ENVIRONMETRICS
Environmetrics 2009; **20**: 912–927
Published online 13 October 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/env.962

Dynamic two state stochastic models for ecological regime shifts

Jan Kloppenborg Møller[1,2*,†], Jacob Carstensen[2], Henrik Madsen[1] and Tom Andersen[3]

[1] *DTU Informatics, Richard Pedersens Plads, Technical University of Denmark Building 321, DK-2800 Lyngby, Denmark*
[2] *National Environmental Research Institute, Fredriksborgvej 399, DK-4000 Roskilde, Denmark*
[3] *University of Oslo, Biologisk Institut, Kristine Bonevies hus, Blindernveien 31, 0371 Oslo, Norway*

SUMMARY

A simple non-linear stochastic two state, discrete-time model is presented. The interaction between benthic and pelagic vegetation in aquatic ecosystems subject to changing external nutrient loading is described by the non-linear functions. The dynamical behavior of the deterministic part of the model illustrates that hysteresis effect and regime shifts can be obtained for a limited range of parameter values only. The effect of multiplicative noise components entering at different levels of the model is presented and discussed. Including noise leads to very different results on the stability of regimes, depending on how the noise propagates through the system. The dynamical properties of a system should therefore be described through propagation of the state distributions rather than the state means and consequently, stochastic models should be compared in a probabilistic framework.

KEY WORDS: aquatic ecosystems; bistability; hysteresis; stochastic state space models; multiplicative noise

1. INTRODUCTION

Eutrophication is a result of nutrient enrichment of ecosystems. For aquatic ecosystems the link between algae blooms and eutrophication was recognized in the 1960s (Schindler, 2006). The main nutrient sources causing eutrophication are generally nitrogen (N) in marine systems and phosphorus (P) in freshwater systems. The human contribution to nutrient inputs derives from industry and households (dominant sources of P), as well as agriculture (dominant source of N).

Holling (1973) and May (1977) are early examples of analyses of systems with multiple stable states in biology. In more recent years this hypothesis has also been proposed for aquatic systems (see Schindler, 2006 for a historic presentation). The presence of alternative stable states is potentially

*Correspondence to: J. K. Møller, DTU Informatics, Richard Pedersens Plads, Technical University of Denmark Building 321, DK-2800 Lyngby, Denmark
†E-mail: jkm@imm.dtu.dk

Received 27 February 2008
Accepted 2 September 2008

reflected in hysteresis effects or even irreversible changes of a system caused by excessive nutrient loading (eutrophication) on the system. There is a bulk literature on this subject in the context of lakes, e.g., Scheffer *et al*. (2003), Carpenter (2005), and others. These kinds of shifts, where the dynamics of the system change suddenly, will be referred to as regime shifts.

An example of a simple system with regime shifts is given in Carpenter *et al*. (1999), where P in the water column is modeled as a function of the loadings. The dynamic part includes a non-linear recycling term. This construction can give rise to alternative stable states, hysteresis effects and irreversible changes can be observed.

Others (e.g., Scheffer *et al*., 2003) give examples of two dimensional (e.g., floating and submerged plants) systems with bistable dynamics. Scheffer *et al*. (2003) offer a conceptual empirical verification of the bistable hypothesis in controlled experiments. Such verification is, however, difficult in real ecosystems. The presence of different time constants in the system may (as pointed out by Carpenter, 2005) appear to be hysteresis effects or even irreversible changes of the system. Systems may take hundreds or even thousands of years to recover after heavy loading has caused a break down of the system over a matter just of years or decades. Therefore,it can also be argued that, from a practical point of view, the system has experienced an irreversible change.

An important issue for systems with alternative stable states is the stability of the states. Such analysis can be addressed with bifurcation analysis (see e.g., Scheffer *et al*., 2003), and this will show how far the system is from an abrupt change. Since ecosystems are very complex systems that cannot be encapsulated in a simple mathematical formulation, there is bound to be noise in the system ("system noise"). In addition to this, measurements of variables related to these systems are uncertain. An adequate description of these systems is therefore a stochastic one. The stochastic nature of the system implies that we can only assign a probability of a regime shift appearing in the next time interval, and the best strategy might therefore be to use a precautionary policy.

Carpenter *et al*. (1999) and Ludwig *et al*. (2003) support this approach and they explore management strategies for lakes using an economic utility model added to the dynamical system description. Both papers advocate a precautionary policy due to the uncertainty in the system. However, studies examining the effect of stochastic perturbations in aquatic models are still few and far between, and mostly concerned with stochastic variations in model forcing. For example, Ludwig *et al*. (2003) analyzed the sensitivity of a simple lake model having two states, phosphorus in the water column and in the sediments respectively, with respect to a log-normal distributed phosphorus loading. Carpenter and Brock (2006) analyze a full stochastic model in the SDE (stochastic differential equations) framework, this analysis led to a proposed management strategy based on variance monitoring.

The objective of this study is to illustrate hysteresis and regime shifts in a simple model consisting of two states for pelagic and benthic vegetation, respectively. Moreover, we will analyze the model with stochastic perturbations added to the forcing as well as in the system equations. Guttal and Jayaprakash (2007) analyze the effect of stochastic forcing in two one-dimensional systems with hysteresis effects and deduce some results on the assymptotic behavior of these systems. We will analyze a two-dimensional system and further we will focus more on the transients. The difference between observation noise and system noise will also be emphasized. The adequate description of such systems is a state space formulation.

We will present some characteristics of this model and emphasize where caution is needed to ensure stability. The model is formulated in a discrete time setting and not, as is often the case, by means of differential equations. The state space model and its rationale are introduced in Section 2. The behavior of the deterministic part of the model is analyzed in Section 3. In Section 4 the effect of stochastic perturbations added to model input and system equations is investigated.

2. THE STATE SPACE MODEL

Aquatic systems undergoing the adverse effects of eutrophication are characterized by changes from benthic to pelagic production (Scheffer *et al.*, 2001; Carpenter, 2005). Here, we formulate a general model for aquatic ecosystems with two states for a nutrient (typically nitrogen for marine ecosystems and phosphorus for lakes), and analyze it within a state space framework. This implies that we will separate the random variations into "system noise" and "observation noise," in addition noise in the input is allowed. The aim is to analyze how the different noise components propagate through the system of equations. The stochastic state space model is given by the system equation

$$\begin{bmatrix} X_{p,t} \\ X_{b,t} \end{bmatrix} = N_{ex,t} \begin{bmatrix} a_p \\ 0 \end{bmatrix} + \begin{bmatrix} \xi_{p,t} & 0 \\ 0 & \xi_{b,t} \end{bmatrix} \times \left(\begin{bmatrix} b_{p,t-1} & 0 \\ b_{p,b} f_{t-1} & b_b \end{bmatrix} \begin{bmatrix} X_{p,t-1} \\ X_{b,t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ K f_{t-1} \end{bmatrix} \right) \quad (1)$$

and the observation equation

$$\begin{bmatrix} Y_{p,t} \\ Y_{b,t} \end{bmatrix} = \begin{bmatrix} \lambda_{p,t} & 0 \\ 0 & \lambda_{b,t} \end{bmatrix} \begin{bmatrix} X_{p,t} \\ X_{b,t} \end{bmatrix} \quad (2)$$

where $X_{p,t}$ and $X_{b,t}$ are the state variables describing the nutrient content in the water column ($X_{p,t}$) and in the benthic vegetation ($X_{b,t}$), at time t. The model is forced by the external nutrient loading ($N_{ex,t}$), which can be a measured or estimated time series, this can be a random or a deterministic function. $\boldsymbol{Y}_t = [Y_{p,t} \quad Y_{b,t}]^{\mathrm{T}}$ is the observed state of the system. a_p is a conversion factor from nutrient input to concentrations in the water column, and $b_{p,t}$ and b_b are the proportions of pelagic and benthic biomass, respectively, carried over from one year to the next. For the water column this retention factor $b_{p,t}$ is a time-varying parameter that depends on the benthic vegetation (see below), whereas the parameter b_b is time-invariant. $b_{p,b} f_t$ is the proportion of nutrients from the water column which is used for growth in the benthic region. The nutrient content in the sediments is not modeled explicitly, but is implicitly contained in $b_{p,b} f_t$ and $b_{p,t-1}$. $K f_t$ describes a background level of benthic vegetation, e.g., immigration of seeds from outside the system.

$\{\xi_{j,t}, \lambda_{j,t}\}$ (with $j \in \{p, b\}$) are mutually independent noise terms. For simplicity additive normal distributed noise is commonly considered in the literature, but as we require the state variables to be positive we have chosen multiplicative noise with $P(\text{noise} < 0) = 0$. This ensures that the observations, and the states remain positive at all times. We further require that all the random variables have expectation equal to one and are mutually independent white noise sequences, i.e., sequences of iid. random variables.

The non-linear effects in the model are the functions $b_{p,t} = b_p(X_{b,t})$ and $f_t = f(X_{p,t})$. $b_{p,t}$ is the recycling of nutrient from the sediment to the pelagic region, assuming that it is low when the benthic vegetation is high and high when the benthic vegetation is low. This functional relationship is represented by a sigmoid function

$$b_{p,t} = (1 - \Phi(X_{b,t}))\,\mathrm{d} + t_p = \left(1 - \Phi\left(X_{b,t}; r_p, \tau_p^2\right)\right)\mathrm{d} + t_p \quad (3)$$

DOI: 10.1002/env

where $\Phi(x; r_p, \tau_p^2)$ is the CDF for a Gaussian random variable with mean r_p and variance τ_p^2. This function generates a regime shift and potentially a hysteresis effect. If $\tau_p^2 = 0$ we will refer to this as a hard threshold, otherwise we will refer to this as a smooth threshold. For the limit $\tau_p^2 \to \infty$, $b_{p,t}$ is a constant. The function is constrained between t_p and $t_p + d$.

The function f_t describes the shadowing effect of nutrients in the water column, mainly through particulate organic forms such as phytoplankton and detritus, and is modeled as

$$f_t = 1 - \Phi\left(X_{p,t}; r_b, \tau_b^2\right) \tag{4}$$

resulting in values ranging from no shading at all ($f_t = 1$) to complete shading of the benthic vegetation caused by nutrients in the water column ($f_t = 0$). These are asymptotic values ($X_{p,t} \to \pm\infty$), but if τ_b^2 is small then this will be effective values for $X_{p,t}$ small or large.

A short-hand notation of the state space model is

$$\boldsymbol{X}_t = N_{ex,t}\boldsymbol{a} + \Xi_t(\boldsymbol{B}_{t-1}\boldsymbol{X}_{t-1} + \boldsymbol{k}_{t-1}) \tag{5}$$

$$= N_{ex,t}\boldsymbol{a} + \Xi_t g(\boldsymbol{X}_{t-1}) \tag{6}$$

$$\boldsymbol{Y}_t = \boldsymbol{\Lambda}_t\boldsymbol{X}_t \tag{7}$$

where the matrices are defined by comparison with Equations (1) and (2)

If the noise terms are held constant at one then dynamics of the system are contained entirely in the system equation ($\boldsymbol{X}_t$). Even without stochastic perturbations, model simulations may result in complicated structures.

3. THE DETERMINISTIC MODEL

The purpose of this section is to examine the system in Equation (5), with the noise term held constant (i.e., $V[\Xi_t] = 0$), and the external loading deterministic ($V[N_{ex,t}] = 0$). One important feature of the model is the hysteresis and associated threshold effect. Thus, the model should be able to describe such features with an appropriate parameterization.

The stationary solutions to the model (i.e., points where $\boldsymbol{X}_t = \boldsymbol{X}_{t-1}$) were found for a selected set of parameter values (Table 1) over a wide range of nutrient inputs as forcing (left panel of Figure 1). The stationary solutions clearly illustrate the hysteresis effect with a sudden change in system variables (from a system dominated by benthic vegetation to a system with high pelagic primary production) occurring when the nutrient loading exceeds 0.4. Following this regime shift the nutrient input to the system must be decreased to below 0.25 for the system to return to the original state. The right-hand side of Figure 1 shows a simulated response of the system with a linearly increasing nutrient load followed by a linearly decreasing nutrient load. The system response is more gradual and the hysteresis appears to occur over a wider range of nutrient input, this happens because the system is not allowed to reach the stationary point before the loading is changed. Since constant nutrient loading in ecosystems is a

Table 1. Parameters of the model

Parameter	a_p	$b_{p,b}$	b_b	K	d	t_p	r_p	τ_p	r_b	τ_b
Value	0.8	0.1	0.8	0.2	0.4	0.2	0.4	0.1	0.4	0.1

DOI: 10.1002/env

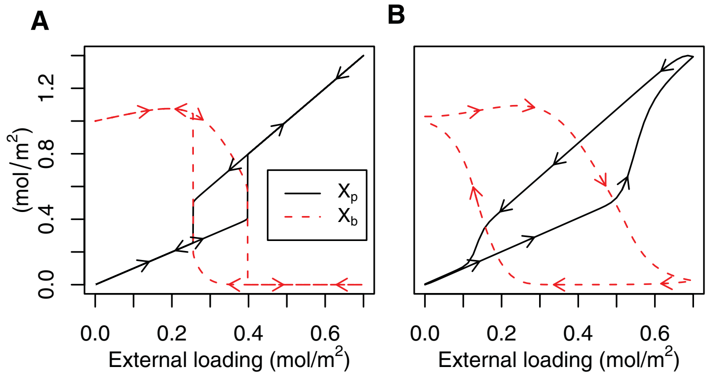

Figure 1. (A) Shows stationary points for the system in Equations (5)–(7) with parameters as in Table 1. (B) Shows the trajectory with a loading increased from 0 to 0.7 over a period of 50 years and then decreased from 0.7 to 0 over a period of 50 years. This figure is available in color online at www.interscience.wiley.com/journal/env

hypothetical case seldom observed in practice, this illustrates some of the difficulties with identifying such systems.

For a nutrient input in the midrange between the two threshold values (i.e., between 0.25 and 0.4) the system will converge towards one of the stability points representing the two different regimes (Figure 2). The border line between the two basins of attractions (the seperatrix) is important for studying an intervention of the system. Figure 2 shows two plots linked to the stability of the processes described above. The left panel of the figure shows the norm of $\boldsymbol{X}_t - \boldsymbol{X}_{t-1}$ for $N_{ex,t} = 0.35$ as a function

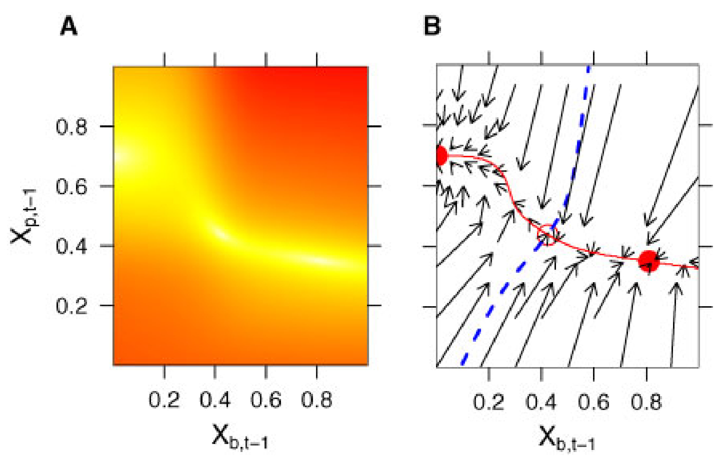

Figure 2. (A) Contour plot of $\|\boldsymbol{X}_t - \boldsymbol{X}_{t-1}\|^2$, with a constant nutrient input, $N_{ex} = 0.35$, and parameters as in Table 1. (B) The corresponding three equilibria; two of these are stable (filled circles) and one is unstable (open circle). The system trajectories towards the equilibrium are indicated by length and direction of the vectors $\boldsymbol{X}_t - \boldsymbol{X}_{t-1}$ (arrows). The solid line indicates the minimum of $\|\boldsymbol{X}_t - \boldsymbol{X}_{t-1}\|^2$ for each X_b, and the dashed line indicates the border of the two basins of attractions (the seperatrix), i.e., sequences starting to the left of this line will converge to the stable equilibrium point indicated with the filled circle to the left and starting points to the right of the dashed line will converge to the filled circle (stable equilibrium point) to the right. This figure is available in color online at www.interscience.wiley.com/journal/env

DOI: 10.1002/env

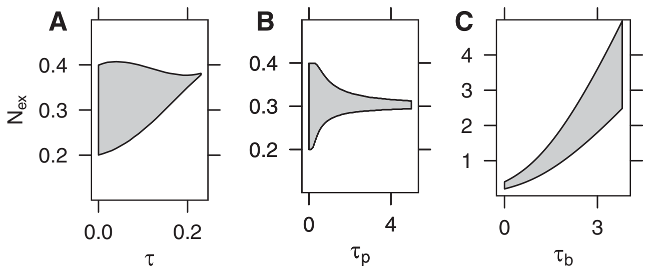

Figure 3. Occurrence of bistability for different nutrient inputs (N_{ex}) and values of the "threshold" parameters: (A) $\tau = \tau_p = \tau_b$, (B) $\tau_b = 0$ and varying τ_p, and (C) $\tau_p = 0$ and varying τ_b. All other parameters are as in Table 1. Note the different scales of the axes

of $\boldsymbol{X}_{t-1}$. The plot shows that there are three equilibria (this is expected from Figure 1). The figure further shows the line of minimum difference connecting these three points. The left panel of the figure shows the direction and length of the same vector, in addition the three equilibrium points are indicated and the blue line indicates the seperatrix. van Nes and Scheffer (2007) propose using the recovery rate from small perturbation as an indicator of how close the system is to a break down, this indicator would be the length of the arrows in Figure 2.

Plots like Figure 2 can help us answer questions about the system. For example, if we are in the left equilibrium point of Figure 2, will it be possible to shift to the other state by removing pelagic biomass from the system? And how much biomass should be removed? A shift from the pelagic dominated state to the benthic dominated state is only feasible provided that $X_b > 0.1$. This may help explain why some bio-manipulation experiments are successful whereas others are not! Another important question that can be addressed with the model is: how much should the loading to the system be changed to cause a regime shift, and how will the transition occur, i.e., will the transition be slow and gradual or more abrupt? Such plots can also illustrate the risk of changing regimes in the more realistic situation when stochastic variation is added to the model.

3.1. Regions of bistability

The region of bistability varies with the parameter values. As stated above, the model will be linear if we set $\tau_b^2 = \tau_p^2 = \infty$ and bistability will not be present. In a regime shift context, this asymptotic behavior is not very informative, since it only explains that there can be no hysteresis effect. It is more important to investigate when bistability actually occurs. Figure 3 shows some combinations of τ's and nutrient load that lead to bistability when all other parameters are held constant. If $\tau_p = \tau_b = \tau$, it is only possible to obtain a bistable system for τ less than about 0.25 (3A). If either τ_p or τ_b is set equal to zero (hard threshold) then a bistable region exists for a larger (possible infinite) range of values for the other threshold parameter (3B,C). It should also be stressed that there are large differences in the bistability regions, depending on which of the τ^2-parameters is set to zero.

4. THE STOCHASTIC MODEL

Models are abstractions of reality simplifying the complexity of ecosystems into a mathematical tool suitable for addressing specific questions, and as such models are bound to be uncertain. To address the

DOI: 10.1002/env

inherent stochastic nature of the system, stochastic models are required. Stochastic perturbations can have large, and sometimes counter-intuitive, impacts on the model behavior. The different uncertainty components described in the introduction propagate differently through the models, and consequently the impacts on model results are different.

The most simple noise component is observation noise, the measurement errors do not propagate through the system equations and therefore the result is random fluctuations added to the deterministic model. System noise on the other hand is mediated through the equations that may lead to complicated correlation structures. Input noise is also filtered through the system equations, but in a more simple way. The difference between input and system noise is discussed in more detail in Appendix A.

4.1. Noisy input

In the following we will let the loading be a random variable, as was the case for the states and the observation, we will also require this to be positive. This is also ensured by using positive multiplicative noise on the input. That is, we set

$$N_{ex,t} = \hat{N}_{ex,t} \cdot u_t \tag{8}$$

where u_t is a white noise process with expectation equal to one and $P(u_t < 0) = 0$, and $\hat{N}_{ex,t}$ is a deterministic function, e.g., an estimated sequence of loadings.

Ludwig *et al*. (2003) used multiplicative $\log t$ distributed noise as input. The $\log t$ distribution does, however, not have a finite expectation, e.g., if $u \sim \log t$ and we would like to compute the expectation of $\boldsymbol{X}_{t|t-1}$ ($\boldsymbol{X}_t$ given $\boldsymbol{X}_{t-1}$), then the expected pelagic nutrient content would be

$$\begin{aligned} E[X_{p,t|t-1}] &= E[N_{ex,t} a_p u_t] + E[\xi_{p,t} b_{p,t-1} X_{p,t-1}] \\ &= N_{ex,t} a_p E[u_t] + b_{p,t-1} X_{p,t-1} E[\xi_{p,t}] \\ &= \infty + b_{p,t-1} X_{p,t-1} = \infty \end{aligned} \tag{9}$$

The full derivation of the non-existence of the expectation of the $\log t$ distribution is given in Appendix B. It could be argued that for all practical purposes (e.g., simulations) this would not be relevant since the probability of extreme events is still small. It is however unfortunate that we are not able to make the most simple inferences about future states. The $\log t$ distribution would further force us to look at quantiles only. For our modeling purposes we want the expectation of the noise terms to equal 1. Consequently, the $\log t$ distribution, as employed by Ludwig *et al*. (2003), cannot be used and the log-normal distribution is considered a better alternative.

Guttal and Jayaprakash (2007) analyze the asymptotic behavior of two one-dimensional bistable ecological models under stochastic forcing. Their analysis shows that the bistable regime is narrowed by the introduction of noise and that for sufficiently large noise there is not two distinct regimes, but a probability distribution for being in one or the other regime. Forcing in biological systems is bound to vary in time and we will focus more on the transient behavior of the process than on the asymptotic behavior. The analysis carried out here will also point to the inclusion of noise in the system equation, this is important when forcing only has direct effect on one level of the system equations.

Simulations with noise on the input show that there is an increasing probability of shifting between regimes as time passes (see Figure 4). Although the simulations start in the "healthy" state, there is no

DOI: 10.1002/env

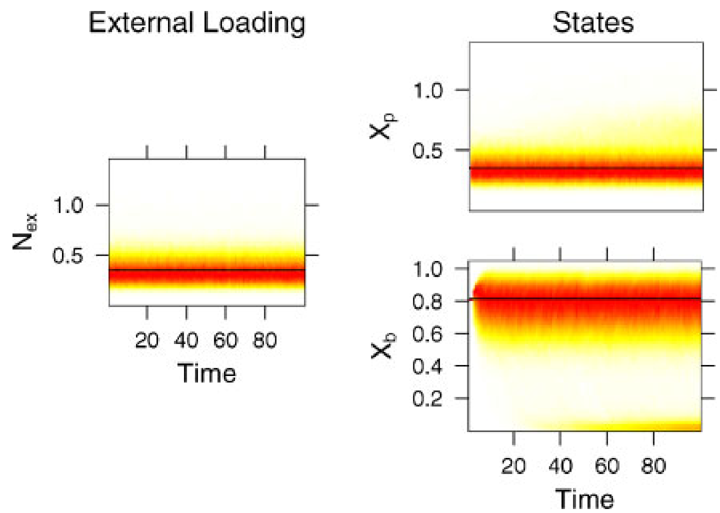

Figure 4. Responses in state variables in a model with input noisy ($V[u_t] = 0.1$) and a constant mean $\hat{N}_{ex,t} = 0.35$, no system noise is added. Solid lines indicate the deterministic response. The noise was simulated by means of a log-normal distribution. The density plots are based on 5000 realizations of the system and are estimated with a Gaussian kernel (the default in "R"). This figure is available in color online at www.interscience.wiley.com/journal/env

guarantee that the system will remain in this state. In fact, the "unhealthy" state is actually absorbing or at least what could be called "effectively absorbing." This means that when the system is in the "unhealthy" state, the probability of returning to the "healthy" state is zero, or at least very close to zero. The onset of this feature is to some extent indicated in Figure 4, but the true nature of the phenomenon reveals itself only after longer simulations (not shown).

We now formally define the two distinct regimes ("healthy" and "unhealthy") for the stochastic model. To do this let $\boldsymbol{X}_{s,\mathrm{I}}$ and $\boldsymbol{X}_{s,\mathrm{II}}$ be the two stable equilibrium points and let $s(X_b)$ be the separator function:

$$s(x_b) = \left\{x_p \in \mathbb{R}_0 : \left\{\boldsymbol{X}_t^{(x_p-\epsilon,x_b)}\right\} \to X_{s,\mathrm{I}}, \left\{\boldsymbol{X}_t^{(x_p+\epsilon,x_b)}\right\} \to X_{s,\mathrm{II}}, \quad t \to \infty, \forall \epsilon > 0\right\} \tag{10}$$

where $\{\boldsymbol{X}_t^{\boldsymbol{y}}\}$ is the deterministic process starting at the point $\boldsymbol{y}$, and $s(x_b)$ is the dashed line in Figure 2. The two regimes are now defined as

$$R_1 = \left\{(x_b, x_p) \in \mathbb{R}_0^2 : s(x_b) > x_p\right\} \tag{11}$$

$$R_2 = \left\{(x_b, x_p) \in \mathbb{R}_0^2 : s(x_b) \leq x_p\right\} \tag{12}$$

Thus, R_1 is the "healthy" regime with low phytoplankton levels and high benthic vegetation, and R_2 is the "unhealthy" regime with high phytoplankton levels and low benthic vegetation.

Figure 4 indicates that regime II could be completely absorbing for $\hat{N}_{ex,t} = 0.35$, i.e., corresponding to $P(X_{b,t} \in R_1, X_{b,t-1} \in R_2) = 0$ and $P(X_{b,t} \in R_2, X_{b,t-1} \in R_2) = 1$. This implies that

DOI: 10.1002/env

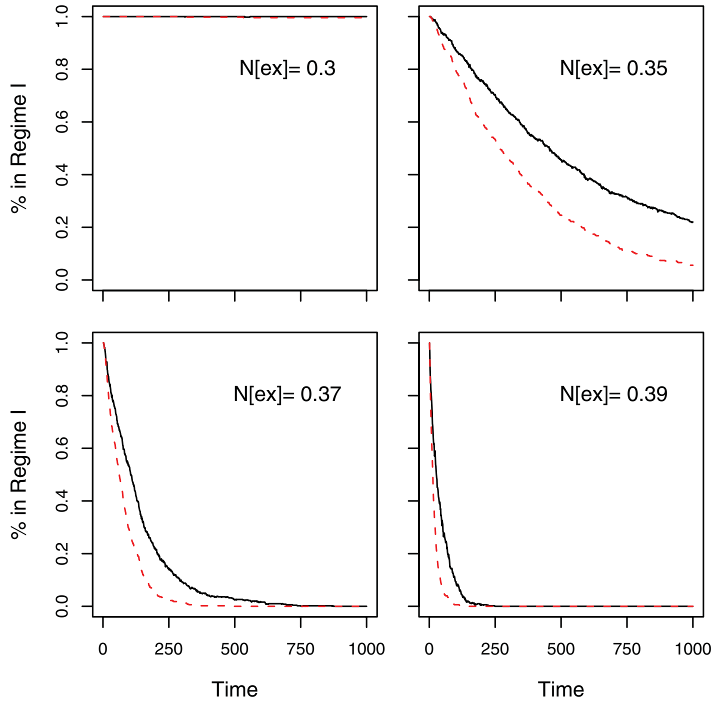

Figure 5. The proportion of sample paths which are in the "healthy" regime (R_1) at time t (solid line), and the proportion of sample paths which have not visited the "unhealthy" regime (R_2) at a time prior to t (dashed line). The plots are based on 500 simulations, of trajectories for $t \in [0, 1000]$. This figure is available in color online at www.interscience.wiley.com/journal/env

$P(X_{b,t} \in R_2) = P(\exists s \in \{1, 2, \ldots, t\} : X_s \in R_2)$, and estimates of these two probabilities are plotted in Figure 5. There are sample paths that visit regime "II" and then return to regime "I." However, it is also clear that regime II is absorbing in the long run for the loading $N_{ex,t} = 0.35$, for higher loadings the first exit time is accelerated and all trajectories end up in regime "II," while for the loading $N_{ex,t} = 0.3$ all trajectories stay in regime "I" for the timespan considered (1000 years). For $N_{ex,t} = 0.35$ the system has not reached the stationary distribution and it is possible that a proportion of the trajectories will remain in regime I in the long run.

In stochastic analysis, focus is primarily on the stationary distributions, but the time before the system settles in the steady state can be very long (Figure 5). Loadings are not constant in time since these are subject to natural variations, changes in land use, and political decisions to mitigate eutrophication. It is therefore relevant to look at the transient behavior of the system, since environmental policies should have an effect within a reasonable time horizon. It is e.g., not relevant for the policy making that a change in loading will lead to a healthy system in 1000 years.

4.1.1. A one step distribution. When the only noise component is on the input we can actually calculate the probability of moving to (or staying in) regime I during the next time step, given the state. Let F_u

DOI: 10.1002/env

be the distribution function of u_t then the probability is

$$P(\boldsymbol{X}_t \in R_2|\boldsymbol{X}_{t-1}) = P(X_{p,t} > s(X_{b,t|t-1})|\boldsymbol{X}_{t-1}) \tag{13}$$

$$= P(u_t\hat{N}_{ex,t}a_p > s(X_{b,t|t-1}) - b_p(X_{b,t-1})X_{p,t-1}) \tag{14}$$

$$= P\left(u_t > \frac{s(X_{b,t|t-1}) - b_p(X_{b,t})X_{p,t-1}}{\hat{N}_{ex,t}a_p}\right) \tag{15}$$

$$= 1 - F_u\left(\frac{s(X_{b,t|t-1}) - b_p(X_{b,t-1})X_{p,t-1}}{\hat{N}_{ex,t}a_p}\right) \tag{16}$$

$$= 1 - F_u(g(\boldsymbol{X}_{t-1})) \tag{17}$$

where $X_{b,t|t-1}$ is a deterministic function of $\boldsymbol{X}_{t-1}$ and the distribution of u is given. $s(X_{b,t|t-1})$ is determined by numerical methods. By solving the equation $g(\boldsymbol{X}_{t-1}) = F_u^{-1}(1-p)$ we can find the functions that describe the probability of moving to regime II during the next time step (this was also

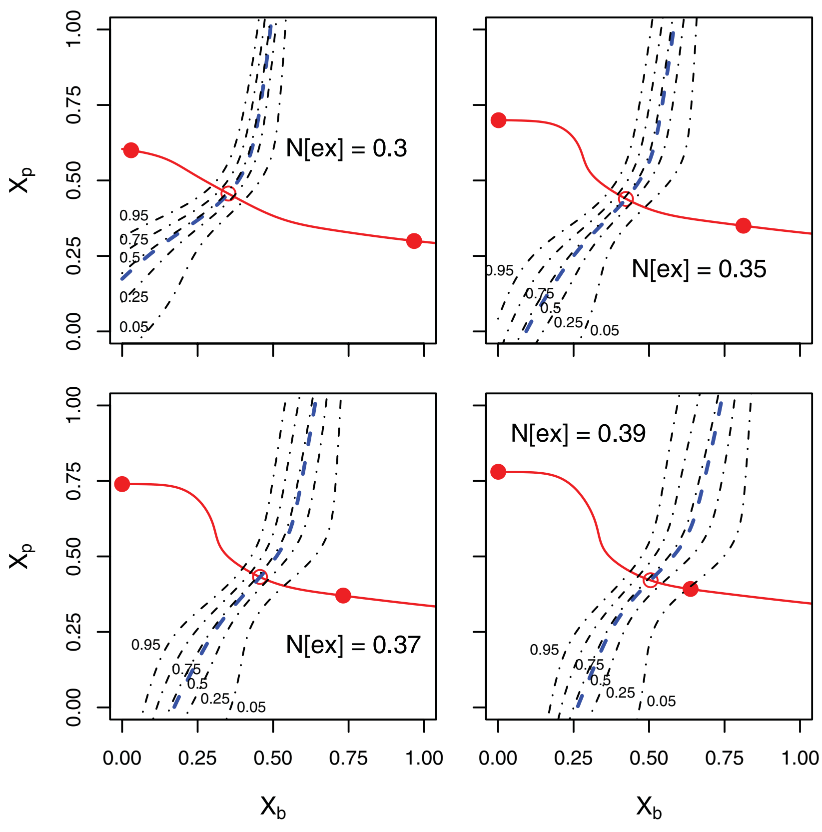

Figure 6. Probability of being in regime II in the next time step as a function of the state variables with four different loadings subject to input noise ($V[u_t] = 0.1$ and $u_t \sim LN$), no system noise. This figure is available in color online at www.interscience.wiley.com/journal/env

DOI: 10.1002/env

done by numerical methods). Solutions of this are shown for a number of different p values and for different loadings in Figure 6. This figure should be compared with Figure 5, and it is seen that even though the probability of moving from $\mathbf{X}_{S,I}$ to regime II in one step is very small ($<< 5\%$), it can still behave like an absorbing state, as is the case for e.g., $N_{ex,t} = 0.37$. The probability of leaving regime I given that the process starts in $\boldsymbol{X}_{S,\mathrm{I}}$ and the loading is 0.3 is very close to zero. For a loading of 0.35 we see that the trajectories are slowly leaving regime I, and further we see that most of the trajectories that have left regime I stay in regime II.

From the plots in Figures 5 and 6 we can conclude that regime II is what we can call effectively absorbing. It is however important to stress that the analysis so far only included noise in the loading, this noise has a direct effect on the pelagic level and only an indirect effect on the benthic level.

4.2. System noise

As discussed earlier, system noise will propagate through the system in a much more complicated way than input noise. The multiplicative noise at time t will be multiplied with system noise at all prior times (see Appendix A). As seen in the previous section, input noise is able to drive the system, in the sense that the system with input noise is very different from a system that fluctuates around the deterministic system (corresponding to observational noise only). As we will see adding noise in different levels in the system will also have very different implications. Noise in the pelagic level propagates through the system in a similar way as input noise, while benthic noise alters the system in a quite different way.

Figure 7 shows a simulation of the model with system and input noise, it is seen that the decay to regime II is much faster than for a system with only input noise (Figure 4). Further it is seen that the spread in the state variables is larger in this system. Quantifying the deviation between two stochastic systems is much more involved than quantifying the deviation between two deterministic systems. In

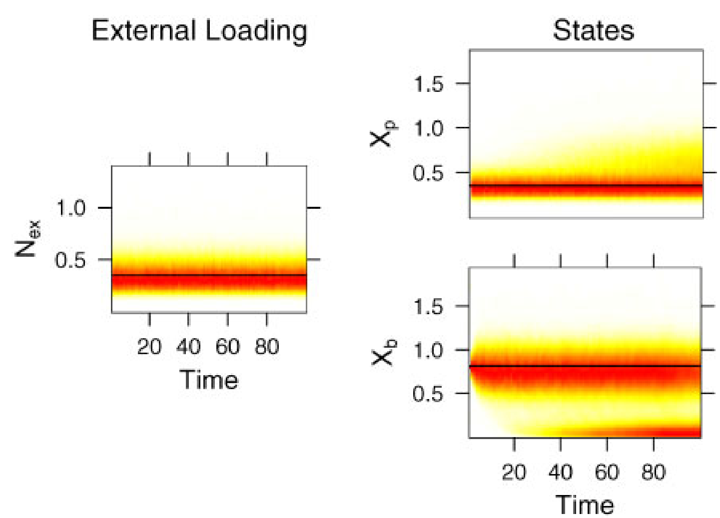

Figure 7. System with input noise and system noise. The solid lines are the deterministic response. Noise was simulated by a log-normal distribution with $V[u_t] = 0.1$, $V[\xi_p] = 0.05$, $V[\xi_b] = 0.01$ and $\hat{N}_{ex,t} = 0.35$ for all t. The density plots are based on 5000 realizations of the system and estimated with a Gaussian kernel (the default in "R"). This figure is available in color online at www.interscience.wiley.com/journal/env

DOI: 10.1002/env

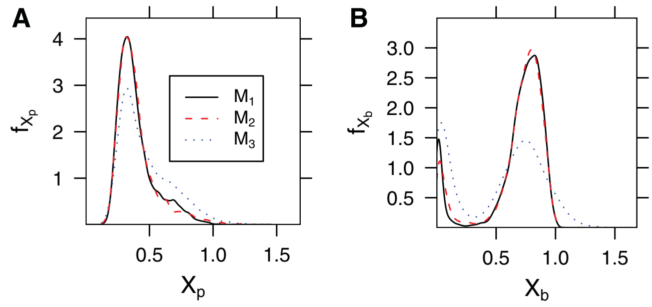

Figure 8. Estimated conditional densities after 10 years, all processes start at the point $\boldsymbol{X}_{S,\mathrm{I}}$ with a loading of $N_{ex,t} = 0.35$. M_1 is a model with only noise in the input ($V[u_t] = 0.1$), M_2 is the density with noise in the input and in the pelagic level ($V[\xi_p] = 0.05$). Finally M_3 is the density for a model with noise in the input, the pelagic level and in the benthic level ($V[\xi_b] = 0.01$). This figure is available in color online at www.interscience.wiley.com/journal/env

deterministic systems it is enough to follow the trajectory of the state variables and measure how fast these diverge. For stochastic systems we are actually interested in the divergence distribution, i.e., how the distribution evolves.

Chan and Tong (2001) propose an invariant measure for the divergence of distributions. However, it seems that this requires a very large number of simulations. We will therefore focus on differences in the first exit-times and the mean value and variances of trajectories and note that these can actually be equal for different distributions. It is clear, however, that two distributions with different moments or exit-times are indeed different.

Figure 8 shows the estimated densities for models with noise in different levels of the model, it is seen that adding noise in the pelagic level does not change the conditional distribution much compared to the model with noise only in the input. In contrast adding noise in the benthic level changes the model significantly, with the probability mass being much more evenly distributed over the range of possible values. It is worth noting that this happens even though the variance of the benthic noise is small compared to the noise in the other levels.

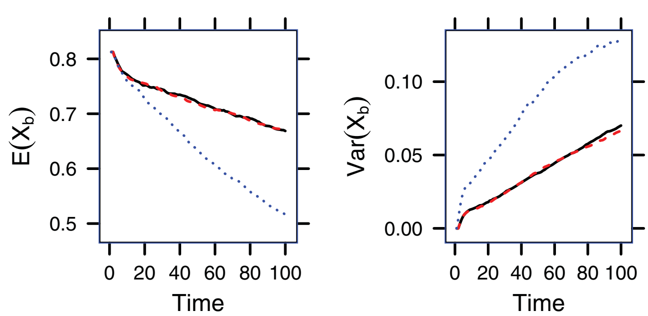

Figure 9. Estimated conditional (on $\boldsymbol{X}_0$) mean values and variances as a function of time, all processes start i $\boldsymbol{X}_{S,\mathrm{I}}$ with a loading of $N_{ex,t} = 0.35$. The solid line is a model with only noise in the input ($V[u] = 0.1$) the dashed line is the mean/variance with noise in the input and in the pelagic level ($V[\xi_p] = 0.05$), and the dotted line is the mean/variance for a model with noise in the input, the pelagic level and in the benthic level ($V[\xi_b] = 0.01$). This figure is available in color online at www.interscience.wiley.com/journal/env

DOI: 10.1002/env

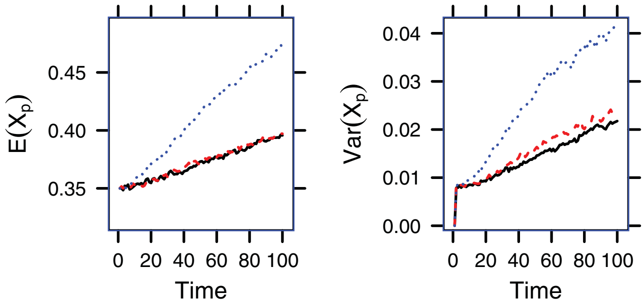

Figure 10. Estimated conditional (on X_0) mean values and variances as a function of time, all processes start at $X_{S,I}$ with a loading of $N_{ex,t} = 0.35$ for the deterministic model. The solid line is a model with only noise in the input ($V[u] = 0.1$), the dashed line is the mean/variance with noise in the input and in the pelagic level ($V[\xi_p] = 0.05$) and the dotted line is the mean/variance for a model with noise in the input, the pelagic level and in the benthic level ($V[\xi_b] = 0.01$). This figure is available in color online at www.interscience.wiley.com/journal/env

Figures 9 and 10 show the mean value and variance for X_b, respectively X_p, as a function of time. It is seen that both the mean value and the variances diverge for the model with noise in all levels, while the model with only noise in the input and the model with noise in the input and in the pelagic level are close at all considered timepoints. The conclusion of this plot is therefore the same as for Figure 8.

Figure 11 shows the time before half of the trajectories have left the healthy region given that $X_0 = X_{S,I}$. The conclusion from this plot is essentially the same as for the mean and variance plots, namely that input noise and pelagic noise propagate through the system in very similar ways, while noise in the benthic level changes the system in a quite different way.

Even though moments and exit-times do not give the full information on the system, it seems that we will be able to discriminate the distributions on the basis of these statistics.

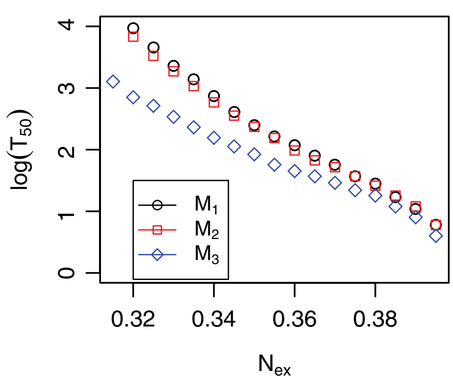

Figure 11. Estimated first exit-time as a function of loading, all processes start at $X_0 = X_{S,I}$ with loading of $N_{ex,t} = 0.35$ for the deterministic model. M_1 is the model with only noise in the input ($V[u] = 0.1$), M_2 is the exit time with noise in the input and in the pelagic level ($V[\xi_p] = 0.05$), and M_3 is the exit time for a model with noise in the input, the pelagic level and in the benthic level ($V[\xi_b] = 0.01$). This figure is available in color online at www.interscience.wiley.com/journal/env

DOI: 10.1002/env

5. CONCLUSION AND DISCUSSION

We have presented a non-linear two state stochastic model for studying the interaction between pelagic and benthic production as a function of nutrient loading. It has been shown how this model can give rise to bistability and hysteresis effects, the regions of bistability are shown to comprise a small subset of the parameter space. Since there will be uncertainties in the parameter estimates, a consequence of this observation is that it is hard to prove the existence of hysteresis or bistability, as it has to be documented that the parameters belong to the region of bistability for the model.

A main topic in the analysis is the propagation of noise. It is illustrated that noise can be an important driver for the dynamics in the system. Guttal and Jayaprakash (2007) analyze the asymptotic behavior of systems under stochastic forcing, their findings are similar to ours, namely that break down of systems can be accelerated by noise. It is therefore important to consider noise within the system rather than only analyzing the deterministic skeleton of the system and perceive noise as a random fluctuation around the deterministic trajectories. We analyze a two-dimensional system rather than a one-dimensional system and distinction between system noise and input noise is then more important. The focus in this analysis is more on densities in the transient than the asymptotic behavior of the systems.

In order to analyze a stochastic model we need to look at the probabilistic properties of the model and the propagation of densities rather than the trajectories of the deterministic model. Comparing densities is, however, not a simple task, and this will typically require a very large number of simulations, hence we suggest comparing simpler statistics such as the mean and the variance of the (time-varying) densities.

In the setting of aquatic system, these points imply that small fluctuations in the pelagic production may increase the probability of a system collapse towards complete dominance of pelagic production. Once the system has collapsed, it may be very difficult to return the system to the original state. It is therefore important to include stochastic components in analysis of such systems and manage aquatic ecosystems on the basis of a precautionary principle.

Carpenter and Brock (2006) and van Nes and Scheffer (2007) advocate two different approaches to identifying regime-shifts before they actually occur, another approach is to estimate parameters in models like the one presented in this article and the base predictions of densities on simulations studies.

ACKNOWLEDGMENTS

This study is a contribution to the EU-funded project Thresholds (GOCE-003933). The study was supported by the Danish Graduate School in Biostatistics.

REFERENCES

Carpenter SR, Ludwig D, Brock WA. 1999. Management of eutrophication for lakes subject *t* potentially irreversible changes. *Ecological Applications* **9**(3): 751–771.

Carpenter SR. 2005. Eutrophication of aquatic ecosystems: bistability an soil phosphorus. *PNAS* **102**(29): 10002–10005.

Carpenter SR, Brock WA. 2005. Rising variance: a leading indicator of ecological transition. *Ecology Letters* **9**: 311–318.

Chan K-S, Tong H. 2001. *Chaos: A Statistical Perspective*. Springer-Verlag, New York, Inc.

Guttal V, Jayaprakash C. 2007. Impact of noise on bistable ecological systems. *Ecological Moddeling* **201**(3): 420–428.

Holling CS. 1973. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics* **4**: 1–23.

Ludwig D, Carpenter SR, Brock WA. 2003. Optimal phosphorus loading for a potentially eutrophic lake. *Ecological Applications* **13**(4): 1135–1152.

May RM. 1977. Thresholds and breakpoints in ecosystems with multiplicity of stable states. *Nature* **269**: 471–477.

van Nes EH, Scheffer M. 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *The American Naturalist* **169**(6): 738–747.

DOI: 10.1002/env

Scheffer M, Carpenter S, Foley JA, Folke C, Walker B. 2001. Catastrophic shifts in ecosystems. *Nature* **413**: 591–596.
Scheffer M, Szabó S,van Nes EH, Rinaldi S, Kautsky N, Norberg J, Roijackers RMM, Franken RJM. 2003. Floating plant dominance as a stable state. *PNAS* **100**(7): 4040–4045.
Schindler DW. 2006. Recent advances in the understanding and management of eutrophication. *Limnology and Oceanography* **51**(1): 356–363.

APPENDIX A: DERIVATION OF NOISE FILTRATION

The purpose of this section is to derive theoretical expressions of how noise at different levels of the state space model is filtered through the system. The state space model is given (as in Section 2) by

$$N_{ex,t} = \hat{N}_{ex,t} \cdot u_t \tag{18}$$

$$\boldsymbol{X}_t = N_{ex,t}\boldsymbol{a} + \Xi_t(\boldsymbol{B}_{t-1}\boldsymbol{X}_{t-1} + \boldsymbol{k}_{t-1}) \tag{19}$$

$$= N_{ex,t}\boldsymbol{a} + \Xi_t g(\boldsymbol{X}_{t-1}) \tag{20}$$

$$\boldsymbol{Y}_t = \boldsymbol{\Lambda}_t\boldsymbol{X}_t \tag{21}$$

The aim here is to give an expression of $\boldsymbol{X}_t$ given $\boldsymbol{X}_0$ ($\boldsymbol{X}_{t|0}$) and a sequence of predicted loadings $\{\hat{N}_{ex,i}\}_{i=1}^{t}$. To do this we use induction of $\hat{\boldsymbol{X}}_t$. The result is

$$\boldsymbol{X}_{t|0} = \boldsymbol{a}\hat{N}_{ex,t}u_t\Xi_t g(\boldsymbol{X}_{t-1|0}) \tag{22}$$

$$= \boldsymbol{a}\hat{N}_{ex,t}u_t\Xi_t(B_{t-1}\boldsymbol{X}_{t-1|0} + \boldsymbol{k}_{t-1}) \tag{23}$$

$$= \boldsymbol{a}\hat{N}_{t,ex}u_t\Xi_t\boldsymbol{k}_{t-1} + \Xi_t\boldsymbol{B}_{t-1}(\boldsymbol{a}N_{ex,t-1}u_{t-1}\Xi_{t-1}(\boldsymbol{B}_{t-2}\boldsymbol{X}_{t-2|0}\boldsymbol{k}_{t-2})) \tag{24}$$

$$= \boldsymbol{a}\hat{N}_{ex,t}u_t\Xi_t\boldsymbol{B}_{t-1}\boldsymbol{a}\hat{N}_{ex,t-1}u_{t-1}\Xi_t\boldsymbol{k}_{t-1} + \Xi_t\boldsymbol{B}_{t-1}\Xi_{t-1}\boldsymbol{k}_{t-2} + \Xi_t\boldsymbol{B}_{t-1}\Xi_{t-1}\boldsymbol{B}_{t-2}\boldsymbol{X}_{t-2|0} \tag{25}$$

$$\vdots$$

$$= \sum_{i=0}^{t}\Big(\prod_{j=i}^{t-1}\Xi_{j+1}\boldsymbol{B}_j\Big)\hat{N}_{ex,i}u_i\boldsymbol{a} + \Xi_t\sum_{i=1}^{t}\Big(\prod_{j=i}^{t-1}\Xi_j\boldsymbol{B}_j\Big)\boldsymbol{k}_{i-1} + \Big(\prod_{i=1}^{t}\Xi_i\boldsymbol{B}_{i-1}\Big)\boldsymbol{x}_0 \tag{26}$$

where $\prod_{i=k}^{l} A_i = I$ for $k > l$. With this derivation it is seen that noise at different levels will filter differently through the system. The most simple noise i.e., the observation noise, does not influence the dynamics of the system. On the contrary noise in the input is filtered by products of $\boldsymbol{B}_s$'s which are themselves non-linear functions of the state variables. Finally system noise will filter with products of the $\boldsymbol{B}_s$'s and itself. Even if the system is linear ($\boldsymbol{B}_s = \boldsymbol{B}$ and $\boldsymbol{k}_s = \boldsymbol{k}$ for all s), then the complexity introduced by the multiplicative noise would make direct derivation of expectations and variances very complicated. We will therefore rely on the simulation studies presented in this article.

APPENDIX B: THE LOG *t*-DISTRIBUTION

In this section we will present the derivations needed to prove that the log t-distribution does not have an expectation. The pdf of a t-distributed random variable with $\nu > 0$ degrees of freedom is

$$f(x) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \tag{27}$$

Now if X follows a t-distribution then $Y = e^X$ follows a log t-distributed and we can calculate the expectation of Y as

$$E[Y] = \int_{-\infty}^{\infty} y(x) f(x)\,\mathrm{d}x \tag{28}$$

$$= \int_{-\infty}^{\infty} \frac{e^x\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \mathrm{d}x \tag{29}$$

$$= \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\Gamma(\nu/2)}\int_{-\infty}^{\infty} \frac{e^x}{\left(1+x^2/\nu\right)^{(\nu+1)/2}}\,\mathrm{d}x \tag{30}$$

The denominator under the integral is a polynomial of degree $\nu + 1$ while the enumerator is a polynomial of degree ∞ and we can conclude that $E[Y] = \infty$.

DOI: 10.1002/env