ECOLOGY LETTERS

Ecology Letters, (2017) doi: 10.1111/ele.12876

LETTER

Probing the limits of predictability: data assimilation of chaotic dynamics in complex food webs

Elias C. Massoud,[1*] Jef Huisman,[2] Elisa Benincà,[3] Michael C. Dietze,[4] Willem Bouten,[2, 5] and Jasper A. Vrugt,[1, 6]

Abstract

The daunting complexity of ecosystems has led ecologists to use mathematical modelling to gain understanding of ecological relationships, processes and dynamics. In pursuit of mathematical tractability, these models use simplified descriptions of key patterns, processes and relationships observed in nature. In contrast, ecological data are often complex, scale-dependent, space-time correlated, and governed by nonlinear relations between organisms and their environment. This disparity in complexity between ecosystem models and data has created a large gap in ecology between model and data-driven approaches. Here, we explore data assimilation (DA) with the Ensemble Kalman filter to fuse a two-predator-two-prey model with abundance data from a 2600+ day experiment of a plankton community. We analyse how frequently we must assimilate measured abundances to predict accurately population dynamics, and benchmark our population model's forecast horizon against a simple null model. Results demonstrate that DA enhances the predictability and forecast horizon of complex community dynamics.

Keywords

Data assimilation, ecological models, ecosystems, food webs, forecast horizons, plankton, predator-prey.

Ecology Letters (2017)

INTRODUCTION

Ecosystems constitute a complex network of living organisms, which are interconnected and linked with their environment through a myriad of nutrient cycles, mass and energy flows. This may give rise to emergent and self-organised behaviour, where organisms exhibit complex spatial and temporal patterns. Faced with this daunting complexity, ecologists have developed mathematical models to gain understanding of ecosystem functioning, to simulate large-scale experiments, and to predict ecological processes into the future. However, most ecological models may explain only a small fraction of the dynamics of the actual ecosystem (Wintle *et al.* 2003). What is more, due to process abstraction and spatial aggregation, ecological model parameters often do not represent directly measurable ecosystem quantities and must therefore be estimated indirectly through calibration using measurements of ecosystem inputs and outputs. The 'calibrated' model can then be used to predict ecological processes over longer periods of time (Fasham *et al.* 1990; Sitch *et al.* 2003).

During the past few decades increasingly larger volumes of ecological data have become available in response to continued advances in measurement techniques and the rapid expansion of long-term monitoring networks (Running *et al.* 1999; Aanensen *et al.* 2009; LaDeau *et al.* 2017). This ever increasing wealth of data provides unique opportunities for ecologists to enhance ecosystem understanding and characterisation (Ter Braak & Van Tongeren 1995; Reichman *et al.* 2011). Yet, ecological data are often complex, high-dimensional, and scale-dependent as governed by local interactions and feedback loops between organisms individually and their environment, as well as both predictable (e.g. periodic) and stochastic changes to the ecosystem (Conway *et al.* 1970; Benincà *et al.* 2009). An important example of complex dynamics is provided by Benincà *et al.* (2008) who offered the first long-term experimental demonstration of chaotic population dynamics in a complex food web, with species abundances that showed striking fluctuations over several orders of magnitude. This study raises important questions about our understanding of the predictability of ecological processes. First, given the few examples of chaotic dynamics in complex systems, how often must we census populations to accurately describe and forecast their dynamics? Benincà *et al.* (2008) relied on twice-weekly sampling for a period of eight years, a level of effort difficult to replicate in other systems. Would chaos be detectable with a lower sampling frequency? Are there general insights about sampling frequency that could be extended from this data set to other systems? Second, the predictability of this system was previously assessed using mechanism-free

[1]*Department of Civil and Environmental Engineering, University of California Irvine, Irvine, CA 92697-1075, USA*

[2]*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*

[3]*Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands*

[4]*Department of Earth and Environment, Boston University, Boston, MA 02215, USA*

[5]*Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands*

[6]*Department of Earth System Science, University of California Irvine, Irvine, CA 92697-1075, USA*

**Correspondence: E-mail: emassoud@uci.edu*

neural networks (Benincà *et al.* 2008) and wavelets (Benincà *et al.* 2009). Does the horizon over which we can predict chaotic dynamics change using mechanistic models? Can we improve our inference about vital rates by fitting population models directly to chaotic data?

Answers to these questions rely on our ability to fit models to data. However, traditional curve fitting methods have trouble differentiating the impacts of multiple sources of uncertainty (initial conditions, parameters, process variability, observation error) (Clark & Bjornstad 2004). This is particularly important for chaotic dynamics, which by definition are highly sensitive to variability and where long-term prediction is fundamentally impossible. In this paper, we advocate use of advanced statistical methods to improve treatment of model and measurement errors, and reconcile ecological models with data (Kendall *et al.* 1999; Peng *et al.* 2011; Dietze *et al.* 2013). This includes state-of-the-art parameter estimation (Conroy *et al.* 1995; Ali *et al.* 2016) and data assimilation (DA) methods (Luo *et al.* 2011; Luo & Schimel 2011). In contrast to parameter estimation, DA uses observations of system behaviour to statistically constrain the state variables (e.g. population size), rather than the parameters, within a model of a system. The archetype of this method, the Kalman filter (KF), was developed by Kalman (1960) for optimal control of dynamical systems. DA holds great promise in ecology as it will help close the gap between ecosystems models and data, enhance ecological forecasting, partition measurement and model errors, and provide guidance on 'optimal' experimental design (McMahon *et al.* 2009; Dietze 2017b). For example, via integration with inferential ecosystem models, DA can help further refine wireless sensor networks by weighing the value of an observation against cost of data collection (Clark *et al.* 2001).

Data assimilation has found widespread application and use in many fields of research, including oceanography (Bertino *et al.* 2003; Gehlen *et al.* 2015), marine ecology (Lawson *et al.* 1996; Natvik *et al.* 2001; Dowd 2007; Xiao & Friedrichs 2014), hydrology (Vrugt *et al.* 2005), glaciology (Granzow 2014), and satellite remote sensing (Dorigo *et al.* 2007), to name a few. Furthermore, DA has received operational status in real-time weather, traffic, tsunami and flood prediction systems because of its proven ability to enhance significantly the forecast skill of dynamic system models. In ecology, the interest in DA has grown rapidly during the past decade (Williams *et al.* 2005; Chen *et al.* 2008; Mo *et al.* 2008; Quaife *et al.* 2008), to support ecological analysis (Zobitz *et al.* 2011), to account for model structural, input and output errors (Luo *et al.* 2011), to improve ecological prediction (Niu *et al.* 2014), to shed new light on model structural errors and provide guidance on model improvement and data informativeness (Sitz *et al.* 2002; Vrugt *et al.* 2005; Keenan *et al.* 2011). In fact, next generation ecological models are developed conscientiously in anticipation of DA applications (Williams *et al.* 2009; Wu *et al.* 2009; Peng *et al.* 2011; Xu *et al.* 2012).

To date, few authors have used DA to analyse predator-prey dynamics in complex microbial food webs. Lawson *et al.* (1995) developed an adjoint DA method for a simple predator-prey model but assimilated simulated data, and Vallino (2000) used DA to constrain organic matter production and consumption in a compartment-type food web model. Here, we present the first application of DA to long-term chaotic population dynamics (Benincà *et al.* 2008). We then investigate how frequently we must assimilate measured abundances to accurately describe population dynamics and evaluate the ecological forecast horizon (EFH) (Petchey *et al.* 2015) of our population model to determine how far into the future useful forecasts can be made (Simmons & Hollingsworth 2002). We benchmark our results against a constant forecast to quantify the added skill of the two-predator-two-prey model.

MATERIALS AND METHODS

This section presents our application of DA to provide guidelines for researchers new to model-data fusion. We first review our food web data, then discuss the Vandermeer (VD) population model (Vandermeer 2004), and then describe the parameter estimation and DA methods used to close the gap between observed and simulated abundances. We conclude with a description of the EFH (Petchey *et al.* 2015) to evaluate DA with different VD model parameterisations.

The data set

We use a microbial plankton community isolated from the Darss-Zingst estuary in the southern Baltic Sea. The structure of this food web is depicted graphically in Benincà *et al.* (2008) and consists of bacteria, several phytoplankton species, herbivorous and predatory zooplankton species, and detritivores. The plankton community was cultured in a laboratory mesocosm under constant external conditions and sampled twice weekly for a period of more than 8 years to count population abundances of the functional groups. A detailed description of the mesocosm experiment appears in the Supplementary information (SI) of Benincà *et al.* (2008) and data preprocessing is discussed in Benincà *et al.* (2009). The final data set comprises 2656 days and consist of $n = 794$ observations of each species' population count (in $(\text{mg fwt L}^{-1})^{1/4}$) with constant measurement interval of $\Delta t = 3.35$ days. The fourth-root power transformation homogenises the variance of each species population counts by suppressing sharp ups and downs.

In keeping with Benincà *et al.* (2008), we focus our attention to the rotifers, calanoid copepods, picocyanobacteria and nanoflagellates. These $k = 4$ functional groups have a relatively large presence in our food web, and their abundances govern much of the population dynamics with oscillations typical for coupled predator–prey interactions (Benincà *et al.* 2009). The measured abundances are stored in a $k \times n$ matrix, $\tilde{\mathbf{x}}$, where $\tilde{}$ signifies observed data.

Coupled predator-prey model

We analyse the dynamics of the two-predator-two-prey communities, using the population model of Vandermeer (1982, 2004). This VD model assumes that the prey species interact through Lotka–Volterra competition, and are consumed by predators according to a saturating functional response. If P_1

and P_2 denote the abundances of the nanophytoplankton and picophytoplankton preys, and Z_1 and Z_2 the abundances of the calanoid copepods and rotifers (competing predators), respectively, then the VD model is given by

$$\begin{aligned}\frac{\mathrm{d}P_1}{\mathrm{d}t}&=r_1P_1\left(1-\frac{1}{K}(P_1+\alpha P_2)\right)-\left(\frac{gP_1Z_1}{H+(P_1+\beta P_2)}+\frac{g\beta P_1Z_2}{H+(\beta P_1+P_2)}\right)\\ \frac{\mathrm{d}P_2}{\mathrm{d}t}&=r_2P_2\left(1-\frac{1}{K}(\alpha P_1+P_2)\right)-\left(\frac{g\beta P_2Z_1}{H+(P_1+\beta P_2)}+\frac{gP_2Z_2}{H+(\beta P_1+P_2)}\right)\\ \frac{\mathrm{d}Z_1}{\mathrm{d}t}&=\frac{g(P_1+\beta P_2)Z_1}{H+(P_1+\beta P_2)}-mZ_1\\ \frac{\mathrm{d}Z_2}{\mathrm{d}t}&=\frac{g(P_2+\beta P_1)Z_2}{H+(\beta P_1+P_2)}-mZ_2.\end{aligned} \tag{1}$$

This system of four ordinary differential equations describes the coupled interactions of our two competing predators and two prey species under constant environmental conditions. The VD model has four state variables, namely $\mathbf{x}=\{Z_1, Z_2, P_1, P_2\}$ and $d = 8$ parameters, whose values we store in vector $\boldsymbol{\theta}$. This includes the two unitless coefficients $\alpha \geq 0$ and $\beta \in [0, 1]$ which characterise the competition between the zooplankton and their predator selectivity, respectively, the growth rates of the first and second prey, r_1 (day^{-1}) and r_2 (day^{-1}), respectively, the grazing rate, g (day^{-1}), the carrying capacity, K ((mg fwt L^{-1})$^{1/4}$), the mortality rate, m (day^{-1}) and functional response, H ((mg fwt L^{-1})$^{1/4}$).

We write the VD model as a function

$$\mathbf{X} = \mathcal{VD}(\boldsymbol{\theta}, \mathbf{x}_0), \tag{2}$$

which uses as input the d-vector $\boldsymbol{\theta} = \{\alpha, \beta, r_1, r_2, g, K, m, H\}$ of VD model parameters, and a $k \times 1$ vector $\tilde{\mathbf{x}}_0$ of initial population counts (at $t = 0$), and returns as output a $k \times n$ matrix, $\mathbf{X}$ of simulated abundances of the two predators and two preys, respectively, during the 2656 days experimental period. Bifurcation analyses of $\mathbf{X}$ illustrates a range of different dynamics, including stable equilibria, limit cycles and chaos, depending on the values of the VD parameters (Figs S1–S3). In particular, chaotic dynamics is quite ubiquitous in the VD model, resulting in complex patterns of synchronous and asynchronous fluctuations of the four species. Vandermeer (1982) and Vandermeer (2004) provide a detailed explanation of the VD model, and Benincà *et al.* (2009) analyzes the coupled predator-prey cycles in our food web.

In this study, we examine three different VD parameterisations. The first originates from wavelet analysis by Benincà *et al.* (2009) and is listed in Table 1 under 'WAVE'. The second parameterisation (SODA) is obtained from the measured data using joint parameter and state estimation with SODA (Vrugt *et al.* 2005). The third parameterisation, coined INTEL, is also based on the SODA method but restricts state estimation to those 20% of the observations with largest population fluctuations. The underlying methodology is described in the next sections.

Bayesian inference of the VD model parameters

In recent decades, Bayesian inference has emerged as a working paradigm for modern probability theory, parameter and state estimation, model selection and hypothesis testing (Vrugt 2016). Bayesian inference allows for an exact description of parameter uncertainty (and other sources of uncertainty) by treating the parameters as probabilistic variables with joint posterior pdf, $p(\boldsymbol{\theta}|\tilde{\mathbf{X}})$. According to Bayes' theorem, the posterior parameter distribution depends upon the prior distribution, $p(\boldsymbol{\theta})$, which captures our initial beliefs about the values of the model parameters, and a likelihood function, $L(\boldsymbol{\theta}|\tilde{\mathbf{X}})$, which quantifies the confidence in the parameter values, $\boldsymbol{\theta}$, in light of the observed data, $\tilde{\mathbf{X}}$. In the absence of detailed knowledge about the parameter values of our food web, we assume a uniform prior distribution, $p(\boldsymbol{\theta})$, with ranges listed in Table 1. If we further assume the measurement errors of the population counts to be independent, zero-mean normally distributed, $\mathcal{N}(0, \hat{\sigma}_{\mathbf{v}}^2)$, with constant variance, $\hat{\sigma}_{\mathbf{v}}^2$, then the likelihood function becomes

$$L(\boldsymbol{\theta}|\tilde{\mathbf{X}}, \mathbf{x}_0, \hat{\sigma}_{\mathbf{v}}^2) = \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi\hat{\sigma}_{\mathbf{v}}^2}}\exp\left[-\frac{1}{2}\hat{\sigma}_{\mathbf{v}}^{-2}\sum_{j=1}^{k}\sum_{t=1}^{n}\left(\tilde{x}_{jt}-x_{jt}(\boldsymbol{\theta},\mathbf{x}_0)\right)^2\right], \tag{3}$$

where $\tilde{x}_{jt}$ and $x_{jt}(\boldsymbol{\theta}, \mathbf{x}_0)$ signify the observed and predicted abundances of the jth species at time t. The summation term inside the exponent is equivalent to the sum of squared residuals commonly used in model fitting. The initial state, $\mathbf{x}_0$, is set equal to the measured population counts of the two predators and two preys at $t = 0$.

We generate samples of the posterior parameter distribution, using Markov Chain Monte Carlo (MCMC) simulation with the differential evolution adaptive metropolis (DREAM) algorithm. A detailed description of DREAM appears in Vrugt (2016). We use ten different Markov chains with starting points drawn from the parameter ranges listed in Table 1. Convergence of the sampled chains to $P(\boldsymbol{\theta}|\tilde{\mathbf{X}})$ is monitored using the $\hat{R}$-convergence diagnostic Gelman & Rubin (1992). This diagnostic compares the within-chain and between-chain variances of the VD parameters. Convergence is achieved when $\hat{R}_j < 1.2$ for all parameters, $j = \{1, \ldots, d\}$.

Simultaneous parameter and state estimation

The initial state of our food web may be known accurately for our mesocosm data set, but the assumption of a perfect model cannot be justified. Indeed, the VD model is a highly simplified description of the population dynamics in our food web, and consequently, it may not be able to mimic accurately the observed abundances. One should therefore not expect the residuals to satisfy assumptions of normality and independence, but instead exhibit considerable variation in bias, variance, and serial correlation under different population counts. We therefore consider the SODA method of Vrugt *et al.* (2005), which combines state and parameter estimation. Thus, we relax the assumption of a perfect model and account implicitly for structural errors of the VD model during parameter estimation.

For the time being, let us assume that the VD parameter values are known. We write the VD model in a state-space formulation

$$\mathbf{x}_t^{\mathrm{f}} = \mathcal{VD}(\boldsymbol{\theta}, \mathbf{x}_{t-\Delta t}) + \mathbf{q}_t, \tag{4}$$

where $\mathbf{x}_t^{\mathrm{f}}$ is a 4×1 vector with forecasted abundances of the two predator and two prey species, respectively, $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_d\}$ is

Table 1 Description of the Vandermeer (VD) model parameters, including their lower and upper values, and units

Symbol	Description	Lower	Upper	Units	WAVE	SODA	INTEL
β	Predator coefficient	10^{-4}	1.0	–	0.1	0.12 (0.26)	0.29 (0.14)
α	Prey coefficient	0.0	2.0	–	1.5	1.75 (0.21)	1.36 (0.16)
r_1	Growth rate of first prey	0.01	2.5	day^{-1}	0.66	0.08 (0.03)	0.13 (0.16)
r_2	Growth rate of second prey	0.01	2.5	day^{-1}	0.66	0.09 (0.04)	0.12 (0.14)
g	Grazing rate	0.1	2.5	day^{-1}	1.0	0.11 (0.04)	0.24 (0.06)
K	Carrying capacity	0.5	2.5	$(mg\ fwt\ L^{-1})^{1/4}$	1.0	2.47 (0.09)	2.07 (0.12)
m	Mortality rate	0.01	0.7	day^{-1}	0.066	0.05 (0.06)	0.08 (0.08)
H	Parameter of functional response	1.0	3.0	$(mg\ fwt\ L^{-1})^{1/4}$	0.8	1.16 (0.20)	2.39 (0.13)
ρ	First-order autocorrelation of model errors	−1.0	1.0	–	N/A	0.69 (0.26)	−0.35 (0.26)
$\hat{\sigma}^2_w$	Variance of model error	0.0	1.0	–	N/A	0.35 (0.07)	0.20 (0.16)
	RMSE of VD model (no DA: open loop)			$(mg\ fwt\ L^{-1})^{1/4}$	0.63	0.84	0.82
	RMSE of VD model (100% DA)			$(mg\ fwt\ L^{-1})^{1/4}$	0.56	0.21	0.22
	Ensemble mean innovation (100% DA)			$(mg\ fwt\ L^{-1})^{1/4}$	0.38	0.14	0.14

The columns WAVE, SODA and INTEL lists the (optimised) parameter values derived from wavelet analysis, and joint parameter and state estimation with SODA using 100% and 20% of the measured abundances for state updating. The bottom part of the table reports the RMSE of the different VD model parameterisations with and without DA, and the ensemble mean innovation.

the $d = 8$-vector of parameters, and $\mathbf{q}_t$ is a 4×1 process error vector that accounts for structural inadequacies of the VD model. The time step, Δt, is equivalent to the interval of 3.35 days between measured abundances. Thus, rather than using the VD model to simulate the entire 2656 day experiment, the VD model now predicts the one-observation-ahead abundances of the coupled two-predator-two-prey system in eqn 1 based on the state, $\mathbf{x}_{t-\Delta t}$, at time $t - \Delta t$ and the values of the parameters, $\boldsymbol{\theta}$. If we treat the probability distribution of $\mathbf{x}^{\mathrm{f}}_t$ as our prior estimate of the system state at time t, then at any time data, $\tilde{\mathbf{x}}_t$, becomes available, we can derive the posterior forecast state via Bayes' theorem (see e.g. Vrugt *et al.* 2013). If the model in eqn 4 is linear, and the measured and forecast state distribution are multivariate Gaussian, then the KF (Kalman 1960) provides an analytic solution for the optimal estimate of the population counts in our food web at each measurement time t as follows

$$\mathbf{x}^{\mathrm{a}}_t = \mathbf{x}^{\mathrm{f}}_t + \mathbf{K}_t(\tilde{\mathbf{x}}_t - \mathbf{x}^{\mathrm{f}}_t), \tag{5}$$

where $\mathbf{x}^{\mathrm{a}}_t$ is a 4×1 vector of the posterior state, and $\mathbf{K}_t \in [0, 1]$ is a 4×4 matrix of weights called the Kalman gain. The posterior, or analysis, state is our best estimate of the abundances in the food web and is a weighted average of the state forecast, $\mathbf{x}^{\mathrm{f}}_t$, and state measurement, $\tilde{\mathbf{x}}_t$, with weights determined by their individual uncertainties. The second term at the right-hand-side of eqn 5 is also called the state innovation, and equates to $\mathbf{x}^{\mathrm{a}}_t - \mathbf{x}^{\mathrm{f}}_t$.

The Kalman gain, $\mathbf{K}_t \in [0, 1]$, is a 4×4 matrix of relative weights calculated using

$$\mathbf{K}_t = \frac{\mathbf{P}_{\mathrm{f}}}{\mathbf{P}_{\mathrm{f}} + \mathbf{R}} = \mathbf{P}_{\mathrm{f}}(\mathbf{P}_{\mathrm{f}} + \mathbf{R})^{-1} \tag{6}$$

where $\mathbf{P}_{\mathrm{f}}$ and $\mathbf{R}$ signify the 4×4 covariance matrices of the state forecast distribution and state measurement errors, respectively. If forecast errors are much larger than measurement errors, the analysis state, $\mathbf{x}^{\mathrm{a}}_t$, will track closely the measured abundances, $\tilde{\mathbf{x}}_t$. On the contrary, if the measurement errors are large, the posterior state will be close to $\mathbf{x}^{\mathrm{f}}_t$.

The analysis state now enters the VD model to predict the abundances at $t + \Delta t$, and this two-step process of state prediction and state analysis is repeated for every next observation until all population counts have been assimilated by the VD model. This recursion negates forecast bias, and helps close the gap between observed and simulated abundances.

The KF provides an analytic formula for the forecast error covariance matrix, $\mathbf{P}_{\mathrm{f}}$, based on the process error, $\mathbf{q}_t$, and uncertainty in the analysis state, $\mathrm{X}^{\mathrm{a}}_{\mathrm{t}}$, however this formula only applies to linear models. We therefore make use of the Ensemble Kalman filter (EnKF), which approximates $\mathbf{P}_{\mathrm{f}}$ by the sample covariance of an ensemble of N different state forecasts (Evensen 1994). Figure 1 illustrates this approach and Supplement B explains further our choice for the EnKF rather than the more generic particle filter.

The performance of the EnKF depends in large part on the choice of $\mathbf{R}$ and $\mathbf{Q}_t$, the covariance matrix of the process error, $\mathbf{q}_t$. We make the common assumption that the measurement errors are independent, zero-mean normally distributed with variance, $\hat{\sigma}^2_{\mathrm{v}}$. Consequently, $\mathbf{R} = \hat{\sigma}^2_{\mathrm{v}}\mathbf{I}_4$, where $\mathbf{I}_4$ is the 4×4 identity matrix. We follow Evensen (1994) and use temporally correlated process error, $\mathbf{q}_t$, in eqn 4

$$\mathbf{q}_t = \rho\mathbf{q}_{t-\Delta t} + \mathbf{w}_t, \tag{7}$$

where $\rho \in (-1, 1)$ is the first-order autocorrelation coefficient of the process errors, and $\mathbf{w}_t$ is a 4×1 draw from a zero-mean Gaussian distribution with covariance matrix, $\hat{\sigma}^2_{\mathrm{w}}\mathbf{I}_4$.

Thus far, we have assumed the VD parameter values, $\boldsymbol{\theta}$, and $\hat{\sigma}^2_{\mathrm{v}}$, $\hat{\sigma}^2_{\mathrm{w}}$ and ρ to be known, which is not particularly realistic. We therefore use SODA to jointly estimate the VD model states and parameters. This method uses an inner EnKF loop for recursive state estimation conditioned on an assumed parameter set, and an outer loop with the DREAM algorithm for batch estimation of $p(\boldsymbol{\theta}|\tilde{\mathbf{x}})$. We assume, again, a uniform prior parameter distribution, $p(\boldsymbol{\theta})$, and compute the likelihood, $L(\boldsymbol{\theta}|\mathbf{Z})$, of the $k \times n$ matrix of state forecast errors, $\mathbf{Z}$ (rather than residuals without DA). Consequently, SODA will produce parameter values that minimise the one-observation-ahead forecast errors of the VD model. Note, that if $\mathbf{Q}$ is set to a zero matrix (no process noise) then SODA reduces to regular parameter estimation.

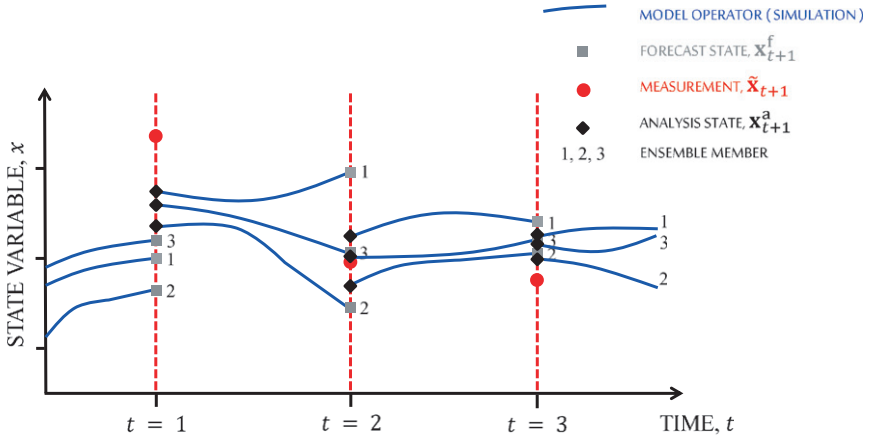

Figure 1 Schematic illustration of the Ensemble Kalman filter (EnKF) with $N = 3$ members for a hypothetical data set with $n = 3$ observations of the system state. The model forecasts the system state from one time step to the next (blue trajectories). Then, when an observation becomes available (red dots), the state forecast of each member (grey squares) is updated to the analysis state (black diamonds). This is known as the analysis or update step. The model then evolves the analysis state of each member to the next observation time. This two-step process of state forecast and state analysis is repeated until all observations of the time series have been assimilated. The state update, or innovation, $\mathbf{i}_t = \mathbf{x}^a_t - \mathbf{x}^f_t$, depends on the variance of the forecast and measurement error, σ^2_w and σ^2_v, respectively. If $\sigma^2_v \ll \sigma^2_w$ then we do not trust much our model forecast, and the analysis state, $\mathbf{x}^a_t$, (black diamonds) will move close to the data (red dots). In this case, the filter is very responsive to the measured data. On the contrary, if $\sigma^w_w \ll \sigma^2_v$ and the measurement errors are relatively large, then the analysis state will remain in close proximity of the forecast (grey squares), $\mathbf{x}^f_t$. If $\sigma^2_w = 0$ then $\mathbf{x}^a_t = \mathbf{x}^f_t$ and $\mathbf{i}_t = 0$ and state estimation has no impact on the modelled state trajectory.

We use an ensemble of $N = 100$ members. Based on replicate measurements of the population counts, we assume that $\hat{\sigma}^2_v = 0.05$, and ρ and $\hat{\sigma}^2_w$ are estimated along with the VD parameters using the measured predator and prey abundances, $\tilde{\mathbf{X}}$.

The 'Intelligent' (INTEL) model

As our food web exhibits complex population dynamics, we should not expect the VD model to track perfectly the observed abundances, in particular the sharp ups and downs in population counts. Yet, state updating may not be strictly necessary during periods with constant abundances. To investigate this further, we consider a second and perhaps more 'intelligent' model, coined INTEL, whose VD parameters and ρ and $\hat{\sigma}^2_w$ are estimated with SODA but with state estimation restricted to 20% of the measured record with largest fluctuations in population counts. In between these 'important' observations, the VD model is executed without state adjustments.

Reduced measurement frequencies

The SODA parameters assume the continued application of DA to the entire 2656 day record. This equates to an average assimilation interval of $\Delta t = 3.35$ days. As data collection may be rather expensive, we investigate how frequently must we assimilate the measured abundances to describe accurately population dynamics in our food web? We evaluate the WAVE, SODA and INTEL parameters with state updates at every $\Delta t = 3.35$, $2\Delta t = 6.70$, $5\Delta t = 16.75$ and $10\Delta t = 33.50$ days, thereby assimilating 100, 50, 20 and 10% of the measured abundances of the two predators and two preys, respectively. This analysis will provide insights into the relationship between the time interval of successive abundance measurements and the forecast skill of the VD model.

Ecological forecast horizon

Data assimilation should enhance the VD model's ability to track the observed population dynamics. Yet, state estimation can only be used 'in sample' as it demands measured abundances. To better understand the 'out of sample' performance of the VD model, we evaluate the EFH of the WAVE, SODA and INTEL parameterisations. Petchey *et al.* (2015) defined the EFH (p. 597) as '*... the dimensional distance for which useful forecasts can be made*'.

To compute the EFH, we need to define a criterion that measures forecast proficiency, and another, related, criterion that classifies a forecast as being 'good' or 'bad'. This second criterion, coined the forecast proficiency threshold, or FPT, by Petchey *et al.* (2015), defines the horizon for which model predictions are deemed good enough, and below which forecasts are considered unacceptable. The EFH now equates to the earliest time it takes for the forecast proficiency to drop below the FPT. The forecast proficiency is set equal to the distance between the observed and predicted abundances. Without formal guidelines on the choice of the FPT, we use FPT = {0.1, 0.2, 0.3, 0.4, 0.5}, and report the corresponding EFH of each species and VD parameterisation. To provide robust EFHs of the predators and preys, we start the VD model at many different times within the experimental record, $\tilde{\mathbf{x}}$ (see Supplement-C for a detailed algorithm).

In practice, one may not use a proficiency threshold to determine whether a forecast is useful or not, but instead select a model because it generates the best predictions. Therefore, we benchmark the EFH of the VD model against that of a simple null model with forecasts equivalent to the measured abundances at initialisation.

RESULTS

To simplify discussion and graphical interpretation, we use colour coding in green, blue and black to differentiate

between the results of the WAVE, SODA and INTEL parameter vectors of Table 1, respectively.

Parameter estimates and model outputs

Table 1 compares the WAVE values of the VD model parameters derived from wavelet analysis (Benincà *et al.* 2009) with their counterparts of SODA and INTEL. The values in parenthesis report the posterior standard deviations of the VD parameters, and the process error variables ρ and $\hat{\sigma}^2_w$. Figure S4 presents box plots of the marginal posterior distributions of the SODA and INTEL parameters.

The SODA and INTEL values of the VD parameters are in good agreement, and correspond reasonably well with their WAVE values. All three parameterisations use values for $\alpha \in [1.36–1.75]$, $\beta \in [0.1–0.3]$ and $m \in [0.05–0.08]$, respectively. This is an encouraging result, as α and β determine the strength of the coupling between the two-predator–prey systems, and thus exert a strong control on the simulated abundances. Note, that WAVE assign much higher values to the prey growth rates, r_1 and r_2, and the grazing rate, g. Different factors may contribute to this disparity. Most importantly, the measured abundances may not contain enough information to back out the gross magnitudes of growth and grazing, and consequently, inference results in many triplets of r_1, r_2 and g that produce similar net effects. This claim is supported by the strong linear dependencies between r_1 and r_2 ($R = 0.99$), and g and m ($R = 0.95$) in the posterior parameter distribution.

SODA also estimates the autocorrelation, ρ, and variance, σ^2_w, of the VD process error. As expected, the autocorrelation is stronger for SODA ($\rho = 0.69$) than for INTEL ($\rho = -0.35$). The process error appears to be 2 (INTEL: $\sigma_w \approx 0.45$) to 2.7 (SODA: $\sigma_w \approx 0.59$) times larger than the measurement error, $\sigma_v = 0.22$. This large process error is required to track the chaotic population oscillations, yet limits the model's ability to accurately predict future abundances. Because the model has no exogenous inputs or heterogeneity, the predictability is governed by chaotic feedbacks and errors in the VD parameters, process representation, and (initial) state (Dietze 2017a). What is currently unresolved is the partitioning of process error, σ^2_w, between model structure and stochasticity in the generation of unstable oscillations in the model. If stochastic variability is independent, the $\rho = 0.69$ of SODA suggests that model structural error dominates. Alternatively, ρ may be capturing unmeasured variability in the mesocosm.

Figure 2 displays the population abundances simulated by the VD model using the WAVE, SODA and INTEL parameters. While the VD model cannot track the observed abundances alone, if we assimilate the measured abundances then the performance of the VD model improves dramatically (Fig. 3). This is especially true for the SODA and INTEL parameters, as they were estimated to minimise the forecast errors, which is confirmed by the lower root mean square error (RMSE) of their one-observation-ahead forecast errors after state estimation (Table 1).

Effect of data assimilation frequency

Figure 4 presents bar charts of the RMSE of the forecast error and mean absolute state innovation of each VD parameterisation. However, DA hardly enhances the performance of the WAVE parameters, the forecast proficiency of SODA and INTEL steadily improves with increasing assimilation frequency. As a consequence, the SODA and INTEL parameters require, on average, smaller state innovations to negate prediction bias and track the observed population counts. Detailed analysis of the state innovations can help pinpoint model structural errors (Vrugt *et al.* 2005). Yet, our attempts to correlate state innovations to each species' population cycle or to other species counts in the food web were unsuccessful.

Altogether, INTEL most efficiently uses DA, with forecast errors that are relatively low per unit effort, if state estimation is applied to every fifth forecast ($\sim$ 17 days). This limit is consistent with the time scale of predictability (15–30 days) inferred by Benincà *et al.* (2008).

Model predictability: forecast horizons

Figure 5 presents the EFH of each species as function of the forecast proficiency threshold (FPT). For completeness, we also include the EFH of a constant forecast (coined NULL) equal to the measured population counts at the start of the prediction period. The null model exhibits a forecast horizon that is systematically larger than the VD model for all species but the Copepods. This is most evident at large FPTs, and testifies to the large process noise of the VD model. The SODA and INTEL parameters exhibit, on average, the smallest forecast errors. In other words, for a given proficiency threshold, the SODA and INTEL parameters can predict further into the future than their WAVE estimates. This translates into a forecast horizon of SODA/INTEL that is just a few days larger than WAVE if FPT = 0.1 and grows rapidly to 20–40 days for FPT = 0.5. What is more, the two zooplankton species exhibit a much larger forecast horizon with SODA, INTEL and NULL than their phytoplankton counterparts. As zooplankton regulate the phytoplankton populations by consumption, they exhibit slower population dynamics, which increases their timescale of predictability. This difference in forecast horizon amounts to a few days if FPT = 0.1 but increases to about 20 days if FPT = 0.5. Note that the forecast horizons increase nonlinearly with FPT.

DISCUSSION

Species abundances in ecological communities can display complex non-equilibrium dynamics (May 1973; Hanski *et al.* 1993; Becks *et al.* 2005; Benincà *et al.* 2008, 2015). A characteristic feature of chaotic systems, is that long-term prediction of the system's trajectory is fundamentally impossible (Strogatz 1995). How then should we make predictions for complex multi-species communities?

The weather may provide a case in point. The chaotic dynamics of weather models (Lorenz 1963), limits the time horizon for reliable weather forecasts to about 1–2 weeks. This forecast skill can be enhanced by assimilating monitoring data from meteorological stations. This DA process uses the incoming observations to update the modelled state variables. Our results demonstrate that ecological models

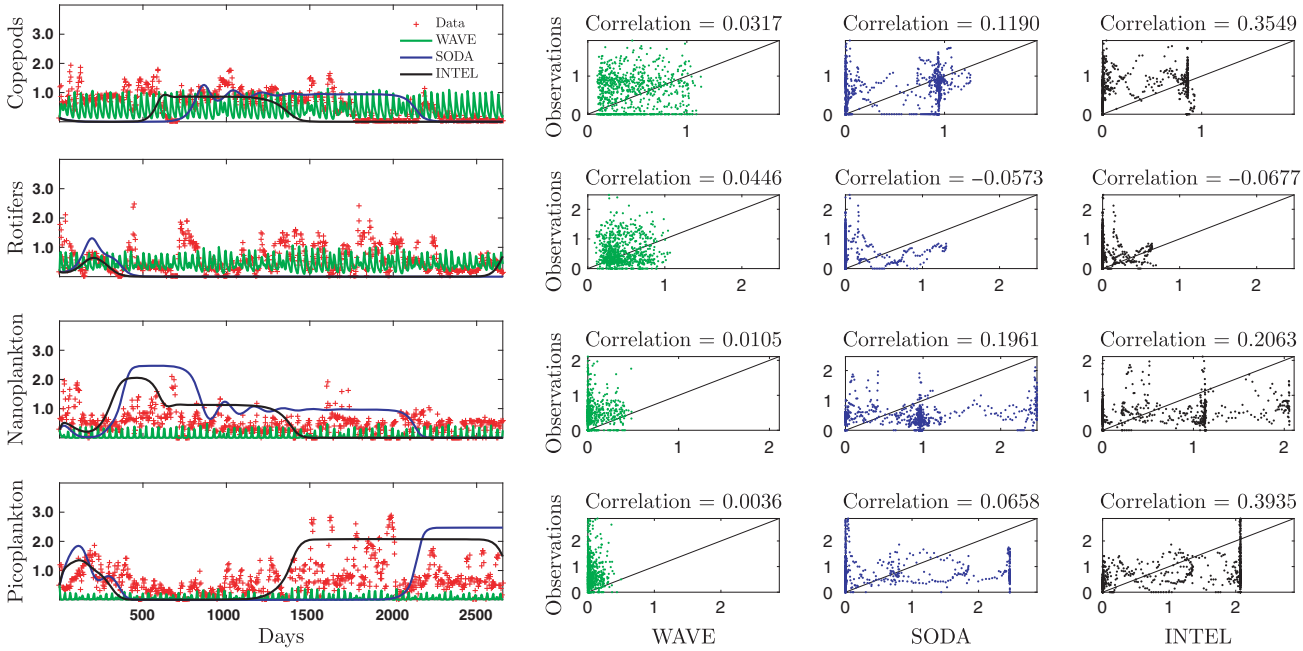

Figure 2 Simulation of the Vandermeer (VD) model for $t \in [0, 2656]$ days using the WAVE, SODA and INTEL parameter values: Time series plots of measured and simulated abundances of the predator and prey species (left panels), and scatter plots of simulated and observed abundances (right panels). The three parameterisations (solid lines) poorly describe the highly nonlinear and chaotic population dynamics of the two predator and two prey species in our microbial food web. Therefore, we investigate the use data assimilation to reduce simulation bias of the VD model, and close the gap between the observed and simulated species abundances.

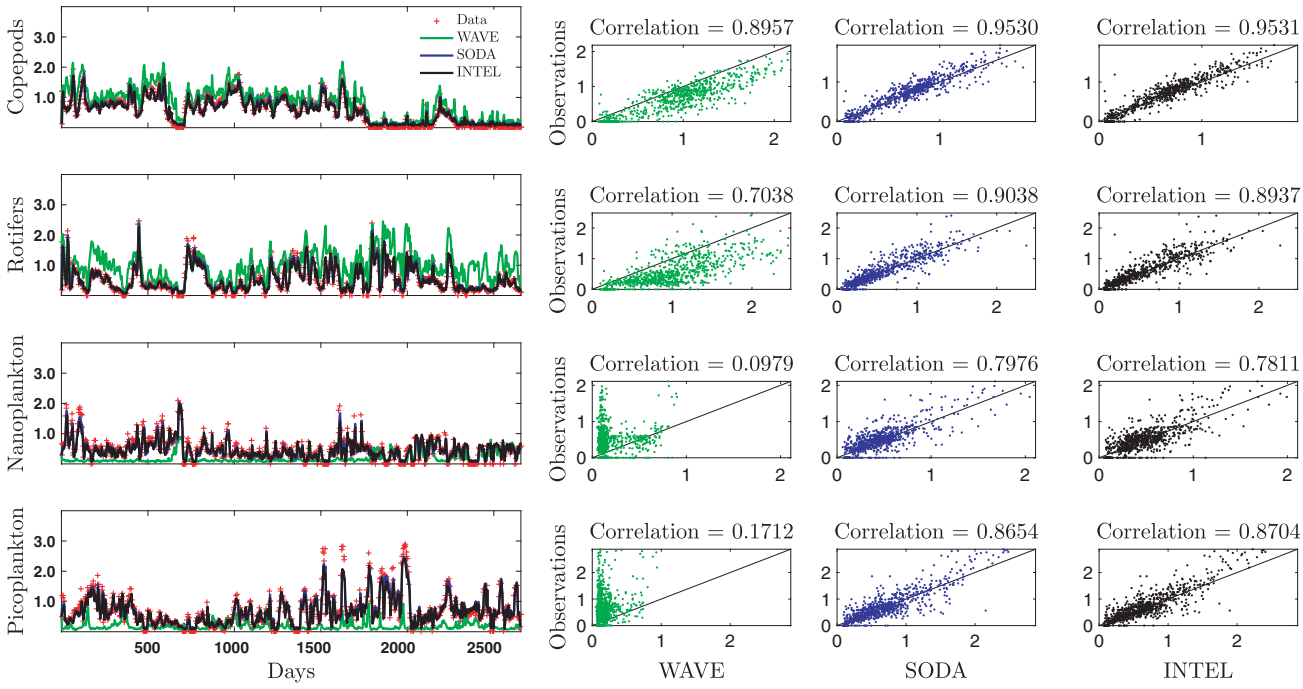

Figure 3 Forecasted abundances of the Vandermeer (VD) model during our 8-year experiment with the WAVE, SODA and INTEL parameters using state estimation with the EnKF. Time series plots of measured and forecasted abundances of the predator and prey species (left panels), and scatter plots of simulated and observed abundances (right panels). Data assimilation enhances significantly the ability of the VD model to track the population dynamics of the predators and preys. This is particularly true for the SODA and INTEL parameterisations, which closely follow the observed population counts.

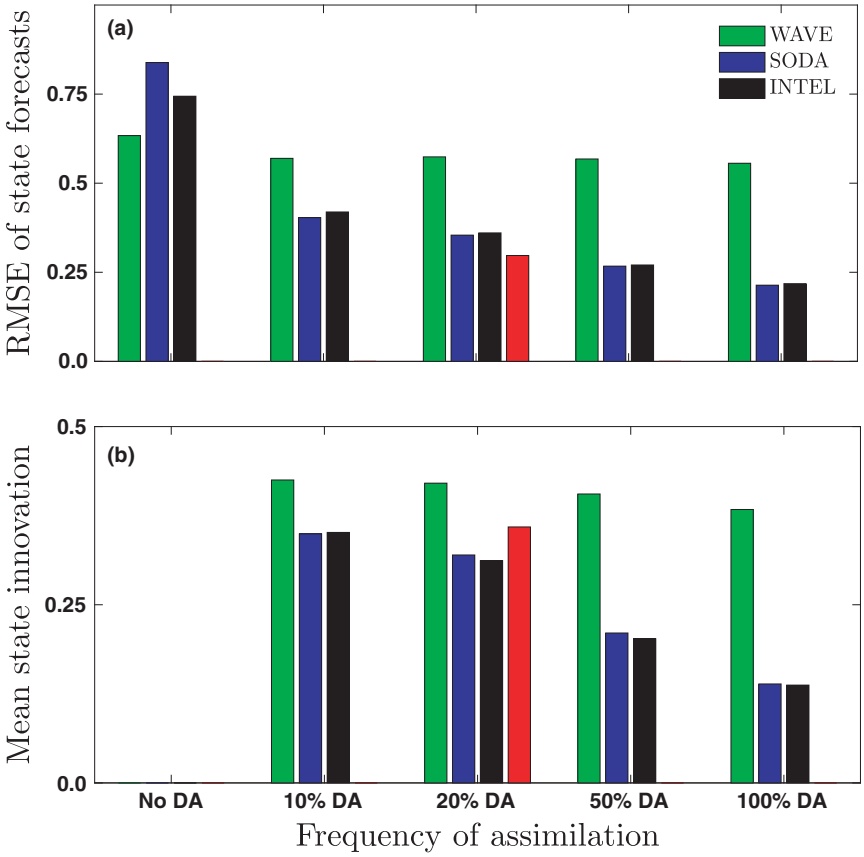

Figure 4 (a) Root mean square error, or RMSE, of the forecasted abundances of the Vandermeer model, using the WAVE, SODA and INTEL parameters with different assimilation frequencies, and (b) ensemble mean value of the state innovation. The red bars signify the results of the INTEL model with state estimation restricted to 20% of the time steps with largest fluctuations in the observed data.

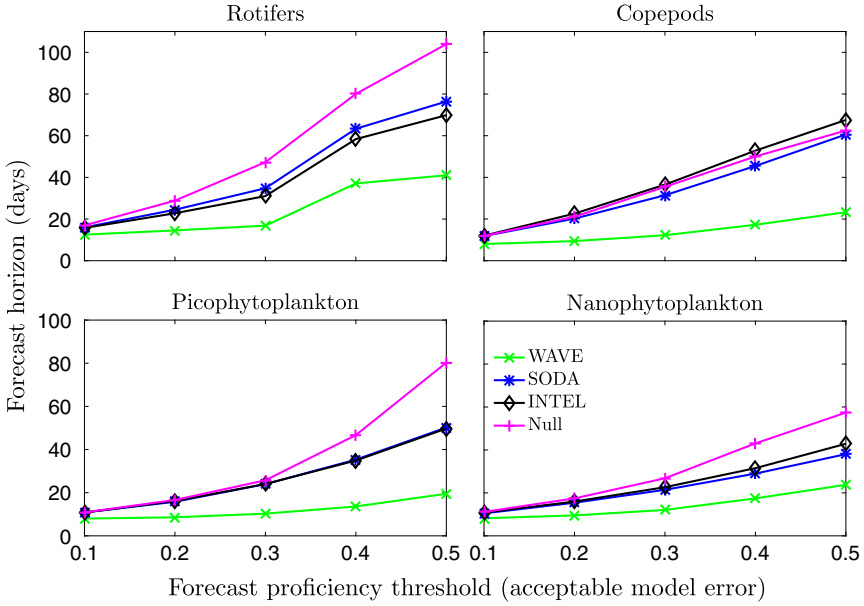

Figure 5 Ecological forecast horizon of the Vandermeer (VD) model derived from forward simulation with the WAVE, SODA and INTEL parameter values using many different initial states drawn from the mesocosm data set. The forecast horizon is defined as the first time from the start of simulation at which the distance between the state forecast and corresponding abundance observation exceeds a threshold, called the forecast proficiency threshold (FPT). The forecast horizons quantify the ability of the VD model to produce 'good' forecasts, without data assimilation. For benchmark purposes, the purple line presents the forecast horizon of a NULL model with a constant forecast equivalent to the measured population count at initialisation.

may benefit from a similar DA approach. We have used DA with the EnKF to fuse the simulated abundances of the two-predator-two-prey VD model with observed population counts. The state adjustments of the EnKF negate, at least in part, a high sensitivity of the VD model to initial conditions (chaos) and process errors, and allow system dynamics to be better described. Indeed, our results with SODA demonstrate that joint parameter and state estimation enhances considerably the predictive skill of the VD model. The SODA approach considers explicitly multivariate parameter correlations and trade-offs, and avoids any bias from individual estimation.

The results provide interesting insights into the biology of the experimental system. In general, a value of $\alpha > 0$ in the VD model illustrates that the two predator-prey systems are coupled through competition, and a value of $\beta > 0$ demonstrates that both predators feed on both prey species (Vandermeer 2004; Benincà *et al*. 2009). For all three parameter sets, we consistently found a value of $\alpha > 1$, which indicates that interspecific competition between the two phytoplankton groups (picocyanobacteria and nanoflagellates) was larger than intraspecific competition. In the absence of predation, this would favour competitive exclusion, where the initial conditions determine which of these two phytoplankton groups would win. Hence, the coexistence of these two phytoplankton groups is mediated by predation. Moreover, all three parameter sets consistently found a relatively low value for β, which implies that there is little diet overlap between the rotifers (feeding mainly on the picocyanobacteria) and calanoid copepods (feeding mainly on the nanoflagellates). According to the VD model, strong coupling through competition (high α) and weak coupling through predation (low β) will lead to anti-phase oscillations of the two predator-prey systems (Vandermeer 2004). That is, when the picocyanobacteria go up, the nanoflagellates go down (and *vice versa*). Similarly, when the rotifers go up, the calanoid copepods go down (and *vice versa*). Indeed, these alternations in species dominance at both trophic levels are observed in the time series data of this experiment (Benincà *et al*. 2009).

To assess how often the mesocosm needs to be sampled to accurately predict population dynamics, we examined the forecast errors of the WAVE, SODA and INTEL parameterisations, using different assimilation frequencies. In line with expectations, higher measurement frequencies decrease the forecast error. Furthermore, measurement frequencies lower than the forecast horizon had noticeably higher RMSE's and much larger state innovations. Ecological monitoring can be expensive, however, and reductions of the measurement frequency may therefore be attractive. To this end, we introduced an 'intelligent' model, called INTEL which restricts state estimation to those measurement times with largest fluctuations in population counts. This approach was shown to be the most efficient use of DA. Hence, our results illustrate that DA not only allows ecologists to fuse their models with data, but may also provide guidance on measurement design and the frequency and timing of observations.

To assess the impact of modelling approach on forecast horizon, we quantified the forecast proficiency of the VD model, using different proficiency thresholds. The SODA and INTEL parameters exhibited much larger forecast horizons than their WAVE counterparts and a previously published neural network (Benincà *et al*. 2008). However, they do not surpass the forecast horizons of a simple null model with a constant forecast of each species population count. We cannot confidently claim why long-term population dynamics in our food web is unpredictable. First, sudden, chaotic state transitions, make it fundamentally impossible to predict system dynamics far into the future. Second, process errors of the VD model limit its forecast horizon even during periods with stable population counts. Third, even well-controlled biotic systems are subject to stochastic perturbations and inhomogeneities. Fourth, parameter errors also contribute to forecast uncertainty.

Process error contributed most to prediction uncertainty, and its autocorrelation, $\rho = 0.69$, suggests the VD model suffers from structural errors. Indeed, the assumption that the two zooplankton groups share the same values of g, H and m may be too simplifying as rotifers and copepods are different organisms, with dissimilar grazing rates, mortality, and prey treatments. A more exhaustive parameterisation may enhance the VD model's ability to describe the coupled predator-prey oscillations. The observed oscillations may also be impacted by variations in nutrients or other organisms in the food web, which were not accounted for in the VD model. Finally, it is possible that the assumed functional forms in the VD model fail to capture the responses observed in nature. Further examination of the time series of state innovations may help diagnose structural errors.

Although DA improves the forecasting ability of ecological models, it does not necessarily yield parameter values that better describe ecological processes. Without DA, for example, the SODA and INTEL parameters do not reproduce the species' population oscillations (Fig. 2), a feature that WAVE replicates.

In conclusion, we have demonstrated that DA is an attractive approach, as it will help close the gap between ecological observations and theory, and increase the forecast horizon of ecosystem models. We look forward to further advances in ecological forecasting, especially related to systems with chaotic dynamics.

ACKNOWLEDGEMENTS

We appreciate the comments of the two anonymous reviewers and the Editor, which have greatly enhanced the quality of our manuscript. The first and last author greatly appreciate the support from UC-Lab Fees Research Program Award 237285 and the UCI Environment Institute, and the fourth author acknowledges support from NSF grant 1638577. We thank R. Heerkloss for the mesocosm data set, which is available in Appendix S1 of Benincà *et al*. (2009). The SODA methodology can be obtained from the last author, (jasper@uci.edu), upon request.

AUTHOR CONTRIBUTIONS

All authors contributed ideas for the study, EM performed model simulations, EM and JAV wrote the first draft of the manuscript. All authors contributed to revisions.

REFERENCES

Aanensen, D., Huntley, D., Feil, E., al Own, F. & Spratt, B. (2009). Epicollect: linking smartphones to web applications for epidemiology, ecology and community data collection. *PLoS ONE*, 4, e6968.

Ali, A., Xu, C., Rogers, A., Fisher, R., Wullschleger, S., Massoud, E. *et al.* (2016). A global scale mechanistic model of the photosynthetic capacity (luna v1.0). *Geosci. Model Dev.*, 9, 587–606.

Becks, L., Hilker, F., Malchow, H., Jürgens, K. & Arndt, H. (2005). Experimental demonstration of chaos in a microbial food web. *Nature*, 435, 1226–1229.

Benincà, E., Huisman, J., Heerkloss, R., Jöhnk, K., Branco, P., Van Nes, E. *et al.* (2008). Chaos in a long-term experiment with a plankton community. *Nature*, 451, 822–825.

Benincà, E., Jöhnk, K., Heerkloss, R. & Huisman, J. (2009). Coupled predator–prey oscillations in a chaotic food web. *Ecol. Lett.*, 12, 1367–1378.

Benincà, E., Ballantine, B., Ellner, S. & Huisman, J. (2015). Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proc. Natl Acad. Sci.*, 112, 6389–6394.

Bertino, L., Evensen, G. & Wackernagel, H. (2003). Sequential data assimilation techniques in oceanography. *Int. Stat. Rev.*, 71, 223–241.

Chen, M., Liu, S., Tieszen, L. & Hollinger, D. (2008). An improved state-parameter analysis of ecosystem models using data assimilation. *Ecol. Model.*, 219, 317–326.

Clark, J. & Bjornstad, O. (2004). Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology*, 85, 3140–3150.

Clark, J., Carpenter, S., Barber, M., Collins, S., Dobson, A., Foley, J. *et al.* (2001). Ecological forecasts: an emerging imperative. *Science*, 293, 657–660.

Conroy, M., Cohen, Y., James, F., Matsinos, Y. & Maurer, B. (1995). Parameter estimation, reliability, and model improvement for spatially explicit models of animal populations. *Ecol. Appl.*, 5, 17–19.

Conway, G., Glass, N. & Wilcox, J. (1970). Fitting nonlinear models to biological data by Marquardt's algorithm. *Ecology*, 51, 503–507.

Dietze, M. (2017a). Prediction in ecology: a first-principles framework. *Ecol. Appl.*, 27, 2048–2060.

Dietze, M.C. (2017b). *Ecological Forecasting*. Princeton University Press, Princeton, NJ.

Dietze, M., LeBauer, D. & Kooper, R. (2013). On improving the communication between models and data. *Plant Cell Environ.*, 36, 1575–1585.

Dorigo, W., Zurita-Milla, R., de Wit, A., Brazile, J., Singh, R. & Schaepman, M. (2007). A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Appl. Earth Obs. Geoinf.*, 9, 165–193.

Dowd, M. (2007). Bayesian statistical data assimilation for ecosystem models using Markov Chain Monte Carlo. *J. Mar. Syst.*, 68, 439–456.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res. Oceans*, 99, 10143–10162.

Fasham, M., Ducklow, H. & McKelvie, S. (1990). A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *J. Mar. Res.*, 48, 591–639.

Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F. *et al.* (2015). Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments. *J. Oper. Oceanogr.*, 8, s168–s187.

Gelman, A. & Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.*, 7, 457–472.

Granzow, G. (2014). A tutorial on adjoint methods and their use for data assimilation in glaciology. *J. Glaciol.*, 60, 440–446.

Hanski, I., Turchin, P., Korpimäki, E. & Henttonen, H. (1993). Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos. *Nature*, 364, 232–235.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.*, 82, 35–45.

Keenan, T., Carbone, M., Reichstein, M. & Richardson, A. (2011). The model–data fusion pitfall: assuming certainty in an uncertain world. *Oecologia*, 167, 587–597.

Kendall, B., Briggs, C., Murdoch, W., Turchin, P., Ellner, S., McCauley, E. *et al.* (1999). Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. *Ecology*, 80, 1789–1805.

LaDeau, S., Han, B., Rosi, E. & Weathers, K. (2017). The next decade of big data in ecosystem science. *Ecosystems*, 20, 274–283.

Lawson, L., Spitz, Y., Hofmann, E. & Long, R. (1995). A data assimilation technique applied to a predator-prey model. *Bull. Math. Biol.*, 57, 593–617.

Lawson, L., Hofmann, E. & Spitz, Y. (1996). Time series sampling and data assimilation in a simple marine ecosystem model. *Deep Sea Res. Part II*, 43, 625–651.

Lorenz, E. (1963). Deterministic nonperiodic flow. *J. Atmospheric Sci.*, 20, 130–141.

Luo, Y. & Schimel, D. (2011). Model improvement via data assimilation toward ecological forecasting 1. *Ecol. Appl.*, 21, 1427–1428.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S. *et al.* (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecol. Appl.*, 21, 1429–1442.

May, R. (1973). *Stability and Complexity in Model Ecosystems*, vol. 6. Princeton University Press, Princeton, NJ.

McMahon, S., Dietze, M., Hersh, M., Moran, E. & Clark, J. (2009). A predictive framework to understand forest responses to global change. *Ann. N. Y. Acad. Sci.*, 1162, 221–236.

Mo, X., Chen, J., Ju, W. & Black, T. (2008). Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an Ensemble Kalman filter. *Ecol. Model.*, 217, 157–173.

Natvik, L., Eknes, M. & Evensen, G. (2001). A weak constraint inverse for a zero-dimensional marine ecosystem model. *J. Mar. Syst.*, 28, 19–44.

Niu, S., Luo, Y., Dietze, M., Keenan, T., Shi, Z., Li, J. *et al.* (2014). The role of data assimilation in predictive ecology. *Ecosphere*, 5, 1–16.

Peng, C., Guiot, J., Wu, H., Jiang, H. & Luo, Y. (2011). Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach. *Ecol. Lett.*, 14, 522–536.

Petchey, O., Pontarp, M., Massie, T., Kéfi, S., Ozgul, A., Weilenmann, M. *et al.* (2015). The ecological forecast horizon, and examples of its uses and determinants. *Ecol. Lett.*, 18, 597–611.

Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B., Disney, M. *et al.* (2008). Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman filter. *Remote Sens. Environ.*, 112, 1347–1364.

Reichman, O., Jones, M. & Schildhauer, M. (2011). Challenges and opportunities of open data in ecology. *Science*, 331, 703–705.

Running, S., Baldocchi, D., Turner, D., Gower, S., Bakwin, P. & Hibbard, K. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and eos satellite data. *Remote Sens. Environ.*, 70, 108–127.

Simmons, A. & Hollingsworth, A. (2002). Some aspects of the improvement in skill of numerical weather prediction. *Q. J. Royal Meteorol. Soc.*, 128, 647–677.

Sitch, S., Smith, B., Prentice, I., Arneth, A., Bondeau, A., Cramer, W. *et al.* (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the lpj dynamic global vegetation model. *Glob. Change Biol.*, 9, 161–185.

Sitz, A., Schwarz, U., Kurths, J. & Voss, H. (2002). Estimation of parameters and unobserved components for nonlinear systems from noisy time series. *Phys. Rev. E*, 66, 016210.

Strogatz, S. (1995). Ordering chaos with disorder. *Nature*, 378, 444–444.

Ter Braak, C. & Van Tongeren, O. (1995). *Data Analysis in Community and Landscape Ecology*. Cambridge University Press, Cambridge.

Vallino, J. (2000). Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *J. Mar. Res.*, 58, 117–164.

Vandermeer, J. (1982). On the resolution of chaos in population models. *Theor. Popul. Biol.*, 22, 17–27.
Vandermeer, J. (2004). Coupled oscillations in food webs: balancing competition and mutualism in simple ecological models. *Am. Nat.*, 163, 857–867.
Vrugt, J. (2016). Markov Chain Monte Carlo simulation using the DREAM software package: theory, concepts, and matlab implementation. *Environ. Model. Softw.*, 75, 273–316.
Vrugt, J., Diks, C., Gupta, H., Bouten, W. & Verstraten, J. (2005). Improved treatment of uncertainty in hydrologic modeling: combining the strengths of global optimization and data assimilation. *Water Resour. Res.*, 41, W01017.
Vrugt, J., ter Braak, C., Diks, C. & Schoups, G. (2013). Hydrologic data assimilation using particle Markov Chain Monte Carlo simulation: theory, concepts and applications. *Adv. Water Resour.*, 51, 457–478.
Williams, M., Schwarz, P., Law, B., Irvine, J. & Kurpius, M. (2005). An improved analysis of forest carbon dynamics using data assimilation. *Glob. Change Biol.*, 11, 89–105.
Williams, M., Richardson, A., Reichstein, M., Stoy, P., Peylin, P., Verbeeck, H. *et al.* (2009). Improving land surface models with fluxnet data. *Biogeosciences*, 6, 1341–1359.
Wintle, B., McCarthy, M., Volinsky, C. & Kavanagh, R. (2003). The use of bayesian model averaging to better represent uncertainty in ecological models. *Conserv. Biol.*, 17, 1579–1590.
Wu, H., Guiot, J., Peng, C. & Guo, Z. (2009). New coupled model used inversely for reconstructing past terrestrial carbon storage from pollen data: validation of model using modern data. *Glob. Change Biol.*, 15, 82–96.
Xiao, Y. & Friedrichs, M. (2014). The assimilation of satellite-derived data into a one-dimensional lower trophic level marine ecosystem model. *J. Geophys. Res. Oceans*, 119, 2691–2712.
Xu, C., Fisher, R., Wullschleger, S., Wilson, C., Cai, M. & McDowell, N. (2012). Toward a mechanistic modeling of nitrogen limitation on vegetation dynamics. *PLoS ONE*, 7, e37914–e37914.
Zobitz, J., Desai, A., Moore, D. & Chadwick, M. (2011). A primer for data assimilation with ecological models using Markov Chain Monte Carlo (MCMC). *Oecologia*, 167, 599–611.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Editor, Frederick Adler
Manuscript received 31 August 2017
First decision made 21 September 2017
Manuscript accepted 7 October 2017