified with the ecological dynamics (2). At time t, $P(z_t|z_{t-1})$ acts to transform the pdf of the ecological state at time $t-1$, $P(z_{t-1})$, to the new pdf at time t, $P(z_t)$.

2. $P(Y|Z)$ is the joint pdf of the observations, Y, conditioned on knowledge of the state Z, or the likelihood function. If the y_t in (3) are conditionally independent given the state z_t, the observation equation (3) implies
$$P(Y|Z) = \prod_{t=1}^{T} P(y_t|z_t). \tag{9}$$
This term therefore incorporates the observational errors.

3. $P(Y)$ is the marginal pdf of the observations. It acts as a normalization factor for the numerator in (6). Expanding using conditional probability and marginalization yields
$$P(Y) = \int_{-\infty}^{\infty} P(Y|Z)P(Z)dZ. \tag{10}$$
Evaluation of this term involves high dimensional integration (equal to the dimension of the state Z). In the past, this was the main difficulty that hindered the application of Bayesian inference. For modern posterior computation using simulation techniques, this normalization constant is not required.

The above decomposition (7) of the quantities in Bayes' formula (ignoring the normalization constant) shows how the prior pdf of the state is updated using the likelihood function to yield the posterior pdf of the state. The posterior pdf can thus be regarded as a synthesis of different sources of information.

If we make the assumption of a linear form for the model and observations equations, as in (2)-(3), and assume additive and independent Gaussian errors, as in (4)-(5), the problem of evaluating $P(Z|Y)$ reduces to the time dependent estimation problem of optimal smoothing (Bryson and Ho 1969, Jazwinski 1970). For this case, the pdf of the state is Gaussian and completely characterized by its mean and covariance. The solution is provided by application of the Kalman filter and associated fixed interval smoother. Unfortunately, most ecological models and measurements fall into the category of nonlinear and non-Gaussian systems, with conditional distributions that are no longer normally distributed. Various types of approximations to, or assumptions on, the pdfs are used to obtain recursive formulas for state estimation. Some of the examples are the extended Kalman filter (Jazwinski 1970), the Gaussian-sum filter (Alspach and Sorenson 1972), the dynamic generalized linear model (West et al. 1985), and the non-Gaussian filter and smoother (Kitagawa 1987, Hodges and Hale 1993). However, these extensions are not always effective (e.g. Evensen 1994). Optimization methods are similarly restricted, estimating only the mean of the state using, say, a variational framework which minimizes the weighted sum of squares of the errors (Bryson and Ho 1969, Bennett 1992). The general solution for $P(Z|Y)$ for nonlinear and non-Gaussian systems relies on Monte Carlo methods (Kitagawa 1987).

Markov Chain Monte Carlo (MCMC) methods have become a well established means for practical applications of Bayesian methods (Gelman et al. 1995). MCMC techniques are general Monte Carlo methods that approximate the generation of samples from the posterior pdf $P(Z|Y)$ when this distribution is so high-dimensional and complex that it cannot be sampled directly. The basic idea of MCMC is to generate a Markov chain with a stationary (or 'limiting' or 'long run') distribution $P(Z|Y)$. This idea of using the limiting behavior of a Markov chain came almost as early as the original Monte Carlo technique, at least in the particle physics literature (Metropolis et al. 1953), but it required computational power that was not available in those early days. Two major techniques devised to create Markov chains with the desired stationary distribution are Gibbs sampling (Geman and Geman 1984, Gelfand and Smith 1990) and the Metropolis-Hasting algorithm (Metropolis et al. 1953, Hastings 1970). The interested reader is referred to Gilks et al. (1996), Gamerman (1997) and Robert and Casella (1999) for more information on the theory and application of MCMC methods. Harmon and Challenor (1997) provide an application to ecological parameter estimation. For the user, important practical issues include ensuring convergence of the Markov chain to stationarity, and making sure that the sampled portion of the Markov chain has traversed the entire parameter space (or achieved 'mixing'). Generally, a 'burn-in' period is allowed for the system to adjust prior to drawing samples, followed by careful assessment of convergence diagnostics as suggested in Cowles and Carlin (1996). This ensures that samples drawn from the chain are truly representative of $P(Z|Y)$. Nonlinear state space models were first examined by Carlin et al. (1992) and Fruehwirth-Schnatter (1994) using the Gibbs sampler. An overview of more advanced MCMC techniques for sequential state-space modelling is given in Doucet et al. (2001).

In this study, we make use of MCMC methods as implemented in the Bayesian Inference Using Gibbs Sampling (BUGS) software (Spiegelhalter et al. 1996). BUGS is available free of charge from

http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml

for the operating systems UNIX, LINUX, and WINDOWS, among others. It comes with complete documentation and two example volumes. BUGS has been used for a Bayesian approach to nonlinear non-Gaussian state space models in various different disciplines, such as fisheries stock assessment (Meyer and Millar 1999, Millar and Meyer 2000), econometrics (Meyer and Yu, 1999), and chaotic dynamical systems in physics (Meyer and Christensen 2000). Our goal is to assess the utility of a Bayesian approach to inverse modelling in the context of a simple, but not unrealistic, coastal marine ecosystem model focused on shellfish growth.

3 Application

Shellfish Ecosystem Model

The Bayesian inverse framework is applied to a shellfish aquaculture ecosystem model. This is a simplified version of the ecosystem model in Dowd (1997) which described a coastal marine ecosystem in a tidal inlet near Lunenburg, Nova Scotia, Canada. The focus is on the prediction of mussel growth and carrying capacity on seasonal time scales in the context of its supporting ecosystem. The physical situation is one of a semi-enclosed tidal inlet exchanging water with the adjacent open ocean. The governing equations for the ecosystem box model are given by the following system of ordinary differential equations:

$$\begin{aligned} \dot{M} &= (\epsilon f_I I - f_R R)\, M && (11) \\ \dot{P} &= -\alpha_P f_I I M N + K(P_\infty - P) && (12) \\ \dot{D} &= -\alpha_D f_I I M N + K(D_\infty - D). && (13) \end{aligned}$$

Here, the biomass of an individual mussel is given by M, with N representing the number, or density, of mussels. The concentration of phytoplankton and detritus are given by P and D, respectively. Their corresponding far-field values are P_∞ and D_∞. The model currency is dry weight carbon (grams carbon or gC). Table 1 summarizes the variables and parameters of the model. Scaling analysis of Dowd (1997) identifies this model as containing the major ecological interactions important at higher mussel densities (e.g. we can ignore P growth within the box). In general terms, this nonlinear model describes predator (M) and prey (P and D) interactions in an open system (Nisbet et al. 1997). More specifically, it is an ecophysiological model of shellfish culture of the type used to assess ecosystem effects, bivalve growth and carrying capacity (Raillard and Ménesguen 1994, Dowd 1997, Chapelle et al. 2000). *Table 1*

The time rate of change in the weight of M is given by (11). This is measured as the net of gains through ingestion and losses due to respiration, and the quantity in brackets represents the scope for growth. The reference rates for ingestion and respiration are given by I and R, respectively. These rates are modulated by the dimensionless functions

$$f_I = \exp\{Q_{10} \cdot \text{Temp}\} \left(\frac{P + D}{k_M + P + D} \right), \qquad f_R = \exp\{Q_{10} \cdot \text{Temp}\}. \tag{14}$$

Both f_I and f_R include a temperature effect controlled by the physiological rate constant Q_{10}. The modulation function for ingestion f_I also includes saturation response to increasing food levels (a Holling type II functional response) and is controlled by the rate constant k_M. The ingested ration in (11) is scaled by a time-varying assimilation efficiency ϵ which measures the fraction of food ingested which is available for growth. It is given by

$$\epsilon = \epsilon_P \alpha_P + \epsilon_D \alpha_D. \tag{15}$$

where ϵ_P and ϵ_D are the assimilation efficiencies for mussels grazing on P and D, respectively. The fraction of P and S in the water at any time are approximated by

$$\alpha_P = \frac{P_\infty}{P_\infty + D_\infty}, \qquad \alpha_D = \frac{D_\infty}{P_\infty + D_\infty}. \tag{16}$$

The time rate of change of the freely-floating P and D components are described by (12) and (13), respectively. Exchange of these components with the adjacent open ocean is proportional to the concentration difference in the state variables between the box (P, D) and the far-field (P_∞, D_∞), scaled by an exchange coefficient K. This gradient-flux relation acts to couple P and D to their far-field values in this open system. Both P and D populations are reduced by the grazing activity of the mussel population (the product of M and N) in direct proportion to their relative abundance in the environment.

A baseline simulation for the model was constructed as follows. The governing equations (11)-(13) were discretized using a Euler method, resulting in a Markovian form for the nonlinear difference equations. The time dependent inputs (P_∞, D_∞ and temperature) are based on idealized annual cycles and are shown in Figure 1. The remaining parameters *Figure 1* required for the simulation, and their numerical values, are given in the second grouping in Table 1. The model was constructed to have a weekly time step, with a view to expediting the Monte Carlo computations at the heart of the probabilistic methods. Simulation runs covered a calendar year. Initial conditions for P and D were set based on their corresponding far-field concentrations, P_∞ and D_∞, at that time. Initial mussel weight takes a value of 0.1 gC (25mm shell length).

Figure 2 shows results from the baseline deterministic simulation. The state variables *Figure 2* P and D are tightly coupled to their far-field values due to the exchange of water between the bay and open ocean. As the year progresses, they become increasingly depressed below these far-field values as individual mussel weight increases, with a corresponding increase in the total mussel biomass, resulting in enhanced grazing of P and D. Note that at the end of the year P nearly equilibrates with P_∞ through exchange processes as it becomes a small fraction of the total seston, and temperature effects slow mussel ingestion. The predicted annual trajectory of individual mussel growth is given in Figure 2a. Superimposed on this curve are observations of average mussel growth based on field sampling of an ensemble of individuals (Grant et al. 1994). After one year, mussel weight increases from its starting value to about 0.4 gC (60mm shell length). The simulated mussel growth trajectory is generally consistent with the observations, but is smoother and under-predicts the observed changes occuring near the end of the year.

Stochastic Simulation

Stochastic simulation predicts the evolution of the probability density of the state variables. It takes into account prior knowledge of the model and parameters, without reference to the measurements. That is, the state space model equation (2) is used to determine the joint pdf of the state, $P(Z)$. According to (8), $P(Z)$ is made up of the pdf of the initial state transformed through multiplication by the time sequence of the various transition probabilities associated with the ecological dynamics. Monte Carlo simulation is used to approximate $P(Z)$. This involves generating a large number of realizations of the time trajectories of the state variables, after which the results are ensemble averaged. A useful analogy is the generation of particle paths according to a random walk, which is also simple Markov process. In the continuum limit, the governing Langevin equation gives rise to an advection-diffusion, or Fokker-Planck, equation describing the time evolution of the pdf of particle position (e.g. Gardiner 1997, Zwillinger 1989).

The discrete version of our ecosystem model (11)-(13) is written as a nonlinear stochastic difference equation, i.e.

$$x_t = g^*(x_{t-1}, \theta)\; e^{n_t^x} \tag{17}$$

where $x_t = (M_t, P_t, D_t)'$. The vector function g^* contains the ecosystem dynamics and at a given time t is functionally dependent on the state at the previous time, x_{t-1}, and the parameters, θ. A multiplicative error term $\exp\{n_t^x\}$ has also been included the right-hand side of (17) to account for discretization errors, as well as system noise or model errors. It is assumed that the n_t^x are zero-mean, normally distributed random vectors whose components are mutually uncorrelated and independent in time. A constant variance σ_x^2 is also assumed. This log-normal form for the multiplicative system noise ensures that the state variables remain non-negative. It also means that the variance of M_t, P_t and D_t scale with their time varying mean population level. This feature has been observed for both phytoplankton (Dowd et al. 2002) and individual shellfish weight (Gangnery et al. 2001). Here, we assume that $\sigma_x^2 = 0.05$ gC^2 m^{-6}.

A subset of the model parameters, θ, in (17) are also considered to be random variables. These have an associated noise process n^{θ}, considered to be independent of n_t^x. For the purposes of this study, parameter uncertainty is assumed to reside only in the mussel ingestion and respiration rates, I and R, so that $\theta = (I, R)'$. Together these quantities set the mussel energy balance, or scope for growth. The remaining parameters were treated as deterministic and take on constant values. The parameters are defined on a finite interval $[\theta_{\min}, \theta_{\max}]$, such that $\theta = \theta_{\min} + n^{\theta}$ where $n^{\theta} = (\theta_{\max} - \theta_{\min})\beta(a, b)$. Here, $\beta(a, b)$ is the standard beta distribution with parameters a and b. In this application $n^{\theta} = (n^I, n^R)'$, and are thus defined with respect to the parameters I and R. For I, the range is $[0, 0.15]$ and the distribution $\beta(5, 2)$ is used. For R, the range is $[0, 0.035]$ and the distribution $\beta(2, 5)$ is used. The range of possible parameter values and is rationalized based on literature values (Griffiths and Griffiths 1987). The prior pdfs for I and R are skewed to the right and left respectively. This assumes that ingestion is generally near its expected value of 0.1 day^{-1}, but may take on low values if grazing activity is reduced or suspended (Cranford and Hill 1999). The expected value of respiration is 0.01 day^{-1} but in situations of physiological stress it may take on much higher values (Griffiths and Griffiths 1987).

Note that the state vector can be redefined such that $z_t = (x_t, \theta)'$, thereby mapping the above model system into standard state space form (2). The noise term in (2) is then correspondingly modified such that $n_t = (n^x, n^{\theta})_t'$. To implement this nonlinear state space model, initial conditions for M, P and D, and the parameters I and R are set following the baseline deterministic simulation. These represent expected values for their associated random variables. The required prior information on the ecological dynamics and parameters has been completely specified. Stochastic simulation is used next to determine the prior pdf, $P(Z)$, of the state.

Probability distributions and ensemble statistics were computed for the state variables M, P and D using the shellfish ecosystem model. These were based on approximately 10^4 independent realizations of the time series of these ecological state variables[1]. Results from

[1]Stochastic simulation was implemented in both MATLAB and BUGS software to ensure consistency in the model implementation. Convergence and mixing of the Markov chain was extremely slow using BUGS. This was due to high correlations amongst successive components of Z which makes Gibbs sampling less efficient. Results for stochastic simulation are based on MATLAB runs.

this stochastic simulation are given in Figure 3. These are reported as (i) time series of the median of the state along with its 95% confidence intervals (Figure 3a,c,e), and (ii) the time evolution of the pdf for each of the state variables (Figure 3b,d,f). *Figure 3*

The time series of median mussel weight, M, has slightly greater values than the corresponding results from the baseline deterministic simulation. This reflects the use of distinct beta distributions as priors for the parameters I and R wherein ingestion tends to be inflated relative to respiration, leading to increased M growth. The 95% confidence interval generally widens with increasing time. This results from the accumulation of uncertainty in M due to parameter uncertainty (the scope for growth computation enters the state multiplicatively at every time step). A lesser contribution is also due to the system noise variance for mussel growth which increases in magnitude as M increases. The time evolution of the pdf of M also clearly shows rapid spreading of the variance. As the simulation progresses, the higher order moments of M, skewness and kurtosis, increase with time. The pdfs are not Gaussian, which is not surprising given the nonlinear model and the use of beta distributions for the parameters and multiplicative lognormal system noise. It is also interesting to note the resemblance of these errors in the mussel growth trajectories to those determined by a population dynamics model of inter-individual variability in a mussel population (Gangnery et al. 2001).

For both phytoplankton, P, and detritus, D, the time evolution of the median follows closely the baseline deterministic simulation. Its variance generally scales with the level of P, reflecting the properties of the system noise variance. However, the variance of D, as well as the skewness and kurtosis, increase with time. This feature results from mussel grazing activity. That is, during weeks 0-20, P and D levels are determined mainly by exchange and their distribution is set by the lognormal system noise. After this period, M grazing begins to affect these variables; the increasing variance in M acts to increase the uncertainty in D as these populations co-evolve. This effect is also seen in P during the period from weeks 20-40. The differential response of the uncertainty in P and D to mussel grazing rests in the strength of the coupling between these food sources and grazing activity. Clearly, the ecological linkages in the dynamics provide a means to couple the co-evolution of the pdfs of the state variables.

The Inverse Model

Inverse methods blend the information contained in the model dynamics and parameters with available observations in order to refine estimates for the state of the ecological system. This procedure was illustrated from a probabilistic perspective using Bayes' theorem (6). It was also demonstrated that the goal for the inverse problem is to evaluate $P(Z|Y)$, the joint posterior density of the state and parameters after taking the observations into account. Inverse methods thus update the prior $P(Z)$, as specified by the state equations (2) and sampled via stochastic simulation in the previous section, to the posterior $P(Z|Y)$ via the observation equation (3). Here, we make use of observations of mussel weight, M, as given in Figure 2a. Note that these are direct measurements of a single state variable. Hence, the observation operator in (3) is defined quite simply, i.e. it is non-null only when associated with the state variable M at the times when observations are available. The observation error v_t in (3) is taken to be additive. It is assumed to be an independent and identically distributed normal random variable with zero mean and variance $\sigma_v^2 = 10^{-4}\mathrm{gC}^2$. This implies that the M measurements are relatively precise.

Markov Chain Monte Carlo methods were used to evaluate the posterior density $P(Z|Y)$. Implementation used the BUGS software (Spiegelhalter et al. 1996). Approximately 10^4 quasi-independent random samples of $P(Z|Y)$ were drawn from the simulated Markov Chain. (This was achieved using a sampling interval greater than the decorrelation time of the chain.) Careful attention was given to ensuring the convergence of the chain to stationarity prior to drawing samples, and to making sure the chain achieved mixing during the Monte Carlo integration. To ensure the robustness and consistency of the results, multiple realizations of the chain were generated using different starting values for the random variables, as well as different random number seeds. The reader is referred to Gilks et al. (1996) for more information on the setup and diagnosis of MCMC simulations.

Figure 4 shows estimates for the posterior pdfs of the parameters I and R, along with their associated priors. Note that the solution has constrained the posterior pdfs to lie in the *Figure 4*
same interval as their respective priors. The posterior pdf for the ingestion rate, I, closely resembles the prior pdf, with only a slight shift in the mode of the distribution. Assuming the posterior follows a beta pdf, the distributional parameters are estimated as $\beta(8.4, 2.3)$,

compared to $\beta(5, 2)$ for the prior. The mean and variance differ only slightly. In contrast, the posterior pdf for the respiration rate, R, differs significantly from its prior. The mean for the posterior and prior are both near 0.01 day^{-1}. However, the variance of the posterior (3.5×10^{-6}day^{-2}) is reduced relative to the prior (3.1×10^{-5}day^{-2}). Most importantly, the shape of the prior has changed to a more symmetric and peaked distribution. The additional information associated with the observations of M have allowed refinement of the posterior pdf of the respiration parameter, but left the pdf of the ingestion rate almost unchanged. The inverse method thus acts to ensure consistency between observed and predicted M by altering the respiration terms in the scope for growth computation.

Figure 5 shows the time evolution of the posterior estimates for the state variables M, P and D. These are reported in terms of time series of the median of the state variables *Figure 5* along with their 95% confidence intervals (Figure 5a,c,e), as well as in terms of the time evolution of the pdf for each of the state variables (Figure 5b,d,f). The posterior pdf of M appears to be nearly Gaussian with a time-dependent mean level and relatively small and constant variance. This is in sharp contrast to the corresponding results from the stochastic simulation, where variance spreading was strongly evident. The additional information contained in the M observations have constrained the posterior density to be much closer to the measurements, which is consistent with the assumption of a small measurement error variance. Another feature of note is the growth in the variance of M in the observation void near mid-record. The envelope of 95% confidence is widest at the greatest distance away from the observations. Deviations from normality are also most pronounced here. The median values for P and S are similar to those from the stochastic simulation, excepting differences near the end of the year. The additional information in the M observations has substantially reduced the overall variance of both P and D during the latter half of the simulation period, where these food sources are being strongly influenced by M grazing activity. (A slight increase in variance is evident at the initial time.) Non-normality in the posterior pdfs is generally lessened compared to the corresponding results from the stochastic simulation. The Bayesian inverse method appears to be an effective means to blend the measurement information with a dynamical model to produce posterior state and parameter estimates.

4 Summary and Conclusions

This study has investigated a probabilistic, or Bayesian, approach to the ecosystem inverse problem. The particular emphasis is on time-dependent models of interacting populations used in marine ecology. The overall goal is to combine (i) prior information from the model dynamics and substantive knowledge about unknown parameters with (ii) information stemming from the observations. This produces posterior estimates of probability density function of the time-varying ecological state variables and parameters. The use of inverse methods in marine ecology is in its infancy, but developing rapidly. Applications to date have used approaches based mainly on optimization methods (Lawson et al. 1995, Vallino 2000, Natvik et al. 2001). Bayesian methods have been used for parameter estimation in ecological models (Dilks et al. 1992, Aldenberg et al. 1995, Steinberg et al. 1997, Harmon and Challenor 1997). Reichart and Omlin (1997) suggest that Bayesian techniques also be used to incorporate uncertainty in the structure of ecological models. Because of their nonlinear and non-Gaussian nature, it seems appropriate to blend time dependent ecological models and measurements using a probabilistic, or Bayesian, framework cast in a general state space representation, a conclusion well supported by the statistical time series literature (Kitagawa 1987, Carlin et al. 1992). This not only allows for incorporation of nonlinear models and measurement operators, but also for the use of very general error processes, such as those having non-normal distributions and entering the system multiplicatively. This study outlines the background to, and an application of, a Bayesian approach to inverse modelling for marine ecosystems.

The ideas behind the Bayesian ecosystem inverse approach were illustrated using a simple biophysical shellfish growth model describing the co-evolution of a mussel population and their food sources, plankton and detritus. This model is of the type used to investigate ecosystem effects of shellfish aquaculture and assess carrying capacity (Raillard and Ménesguen 1994, Dowd 1997, Chapelle et al 2000). This nonlinear time-dependent model was cast in a state space framework, making use of an augmented state vector which included both the state variables, as well as the model parameters (Kitagawa 1998). The specification of prior information is a contentious issue for Bayesian applications. Our choices are based on both parsimony and biological realism, as well as a desire to illustrate the flexibility of the method. Multiplicative non-Gaussian system noise was used in the discretized model equations. This took the form of a lognormal distribution which acts to ensure that the ecological populations remain non-negative. (This contrasts to a common assumption made in oceanographic data assimilation which makes use additive Gaussian noise and does not enforce non-negativity.) A subset of the model parameters were also considered as random variables. These were restricted to non-negative intervals consistent with their literature values, and their prior pdfs were based on beta distributions. Stochastic simulation allowed us to first determine the time evolution of the prior pdf of the state, without reference to any observations. A set of available measurements on one ecosystem component was then used to solve the inverse problem using Bayesian inference. In this manner, the time-varying posterior pdfs of the ecological state and parameters were determined. It was shown that posterior estimates of the parameters and state were all influenced by the information in the measurements, as dictated by the strength of their dynamic coupling. The application demonstrated how ecological dynamics link not only the mean state, but also quantities related to the uncertainties, i.e. in terms of their joint probability density functions. The implication is that models consider the ecological state as a random variable.

While the theory of Bayesian inverse methods is straightforward, practical applications must consider a number of factors. Chief amongst these is the complexity of the dynamic model and its influence on the associated dimensionality of the forward model, and hence the inverse problem. For example, Harmon and Challenor (1997) used a MCMC technique for parameter estimation in an ecosystem model. They concluded that that it was impractical to determine second order properties (covariances) for more than about 10 parameters. In this study, Bayesian posterior computation was performed using the Gibbs sampler and proved suitable for determining the posterior density of the time-varying state and the parameters. The use of a simple time-dependent ecosystem box model with careful attention given to controlling the dimensionality facilitated this application. With minimal extensions (multiple boxes and additional ecosystem components) this model can provide a realistic description of a coastal marine ecosystem (Dowd 1997), and yet still prove amenable to Bayesian inverse methods. The trend in marine ecosystem modelling is, however, towards multi-compartment ecosystem models coupled to ocean circulation models (Hofmann and Lascara 1998). The extent to which probabilistic techniques can be adapted to these very high dimension models is unclear. It is suggested that the use of Bayesian inverse methods with complex ecological models may require the development of specifically tailored Markov Chain Monte Carlo algorithms that are more efficient than the Gibbs sampler. It is also encouraging to note that for complex physical models of ocean and atmospheric circulation patterns, hybrid approximate methods based on probabilistic Monte Carlo methods and optimization techniques have proven useful (Evensen 1994), and are being applied to biological systems (Natvik et al. 2001, Losa et al. 2002). It is still an open question as to whether simplified ecological models with sophisticated data assimilation, such as the Bayesian inverse approach, prove to be more effective than simple inverse methods used with very complex dynamical models.

In summary, the use of a Bayesian inverse approach to marine ecosystem modelling appears promising. Inverse methods are expected to become widely used in marine ecosystem modelling as new, automated observing technologies are developed. An important use of inverse techniques is to provide guidance for the refinement of the mathematical structure of ecological models. Bayesian analysis is particularly well suited to this role, having the capability to make inferences on parameters and the state, as well as carry out hypothesis testing. However, a great deal more work must be carried out to assess the feasibility of Bayesian inverse methods as applied to ecological models. It is hoped that this study provides stimulation for the further development of novel inverse techniques appropriate for ecosystem models.

Acknowledgements

The authors would like to thank Dr. W. Carlisle Thacker of the NOAA Atlantic Oceanographic and Meteorological Laboratory for stimulating discussions on the role of Bayesian methods in data assimilation. This work was supported by Fisheries & Oceans Canada and the Royal Society of New Zealand Marsden Fund. The authors also thank Dr. Svetlana Losa of the Bedford Institute of Oceanography for reviewing the manuscript.

References

Aldenberg, T., J.H. Janse and P.R.G. Kramer. 1995. Fitting the dynamics model PCLake to a multi lake survey using Bayesian statistics. *Ecological Modelling.* 78:83-90.

Alspach, D.L., H.W. Sorenson. 1972. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, AC-17: 439-448.

Batchelder, H.P., J.R. Van Keuren, R. Vaillancourt, and E. Swift. 1995. Spatial and temporal distributions of acoustically estimated zooplankton biomass near the Marine Light-Mixed Layers station (59 degree 30'N, 21 degree 00'W) in the North Atlantic in May 1991. *Journal of Geophysical Research.* 100(C4):6549-6563.

Bennett, A.F. 1992. *Inverse Methods in Physical Oceanography.* Cambridge University Press, New York. 346pp.

Bryson, A.E. and Y. Ho. 1969. *Applied Optimal Control.* Blaisdell, Waltham. 481pp.

Carlin, B.P., N.G. Polson and D.S. Stoffer. 1992. A Monte Carlo approach to nonnormal and nonlinear state space modelling. *Journal of the American Statistical Association.* 87:493-500.

Chapelle A., A. Menesguen, J-M. Deslous-Paoli, P. Souchu, N. Mazouni, A. Vaquer, and B. Millet. 2000. Modelling nitrogen, primary production and oxygen in a Mediterranean lagoon. Impact of oysters farming and inputs from the watershed. *Ecological Modelling.* 127:161-181.

Cowles, M.K. and B.P. Carlin. 1996. Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association.* 91: 883–904.

Cranford, P.J. and P.S. Hill. 1999. Seasonal variation in food utilization by the suspension-feeding bivalve *Mytilus edulis* and *Placopeten magellanicus. Marine Ecology Progress Series.* 190:223-239.

Cullen, J.J., A.M. Ciotti, R.F. Davis, and M.R. Lewis. 1997. Optical detection and assessment of algal blooms. *Limnololgy and Oceanography.* 42:1223-1239.

Dilks, D.W., R.P. Canale and P.G. Meier. 1992. Development of Bayesian Monte Carlo techniques for water quality model uncertainty. *Ecological Modelling.* 62:149-162.

Doucet, A., N. de Freitas, N. Gordon. 2001. *Sequential Monte Carlo Methods in Practice.* Springer, New York, 581pp.

Dowd, M. 1997. On predicting the growth of cultured bivalves. *Ecological Modelling.* 104:113-131.

Dowd, M., J.L. Martin, M.M. LeGresley, A. Hanke and F.H. Page. 2002. Interannual variability in a plankton time series. *Environmetrics.* (in press).

Edwards, A.M. 2001. Adding detritus to a nutrient-phytoplankton-zooplankton model: a dynamical systems approach. *Journal of Plankton Research.* 23(4):389-413.

Evensen. G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research.* 99:10143-10162.

Franks, P.J.S., J.S. Wroblewski, and G.R. Flierl. 1986. Behaviour of a simple plankton model with food-level acclimation by herbivores. *Marine Biology.* 91:121-129.

Fruehwirth-Schnatter, S. 1994. Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis.* 15: 183–202.

Gamerman, D. 1997. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.* CRC Press. 264pp.

Gangnery, A., C. Bacher, and D. Buestel. 2001. Assessing the production and impact of cultivated oysters in the Thau lagoon (Mediterranee, France) with a population dynamics model. *Canadian Journal of Fisheries and Aquatic Sciences.* 58:1012-1020.

Gardiner, C.W. 1997. *Handbook of Stochastic Methods.* Springer. Berlin. 442pp.

Gelfand, A.E. and A.F.M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association.* 85:398-409.

Gelman, A., J. Carlin, H.S. Stern and D.B. Rubin. 1995. *Bayesian Data Analysis.* CRC Press. 552pp.

Geman, S. and D. Geman. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 6:721-741.

Gilks, W.R. S. Richardson and D.J. Spiegelhalter (eds.) 1996. *Markov Chain Monte Carlo in Practice.* London, Chapman and Hall.

Grant, J., M. Dowd, K.R. Thompson, C. Emerson, and A. Hatcher. 1993. Perspectives on field studies and related biological models of bivalve growth. In: *Bivalve filter feeders and marine ecosystem processes*, Dame, R. (ed.), New York, Springer Verlag, pp. 371-420.

Griffiths C.L. and R.J. Griffiths. 1987. Bivalvia. In: *Animal Energetics. Vol 2. Bivalvia through Reptilia.* T.J. Pandian and F.J. Vernberg, (eds.). Academic Press, New York. p. 2-87.

Harmon, R. and P. Challenor. 1997. A Markov chain Monte Carlo method for estimation and assimilation into models. *Ecological Modelling.* 101:41-59.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika.* 57:97-109.

Hodges, P.E. and D.F. Hale. 1993. A computational method for estimating densities of non-Gaussian nonstationary univariate time series. *Journal of Time Series Analysis.* 14: 163–178.

Hofmann, E.E. and C.M. Lascara. 1998. Overview of interdisciplinary modeling for marine ecosystems. In: *The Sea: Ideas and Observations on Progress in the Study of the Seas. Volume 10: The Global Coastal Ocean: Processes and Methods.* Brink, K.H. and A.R. Robinson (eds). New York, Wiley.

Jazwinski, A.H. 1970. *Stochastic Processes and Filtering Theory.* Academic Press, New York. 376pp.

Kahru, M. and B.G. Mitchell. 2001. Seasonal and nonseasonal variability of satellite-derived chlorophyll and colored dissolved organic matter concentration in the California Current. *Journal of Geophysical Research*, 106: 2517-2529.

Kitagawa, G. 1987. Non-Gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association.* 82:1032-1041.

Kitagawa, G. 1998. A self-organizing state space model. *Journal of the American Statistical Association.* 93(443):1203-1215.

Kitagawa, G. and W. Gersch. 1996. *Smoothness Priors Analysis of Time Series.* Lecture Notes in Statistics No. 116. Springer-Verlag, New York. 261pp.

Lawson, L.M., Y.H. Sptiz, E.E. Hofmann and R.L. Long. 1995. A data assimilation technique applied to a predator-prey model. *Bulletin of Mathematical Biology.* 57:593-617.

Losa, S.N., G.A. Kivman and V.A. Ryabchenko. 2002. Weak constraint parameter estimation for a simple ocean ecosystem model: what can we learn about the model and the data. *Journal of Marine Systems.* (to appear)

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chemical Physics.* 21:1087-1091.

Meyer, R. and N.L. Christensen. 2000. Bayesian Reconstruction of Chaotic Dynamical Systems. *Physical Review E.* 62: 3535-3542.

Meyer, R. and R. Millar. 1999. BUGS in Bayesian stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences.* 56: 1078–1087.

Meyer, R. and J. Yu. 2000. BUGS for a Bayesian analysis of stochastic volatility models. *The Econometrics Journal.* 3: 198-215.

Millar, R.B. and R. Meyer. 2000. Bayesian State-Space Modeling of Age-Structured Data: Fitting a Model is Just the Beginning. *Canadian Journal of Fisheries and Aquatic Sciences.* 57: 43-50.

Natvik, L.J., M. Eknes, and G. Evensen. 2001. A weak constraint inverse for a zero dimensional marine ecosystem model. *Journal of Marine Systems.* 28:19-44.

Nisbet, R.M., S. Diehl, W.G. Wilson, S.D. Cooper, D.D. Donaldson, and K. Kratz. 1997. Primary productivity gradients and short-term population dynamics in open systems. *Ecological Monographs.* 67(4):535-553.

Popova, E.E., M.J.R. Fasham, A.V. Osipov, V.A. Ryabchenko. 1997. Chaotic behaviour in an ocean ecoystem model under seasonal external forcing. *Journal of Plankton Research.* 19:1495-1515.

Raillard O. and A. Ménesguen. 1994. An ecosystem box model for estimating the carrying capacity of a macrotidal shellfish ecosystem. *Marine Ecology Process Series.* 115:117-130.

Reichert, P. and M. Omlin. 1997. On the usefulness of overparameterized ecological models. *Ecological Modelling.* 95:289-299.

Robert, C.P. and G. Casella. 1999. *Monte Carlo Statistical Methods.* Springer-Verlag, New York.

Sieracki, C.K., M.E. Sieracki, and C.S. Yentsch. 1998. An imaging-in-flow system for automated analysis of marine microplankton. *Marine Ecology Progress Series.* 168:285-296.

Soudant, D., B. Beliaeff, and G. Thomas. 1997. Dynamic linear Bayesian models in phytoplankton ecology. *Ecological Modelling.*99:161-169.

Spiegelhalter, D.J., A. Thomas, N.G. Best and W.R. Gilks. 1995. *BUGS: Bayesian inference Using Gibbs Sampling, Version 0.50.* MRC Biostatistics Unit, Cambridge.

Steinberg, L.J., K.H. Reckhow, and R.L. Wolpert. 1997. Characterization of parameters in mechanistic models: a case study of a PCB fate and transport model. *Ecological Modelling.* 97:35-46.

Steele, J.H. and E.W. Henderson. 1981. A simple plankton model. *American Naturalist.* 117:676-691.

Tarantola, A. 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Elsevier, New York. 613pp.

Thacker, W.C and R.B. Long. 1988. Fitting dynamics to data. *Journal of Geophysical Research.* 93:1227-1240.

Thompson, K.R., M. Dowd, Y. Lu., and B. Smith. 2000. Oceanographic data assimilation and regression analysis. *Environmetrics.* 11:183-196.

Vallino, J.J. 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *Journal of Marine Research.* 58:117-164.

West, M., P.J. Harrison, and H.S. Migon. 1985. Dynamic generalized linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association* 80: 73–97.

Wunsch, C. 1996. *The Ocean Circulation Inverse Problem.* Cambridge University Press, Cambridge. 442pp.

Zwillinger, D. 1989. *Handbook of Differential Equations*. Academic Press. London. 673pp.

Table 1: Defintion of quantities using in the shellfish ecosystem model. These are grouped according to variable type. For each quantity the following information is given: units, numerical value (or its source), and a brief definition. See text for more details.

Quantity	Units	Value	Definition
(i) State			
M	gC	eq. (11)	individual Mussel Weight
P	gC m^{-3}	eq. (12)	phytoplankton biomass
D	gC m^{-3}	eq. (13)	detritus biomass
(ii) Parameters			
N	m^{-3}	1	number (density) of mussels
K	day^{-1}	0.1	exchange/flushing coefficient
k_M	gC m^{-3}	1	half-saturation constant for ingestion
I	day^{-1}	0.1	ingestion rate
R	day^{-1}	0.01	respiration rate
Q_{10}	$^oC^{-1}$	0.05	physiological temperature effect
ϵ_P	-	0.9	assimilated fraction for ingested P
ϵ_D	-	0.2	assimilated fraction for ingested D
(iii) External Inputs			
P_∞	gC m^{-3}	Fig. 1a	far-field P biomass
D_∞	gC m^{-3}	Fig. 1b	far-field D biomass
Temp	^oC	Fig. 1c	temperature
(iv) Modulating Functions			
f_I	-	eq. (14a)	modulation function for M ingestion
f_R	-	eq. (14b)	modulation function for M respiration
(v) Derived Parameters			
$\epsilon(t)$	-	eq. (15)	assimilated fraction of ingested ration
$\alpha_P(t)$	-	eq. (16a)	fraction of plankton in seston
$\alpha_D(t)$	-	eq. (16b)	fraction of detritus in seston

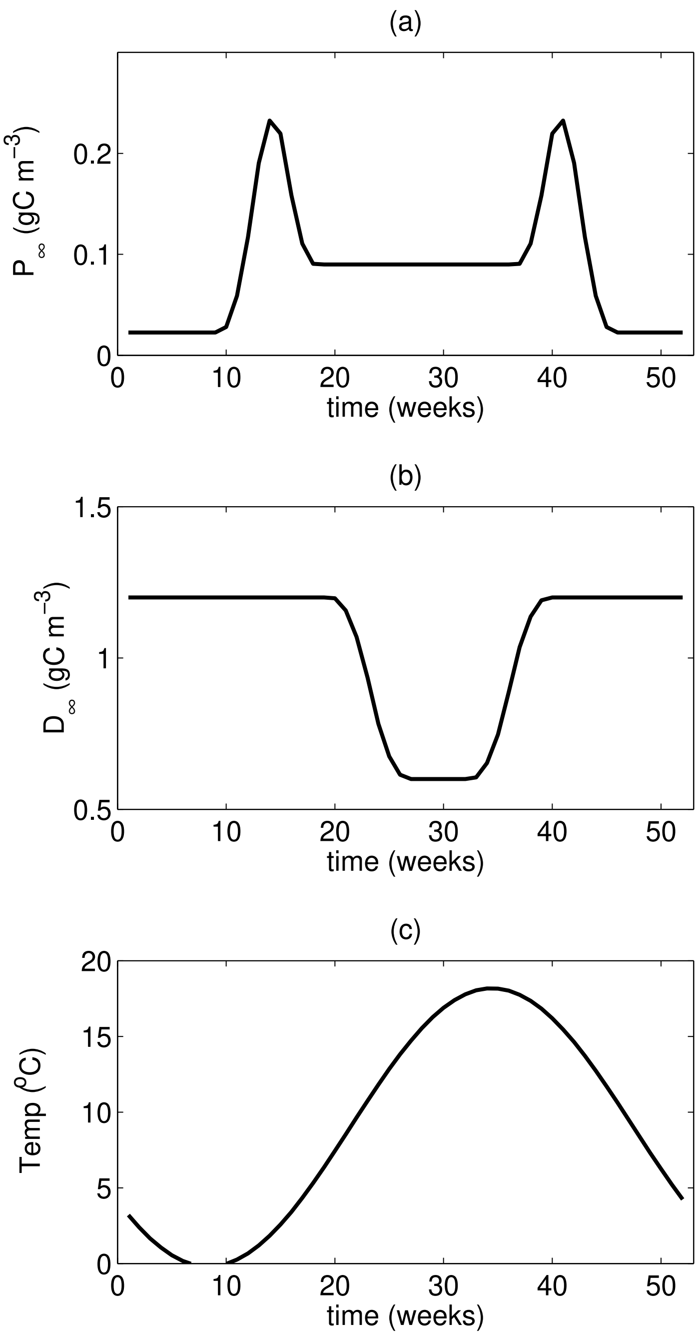

Figure 1: Inputs for the shellfish ecosystem model in the form of idealized annual cycles for: (a) phytoplankton far-field concentration, P_∞, (b) detritus far-field concentration, D_∞, and (c) temperature. The time axis covers the calendar year.

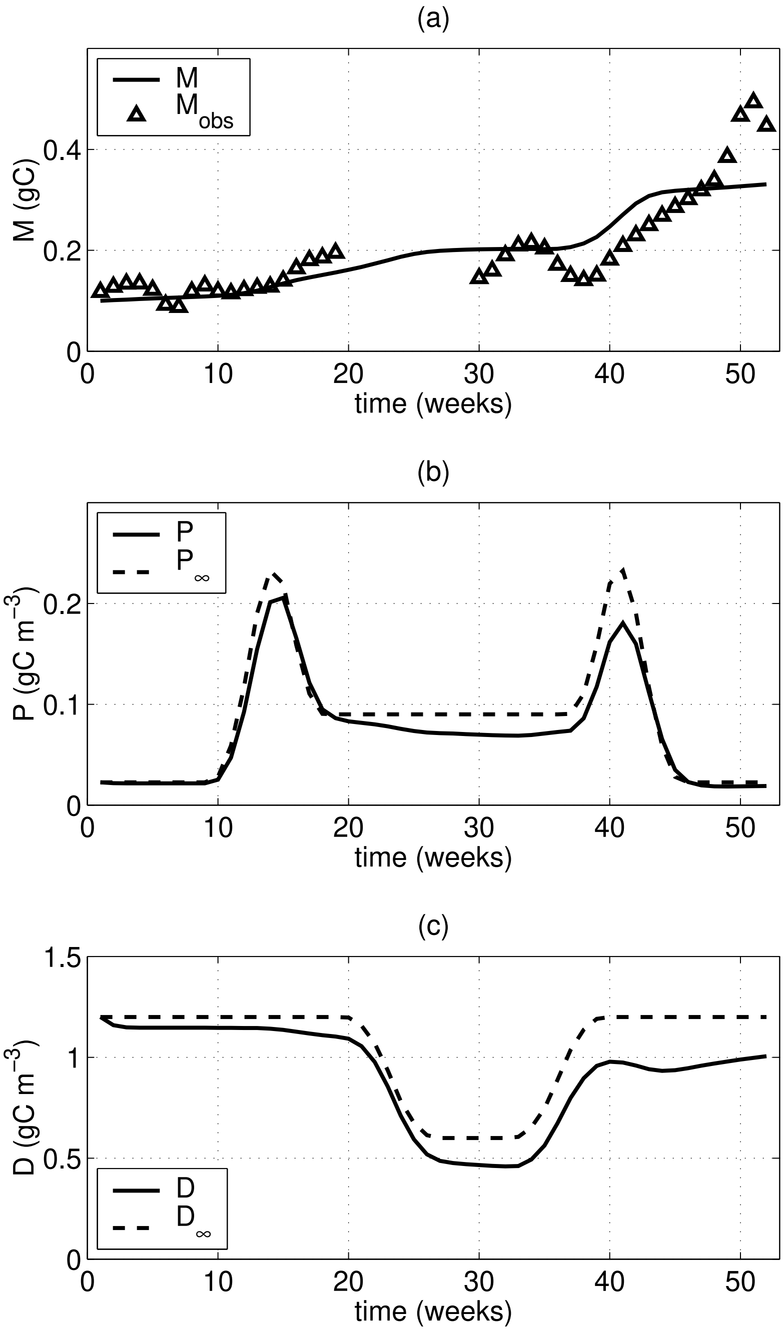

Figure 2: Baseline results from the deterministic simulation of the ecosystem model. (a): Predicted (solid line) and observed (triangle) individual mussel weight M. (b): Simulated P cycle (solid line), with the far field P_∞ shown for reference (dashed line). (c): Simulated D cycle (solid line), with the far field D_∞ shown for reference (dashed line). The time axis covers the calendar year.

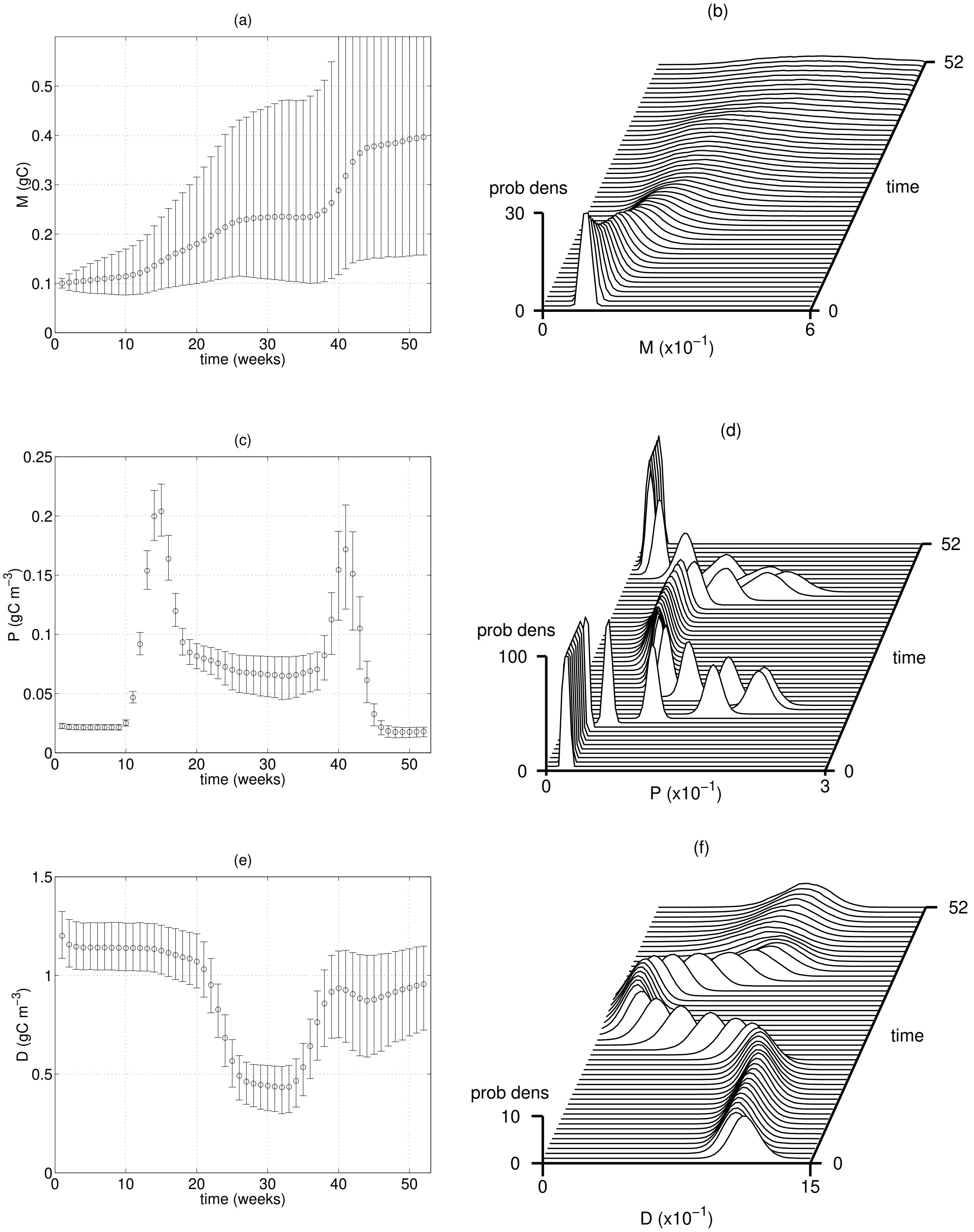

Figure 3: Results from the stochastic simulation of the ecosystem model for the state variables M, P and D. The panels on the right show the time sequence of the median (dots) and the 95% confidence interval. The panels on the left show the time evolution of the probability density functions.

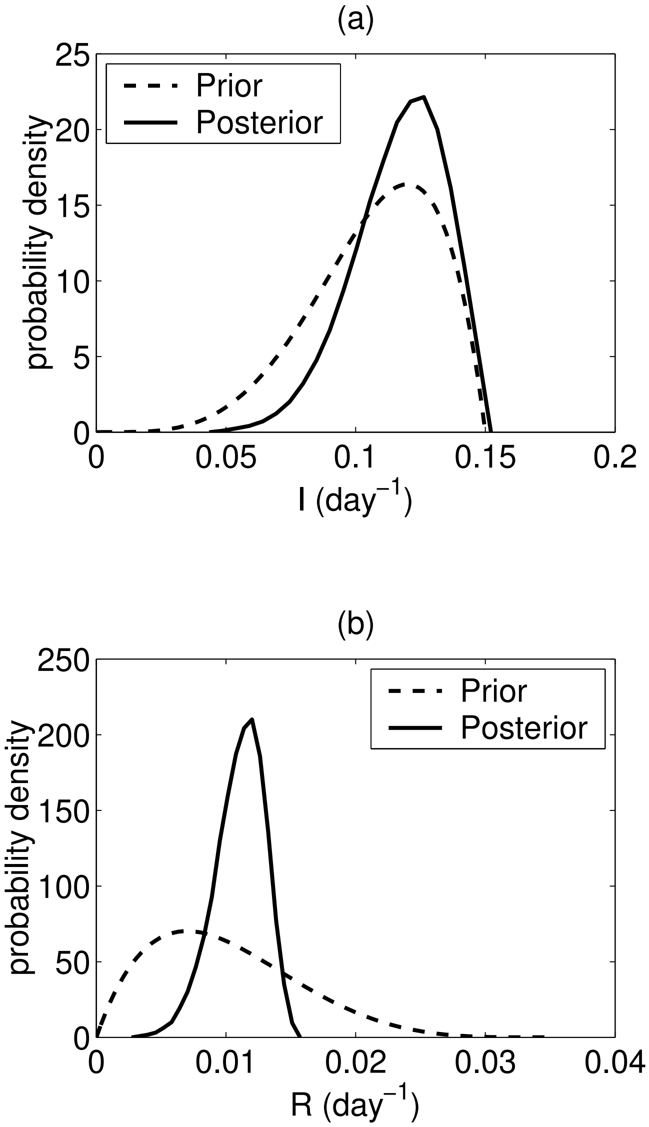

Figure 4: Prior (dashed line) and posterior (solid line) probability density functions for the parameters I (panel a) and R (panel b).

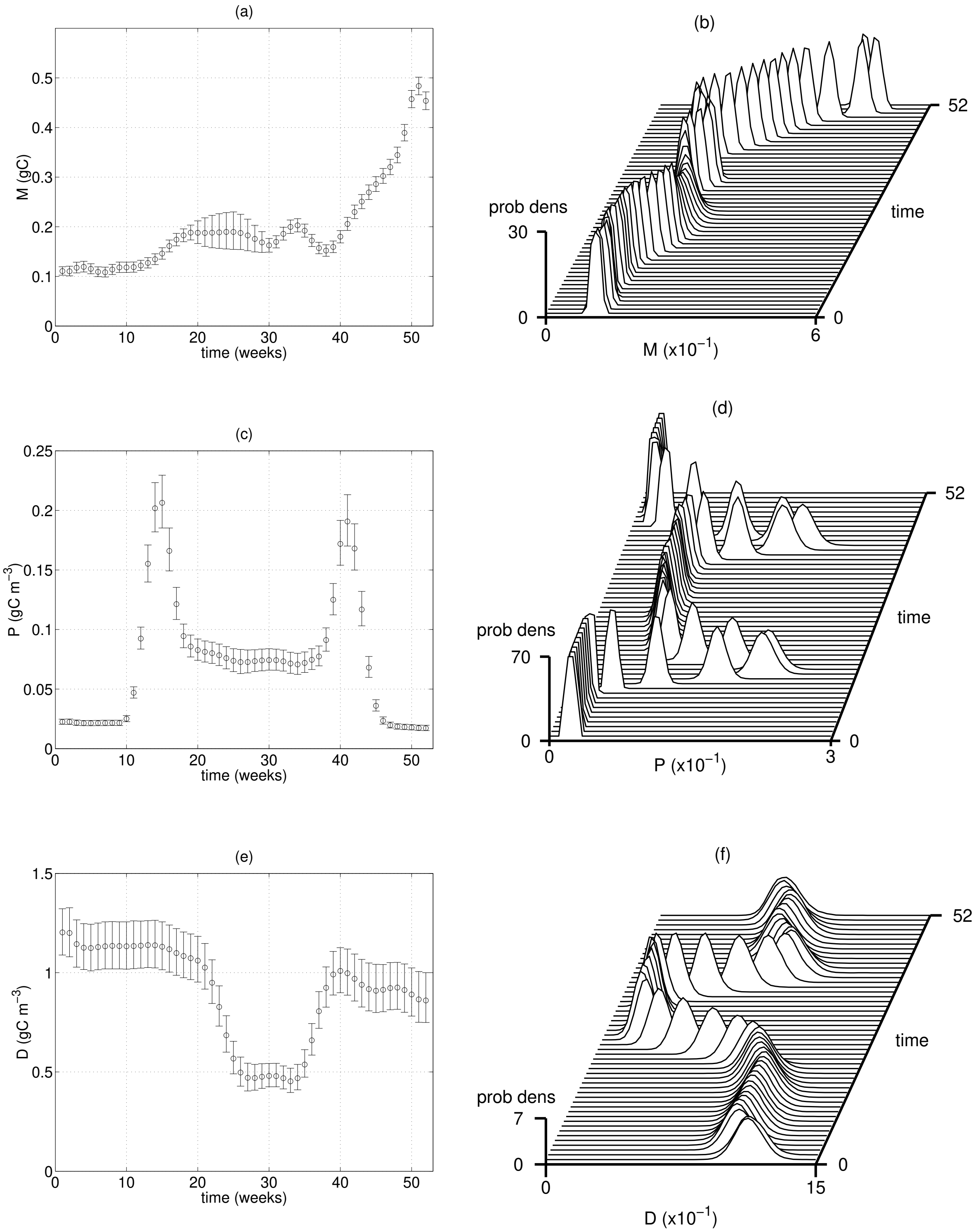

Figure 5: Results from application of the Bayesian inverse method for the state variables M, P and D. The panels on the right show the time sequence of the posterior median (dots) and the 95% confidence interval. The panels on the left show the time evolution of the posterior probability density functions.