Spatially dependent parameter estimation and nonlinear data assimilation by autosynchronization of a system of partial differential equations

Sean Kramer and Erik M. Bollt

Citation: Chaos **23**, 033101 (2013); doi: 10.1063/1.4812722
View online: http://dx.doi.org/10.1063/1.4812722
View Table of Contents: http://chaos.aip.org/resource/1/CHAOEH/v23/i3
Published by the AIP Publishing LLC.

Spatially dependent parameter estimation and nonlinear data assimilation by autosynchronization of a system of partial differential equations

Sean Kramer and Erik M. Bollt
Department of Mathematics, Clarkson University, Potsdam, New York 13669, USA

(Received 27 July 2012; accepted 17 June 2013; published online 8 July 2013)

Given multiple images that describe chaotic reaction-diffusion dynamics, parameters of a partial differential equation (PDE) model are estimated using autosynchronization, where parameters are controlled by synchronization of the model to the observed data. A two-component system of predator-prey reaction-diffusion PDEs is used with spatially dependent parameters to benchmark the methods described. Applications to modeling the ecological habitat of marine plankton blooms by nonlinear data assimilation through remote sensing are discussed. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4812722]

When attempting to model many physical processes, a frequent roadblock is the inability to observe certain variables and parameters. For example, hyperspectral satellite imagery provides a means to model phytoplankton ecology in the ocean. However, a realistic model includes predator-prey interaction with zooplankton, the primary controller of phytoplankton blooms, and unobservable by satellite imagery. Data assimilation, wherein measured data are incorporated into model output, is crucial to obtain a model that fits real-time observation. We address the problem of data assimilation and parameter estimation for a system of reaction-diffusion partial differential equations (PDEs) wherein only one species is observable such that modeling ocean ecology by means of satellite imagery is plausible. This article extends former results in ordinary differential equations by so-called autosynchronization to the PDE case; the results permit an important generalization, estimating parameters that are spatially heterogeneous.

I. INTRODUCTION

Parameter estimation in *ordinary differential equations* (ODEs) and partial differential equations (PDEs) has developed into a vast field in applied mathematics and control engineering. For models representing important physical processes, accurate estimates of appropriate model parameters help facilitate short-term forecasting. However, to forecast a system, one requires not only accurate parameter estimates but also full knowledge of the initial state of the system. There are widely varying and powerful methods for parameter estimation of spatio-temporal systems including, but certainly not limited to, Kalman filtering methods,[1–3] multiple shooting methods,[4,5] and adjoint methods.[6] In fact, parameter estimation based on synchronization has drawn substantial interest.[7–18] Applications include communications and cryptography,[10] electronics and circuit dynamics,[9,19] and cardiac cell dynamics[15] to name just a few. There are currently several methods to estimate parameters based on synchronization, a field that emerged after the seminal work of Pecora and Carroll in Ref. 7. One approach is to optimize a time-averaged synchronization error on which synchronization acts as a regularizing force; the optimization problem of finding the minimum synchronization error in parameter space is well-posed.[9,14–17] Our interest here will be based on an approach to force a response model to adapt to observed data by developing additional equations for the parameters that depend on the synchronization error.[8,13]

To estimate model parameters by synchronization, we exploit a special variation of synchronization called "autosynchronization." For systems of ODEs, an observed scalar time series is coupled to a response system during model simulation. The goal of this feedback is to cause the response system to synchronize to the drive system. Ideally, a proof of convergence follows by demonstration of an appropriate Lyapunov function.[8,11] In Refs. 12 and 13, we see some generalizations of how to derive synchronization schemes for many systems including the case where, *a priori*, we do not know the model form of the drive system.[13] By autosynchronization, we recover the model parameters, the current model state, and in some cases, a model form for an observed system.

Stating an autosynchronization problem in the ODE setting, suppose $\mathbf{u}_t = \mathbf{f}(\mathbf{u}, \mathbf{p})$, where $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{p} \in \mathbf{R^m}$, $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$, and $\mathbf{u}_t$ represents the time derivative of $\mathbf{u}$. Further suppose $\mathbf{u}_t = \mathbf{f}(\mathbf{u}, \mathbf{p})$ is a nonlinear system of differential equations such that the model form of $\mathbf{f}$ is known. We require a drive system

$$\mathbf{u_t} = \mathbf{f}(\mathbf{u}, \mathbf{p}) \tag{1}$$

from which we are able to sample data with (unknown to us) parameters $\mathbf{p} \in \mathbf{R^m}$. Then we must state a response system

$$\mathbf{v_t} = \mathbf{g}(\mathbf{u}, \mathbf{v}, \mathbf{q}), \tag{2}$$

where $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{q} \in \mathbf{R^m}$, $\mathbf{g} : E \to \mathbb{R}^n$, E is an open subset of $\mathbb{R}^{2n}$, and $\mathbf{v}_t$ represents the time derivative of $\mathbf{v}$. Equation (2) has the same model form as the drive system if $\mathbf{q} = \mathbf{p}$. By "same" we mean in as far as possible by our understanding of the underlying physics. Then the goal is that when $\mathbf{u}$ is coupled forward into Eq. (2), then Eq. (2) will synchronize with Eq. (1) and $\mathbf{v} \to \mathbf{u}$. Furthermore, parameter ODEs are given by

$$\mathbf{q_t} = \mathbf{h}(\mathbf{u}, \mathbf{v}, \mathbf{q}), \tag{3}$$

so that $(\mathbf{v}, \mathbf{q}) \to (\mathbf{u}, \mathbf{p})$ as $t \to \infty$.

The idea of synchronization was extended to one-dimensional systems of PDEs in Ref. 19 and two-dimensional systems in Ref. 16, where the authors considered the Grey-Scott and Barkely reaction-diffusion systems respectively. In these works, the authors observed synchronization of an infinite-dimensional system by coupling the drive and response systems at only a finite number of grid points. Further work examines parameter estimation for given PDE systems using optimization over the synchronization error surface. The authors observe single-species assimilation as they drive the PDE system to synchronization while coupling with only one species.[16] However, none of these utilize autosynchronization.

In many systems, it is very reasonable to expect that model parameters need not be spatially homogeneous. For example, taking our problem of interest, spatial inhomogeneity in parameter values may be central when constructing models for coastal algal blooms, since plankton growthrate is affected by near-shore nutrient runoff and upwelling.[20] More to that point, ocean fronts and eddies cause flow-induced long-term inhomogeneities in the ocean which results in a formidable spatial structure for density profiles in the ocean.[20] Whether inhomogeneities be the result of the flow dynamics or of "boundary conditions" from nutrient runoff, they are an important consideration for modelling ecology over large coastal domains. Thus, it is reasonable to argue that a biophysics-based model over the mesoscale should accept spatially dependent parameters.

Parameter estimation by filtering methods adapted for PDEs can be computationally expensive.[5] Furthermore, we have found that some filtering methods suffer during periods of exponential growth, such as might be expected during plankton blooms. Optimizing the time-averaged synchronization error in some function space is far more complicated than the finite-dimensional alternative with scalar parameters as in Ref. 16; optimization methods may not be practical.

Our work aims to extend the method of parameter estimation for PDE systems by synchronization to autosynchronization, especially including autosynchronization with spatially dependent parameters. Thus, we investigate observed data from the PDE drive system

$$\mathbf{u_t}(x, y, t) = \mathbf{f}(\mathbf{u}(x, y), \mathbf{p}(x, y)), \tag{4}$$

with parameters $\mathbf{p}(x, y) \in C^0(\Omega)$ and a response system

$$\mathbf{v_t}(x, y, t) = \mathbf{g}(\mathbf{u}(x, y), \mathbf{v}(x, y), \mathbf{q}(x, y)). \tag{5}$$

We formulate an associated system of PDEs for the parameters of Eq. (5)

$$\mathbf{q_t}(x, y, t) = \mathbf{h}(\mathbf{u}(x, y), \mathbf{v}(x, y)), \tag{6}$$

where $\mathbf{h} : C^0(\Omega) \to C^0(\Omega)$, with the goal that $(\mathbf{v}, \mathbf{q}) \to (\mathbf{u}, \mathbf{p})$ as $t \to \infty$. We design our methods by considering a benchmark system of reaction-diffusion PDEs. Since we know the model form of the drive system and the parameters used to build the observed data, we compare our estimated parameters with the exact parameters.

In preview of the paper layout, we begin by introducing the reaction-diffusion equations that we will use as the drive system. We discuss how this system is solved numerically and the parameters used to simulate complex spatiotemporal dynamics. Next, we implement the response system and show the parameter PDEs used to find autosynchronization. We demonstrate the power of the systems of PDEs to autosynchronize by employing three different spatial functions for the parameters. Next, we show the estimated parameters and the convergence plots for both state variables and parameters to the correct values. Finally, we give an improvement on the response system that admits autosynchronization wherein only one species is sampled, which is an important breakthrough for applications since generally only the phytoplankton is easily observable.

II. THE PARAMETER ESTIMATION METHOD

Spatiotemporal dynamics for aquatic ecosystems including phytoplankton are often modeled by two-species reaction-diffusion predator-prey systems.[21–25] These systems describe predator-prey ecology between two species, often Phytoplankton and Zooplankton, where diffusivities are equal for both species since mixing is largely the result of marine churning.[20] Many systems derived include reaction terms such that zooplankton populations increase by grazing upon phytoplankton with some grazing efficiency,[26] phytoplankton increase from exposure to nutrients, and each species is subject to a mortality rate based on its current population size or predation.

Consider the system of two PDEs as given in Ref. 20,

$$\frac{\partial P}{\partial t} = \triangle P + P(1-P) - \frac{PZ}{P+h}, \quad \frac{\partial Z}{\partial t} = \triangle Z + k\frac{PZ}{P+h} - mZ, \tag{7}$$

where $\triangle$ denotes the two-dimensional Laplacian, $P(x, y, t)$ is the abundance of phytoplankton at a point in the domain at a given time, $Z(x, y, t)$ is the zooplankton abundance, and k, m, and h are dimensionless model parameters. Equation (7) is simulated on a compact connected two-dimensional domain, Ω, with zero-flux boundary conditions.

In terms of the biology of the model, the system represents a dimensionless reaction-diffusion model for phytoplankton-zooplankton predator prey dynamics in a horizontal layer where vertical distributions of plankton are considered uniform. Although shown in dimensionless form, the model is derived from principles in which phytoplankton concentrations obey a logistic growth and are grazed upon by zooplankton, following a Holling-type II functional response. First classified by Holling,[26] the Holling-type II functional response assumes a decelerating growth rate such that the predator, or consumer, is limited by its ability to efficiently process food. Zooplankton grow at a rate proportional to phytoplankton mortality and are subject to a natural mortality rate. In dimensionless form, the growth and death rates for phytoplankton are absorbed in the dimensionless parameters k, m, and h. For a range of parameter values, e.g., $k = 2$,

$h = 0.4$, and $m = 0.6$, this system exhibits spatiotemporal chaos after transient spiral pattern behaviour.[20]

For simulations shown here, we choose Ω to be a rectangle of size 864×288. For homogeneous initial plankton distributions, the system remains in a homogeneous state for all time, therefore, we use the perturbed initial conditions found in Ref. 20.

We solve this system with a finite difference method, using a three-point centered difference stencil for spatial derivatives and forward Euler time stepping. Our simulations use spatial discretization with $dx = 2$ and and Euler time step of $dt = 0.2$. Although Eq. (7) is nonlinear, the choice of step size satisfies the linear stability requirement for the Euler scheme, i.e., that $\frac{dt}{dx^2} + \frac{dt}{dy^2} \leq \frac{1}{2}$. The model output is treated as an image sequence given by a particular (known) model form but with *unknown* parameters k and m, to be determined. Thus, we will mimic our target application of remote sensing oceanographic images of hyperspectral images filtered to reveal plankton blooms. Here, we observe over every point in the grid as would be available from satellite data, however, it has been shown that synchronization is possible by sub-sampling the grid.[16] Moreover, we will be interested to allow k and m to vary spatially as functions, $k(x, y)$ and $m(x, y)$ as they also account for phytoplankton dynamics.

Our interest in this PDE model stems from our work in remote sensing, to build a better understanding of our ocean's ecology. Particularly, we aim to predict short-term behavior of coastal algal blooms. Such a system may in principle be modeled by estimating parameters directly from observed data in the field. This additional data could then be used to validate model predictions informed by remote sensing data. However, hyperspectral satellite imagery provides the observed data to which we would synchronize a response model in hopes of autosynchronization providing good parameter estimates for forecasting. Since phytoplankton are largely affected by spatial inhomogeneities in the ocean such as nitrogen runoff, regions of hypoxia, or upwelling, to name a few parameter inhomogeneity-inducing effects, we wish to allow model parameters to vary spatially. These considerations are especially important since our models will be built over coastal domains where large changes in ocean biology occur spatially, leading naturally to spatially dynamic parameters.

We are only able to observe time series data Eq. (4) as a movie and we know the model form of Eq. (7), but want to estimate the parameters used to create the observed data. The system in Eq. (7) will be taken as the drive system and we form a response system to be synchronized to the observations as,

$$\begin{aligned} \frac{\partial \hat{P}}{\partial t} &= \triangle \hat{P} + \hat{P}(1-\hat{P}) - \frac{\hat{P}\hat{Z}}{\hat{P}+h} + \kappa(P-\hat{P}), \\ \frac{\partial \hat{Z}}{\partial t} &= \triangle \hat{Z} + \hat{k}\frac{\hat{P}\hat{Z}}{\hat{P}+h} - \hat{m}\hat{Z} + \kappa(Z-\hat{Z}), \end{aligned} \tag{8}$$

where we assume $\hat{P}(x,y,0) \neq P(x,y,0), \hat{Z}(x,y,0) \neq Z(x,y,0)$, $\hat{k}(x,y,0) \neq k(x,y)$, and $\hat{m}(x,y,0) \neq m(x,y)$. Thus, we do not know the initial model states, and wish to recover the spatially varying parameters $m(x, y)$ and $k(x, y)$. To derive Eq. (8), a diffusive coupling term is added to each equation in Eq. (7) accounting for the error between the drive and response values with a coupling strength, κ. These additional terms drive $\hat{P} \to P$ and $\hat{Z} \to Z$, so that the PDEs will synchronize after a short time. The synchronization is of identical type and dependent upon the choice of κ, as is the synchronization speed.

III. RESULTS AND SIMULATIONS OF AUTOSYNCHRONIZATION PARAMETER ESTIMATION

We modify the system Eq. (7) as found in Ref. 20 by forcing the parameters to be non-negative $C^0(\Omega)$ functions. Here Ω is the domain, which in the case of our simulations, $\Omega \subset \mathbf{R}^2$ is a compact domain such as a rectangle or even a domain shaped as the Gulf of Mexico. Parameters are updated as diffusively coupled PDEs during the synchronization process as

$$\frac{\partial \hat{k}}{\partial t} = -s(P-\hat{P}) \quad s > 0, \quad \frac{\partial \hat{m}}{\partial t} = -s(Z-\hat{Z}), \tag{9}$$

where we choose $s = 30$ for specificity and for which we observe good convergence results. The *Ansatz* system Eq. (9) was chosen after testing multiple forms. The parameter equations are evolved simultaneously with Eq. (8) with a forward Euler discretization and the same time step. The model form of Eq. (9) was chosen after testing several forms and there may exist other forms for which synchronization is possible. Once the model form was chosen, a parameter search was performed to find $s = 30$. As we vary s and κ, the synchronization manifold may lose stability, a common situation with diffusively coupled systems. Parameters may be updated as reaction-diffusion PDEs, by adding a diffusion term, however, we need to restrict parameters to be nonnegative $C^2(\Omega)$ functions and stability may be affected. To begin the simulation, parameters are arbitrarily initialized as the constant function

$$\hat{k}(x,y,0) = 5, \tag{10}$$

$$\hat{m}(x,y,0) = 5. \tag{11}$$

We evolve Eq. (7) forward and count the model output as observed data. Initial conditions for the response system are $\hat{P}(x,y,0) = 2$ and $\hat{Z}(x,y,0) = 2$. Furthermore, to avoid values outside the normal range of Eq. (7), we enforce that

$$\hat{P} = \begin{cases} \hat{P} & : \quad 0 < \hat{P} < 2 \\ 0 & : \quad \hat{P} \leq 0 \\ 2 & : \quad \hat{P} \geq 2 \end{cases} \quad \text{and} \quad \hat{Z} = \begin{cases} \hat{Z} & : \quad 0 < \hat{Z} < 2 \\ 0 & : \quad \hat{Z} \leq 0 \\ 2 & : \quad \hat{Z} \geq 2 \end{cases}$$

during the simulation.

First, we develop synthetic datasets with spatially varying parameters to challenge our methods. Spatially dependent parameters are chosen to be in the range given in Ref. 20 for spatially irregular behavior. Three different functional forms for the parameters are tested. First, a Gaussian parameter function is defined as

$$k_1(x,y) = ae^{-\left(\frac{(x-n/2)^2}{2\sigma^2}+\frac{(y-m/2)^2}{2\sigma^2}\right)},$$
$$m_1(x,y) = ce^{-\left(\frac{(x-n/2)^2}{2\sigma^2}+\frac{(y-m/2)^2}{2\sigma^2}\right)}, \quad (12)$$

where $a=2$, $c=0.6$, $m=300$, $n=900$, and $\sigma=400$. Appropriate parameters are chosen to maintain $m_1(x,y)$ and $k_1(x,y)$ in the target range. For example, Eq. (12) is displayed in Figures 1(a) and 1(b). Next, we define

$$k_2(x,y) = a\cos(bx+d)\sin(by)+s,$$
$$m_2(x,y) = c\cos(bx+d)\sin(by)+t, \quad (13)$$

where $a=0.2, b=\pi/(m/2), c=0.6, d=\pi/2, s=0.5$, and $t=1.5$, to test the quality of the autosynchronization scheme to resolve fine spatial structures in model parameters. The surfaces produced by Eq. (13) are displayed in Figures 1(c) and 1(d).

Finally, we build a swirly parameter function in order to simulate spiral-like behavior in parameter values as might be expected in turbulent near-coastal regions. We save a time-instance of phytoplankton density from the simulation of the original PDE, Eq. (7), scale it appropriately, and treat it as a parameter function. We refer to these spiral parameters as $k_3(x,y)$ and $m_3(x,y)$, and their forms are shown in Figures 1(e) and 1(f), respectively.

We observe solution data at every time step relative to the response system, Eq. (8), and the parameter system, Eq. (9), to drive $(\hat{P},\hat{Z}) \to (P,Z)$ and $(\hat{m}(x,y),\hat{k}(x,y)) \to (m(x,y),k(x,y))$ as $t\to\infty$. For brevity, only the parameters defined by Eqs. (12) and (13) are shown and compared with their estimated counterparts. We observe autosynchronization for each test set of parameters and the spatial inhomogeneities in each case are effectively resolved, demonstrated in Figures 2–5.

Figure 2 displays time-instances of the simulation of Eqs. (8) and (9), as the response, and Eq. (7), as the drive, with spatially dependent model parameters rendered by Eq. (12). Figure 2(a) shows the initial conditions, $P(x,\ y,\ 0)$ over $\hat{P}(x,y,0)$, Figure 2(c) represents $P(x,\ y,\ 1000)$ over $\hat{P}(x,y,1000)$, and Figure 2(e) represents $P(x,\ y,\ 4788)$ compared to $\hat{P}(x,y,4788)$. Figure 2(b) shows the initial conditions, $Z(x,\ y,\ 0)$ over $\hat{Z}(x,y,0)$, Figure 2(d) represents $Z(x,\ y,\ 1000)$ over $\hat{Z}(x,y,1000)$, and Figure 2(f) represents $Z(x,\ y,\ 4788)$ compared to $\hat{Z}(x,y,4788)$.

Likewise, Figure 3(a) shows the initial conditions, $k_1(x,y)$ over $\hat{k}(x,y,0)$, Figure 3(c) represents $k_1(x,y)$ over $\hat{k}(x,y,1000)$, and Figure 3(e) represents $k_1(x,y)$ compared to $\hat{k}(x,y,4788)$. Figure 3(b) shows the initial conditions, $m_1(x,y)$ over $\hat{m}(x,y,0)$, Figure 3(d) represents $m_1(x,y)$ over $\hat{m}(x,y,1000)$, and Figure 3(f) represents $m_1(x,y)$ compared to $\hat{m}(x,y,4788)$. Additional results will be presented with the same arrangement.

To challenge our methods in resolving spatial heterogeneities in parameters, we simulate Eq. (8) and (9), fed by the drive system Eq. (7) with parameters rendered by Eq. (13). Time-instances of this simulation are shown with species compared in Figure 4 and parameters compared in Figure 5. Time instances are displayed at $t=0$, $t=1000$, and $t=10\,660$. This simulation demonstrates the effectiveness of

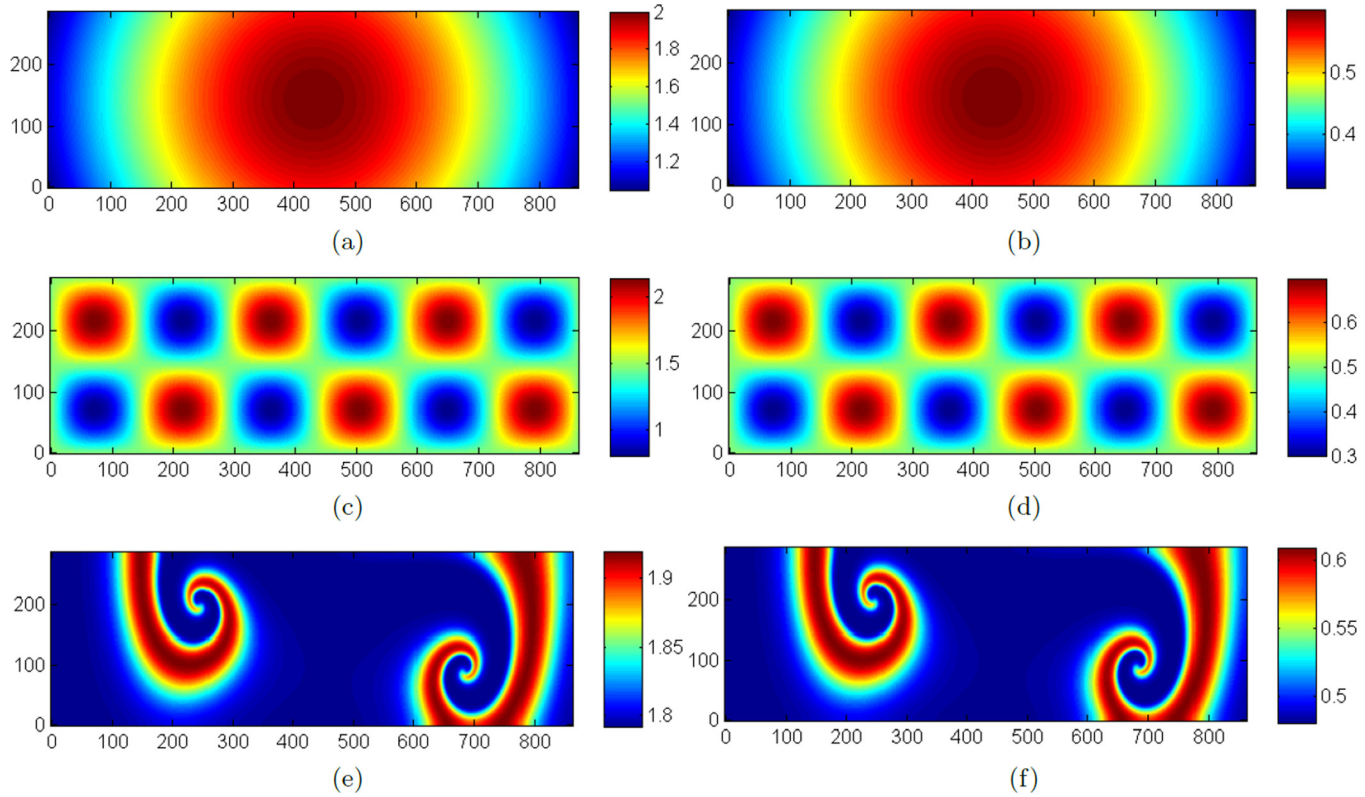

FIG. 1. Three sets of spatially dependent parameters used in simulations. Figures 1(a) and 1(b) are described by Eq. (12), with $k_1(x,y)$ on the left and $m_1(x,y)$ on the right. Below, with the same ordering, are the parameters described by Eq. (13). Finally, the swirly parameters are shown in Figures 1(e) and 1(f).

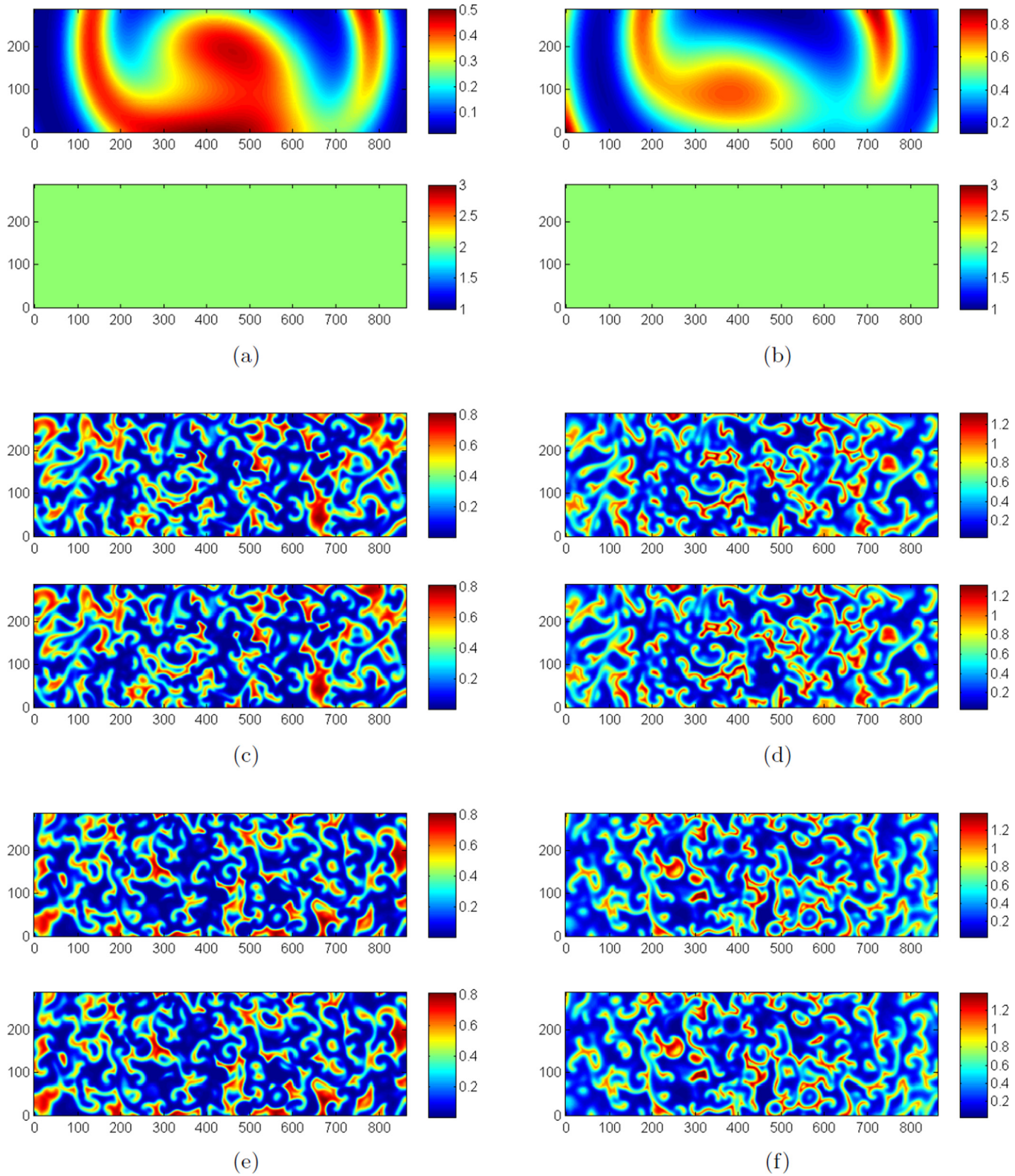

FIG. 2. Autosynchronization of species in Eqs. (8) and (9). Each figure shows drive (top) and response (bottom) pairs. $P(x, y, 0)$ and $\hat{P}(x, y, 0)$ in (a), $P(x, y, 1000)$ and $\hat{P}(x, y, 1000)$ in (c), and $P(x, y, 4788)$ and $\hat{P}(x, y, 4788)$ in (e). $Z(x, y, 0)$ and $\hat{Z}(x, y, 0)$ in (b), $Z(x, y, 1000)$ and $\hat{Z}(x, y, 1000)$ in (d), and $Z(x, y, 4788)$ and $\hat{Z}(x, y, 4788)$ in (f). Model parameters are $k_1(x, y)$ and $m_1(x, y)$.

parameter reconstruction. Similar results were found by testing parameters that vary spatially according to $m_3(x, y)$ and $k_3(x, y)$.

In the top row of Figure 6, the globally averaged relative error between the phytoplankton terms in the drive and response system is driven to less than 3.0×10^{-9} and the error between the zooplankton terms is driven below 2.0×10^{-8}. The globally averaged relative error between true and estimated parameters, both k and m, is driven below 1.0×10^{-5}. The top row corresponds to the simulation shown

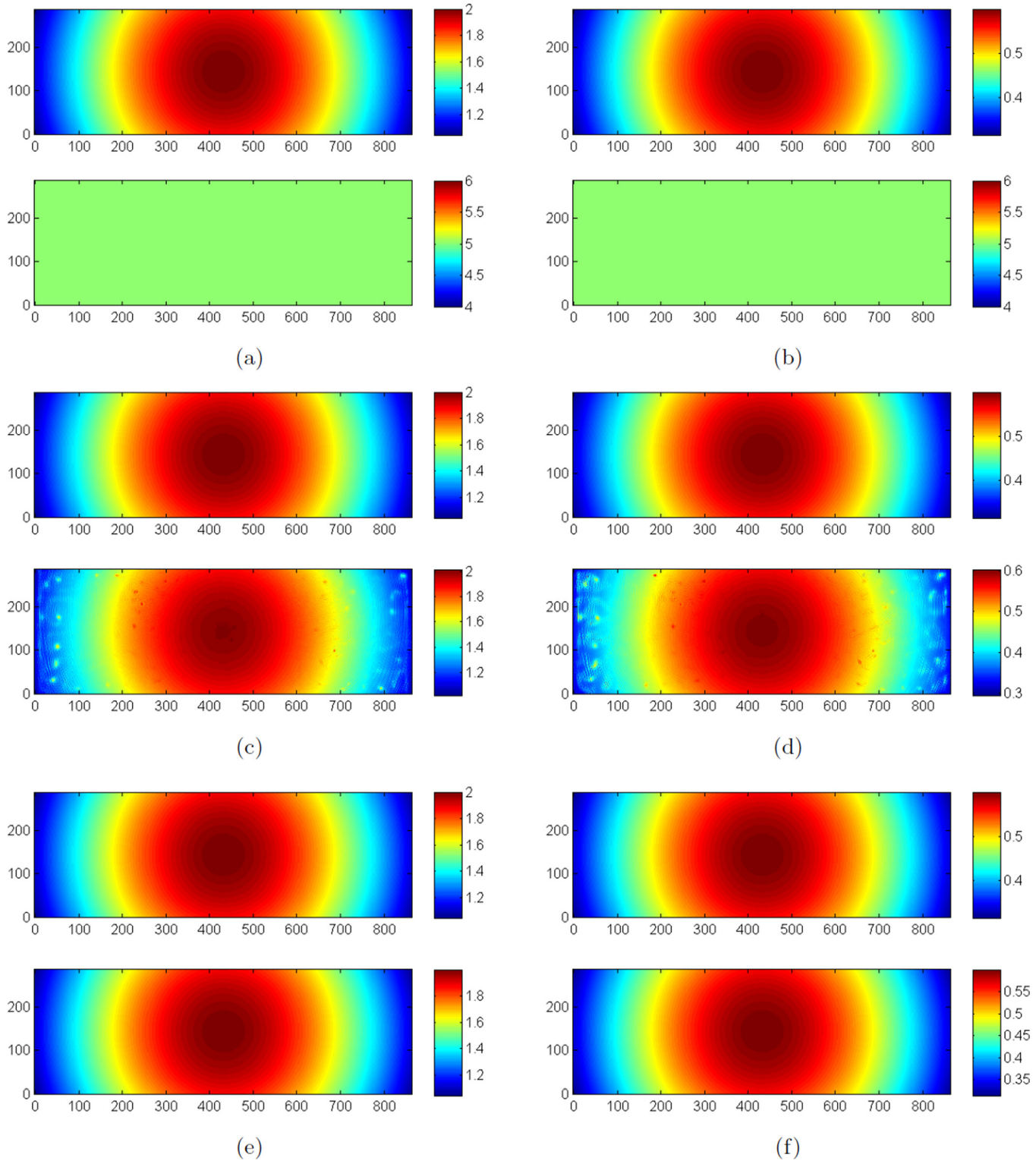

FIG. 3. Autosynchronization of response parameters in Eqs. (8) and (9). Each figure shows drive (top) and response (bottom) pairs. $k_1(x,y)$ and $\hat{k}(x,y,0)$ in (a), $k_1(x,y)$ and $\hat{k}(x,y,1000)$ in (c), and $k_1(x,y)$ and $\hat{k}(x,y,4788)$ in (e). $m_1(x,y)$ and $\hat{m}(x,y,0)$ in (b), $m_1(x,y)$ and $\hat{m}(x,y,1000)$ in (d), and $m_1(x,y)$ and $\hat{m}(x,y,4788)$ in (f).

in Figures 2 and 3, with parameters built using Eq. (12). In the bottom row of Figure 6, globally averaged relative errors are plotted corresponding to the simulations shown in Figures 4 and 5, with parameters corresponding to Eq. (13). All simulations shown here were terminated once parameter errors fell below 1.0×10^{-5}.

IV. SYNCHRONIZATION BY SAMPLING ONLY ONE SPECIES

To this point, an important criticism of our work is that we need to sample both species to drive the response model and parameters. As mentioned above, our interest in

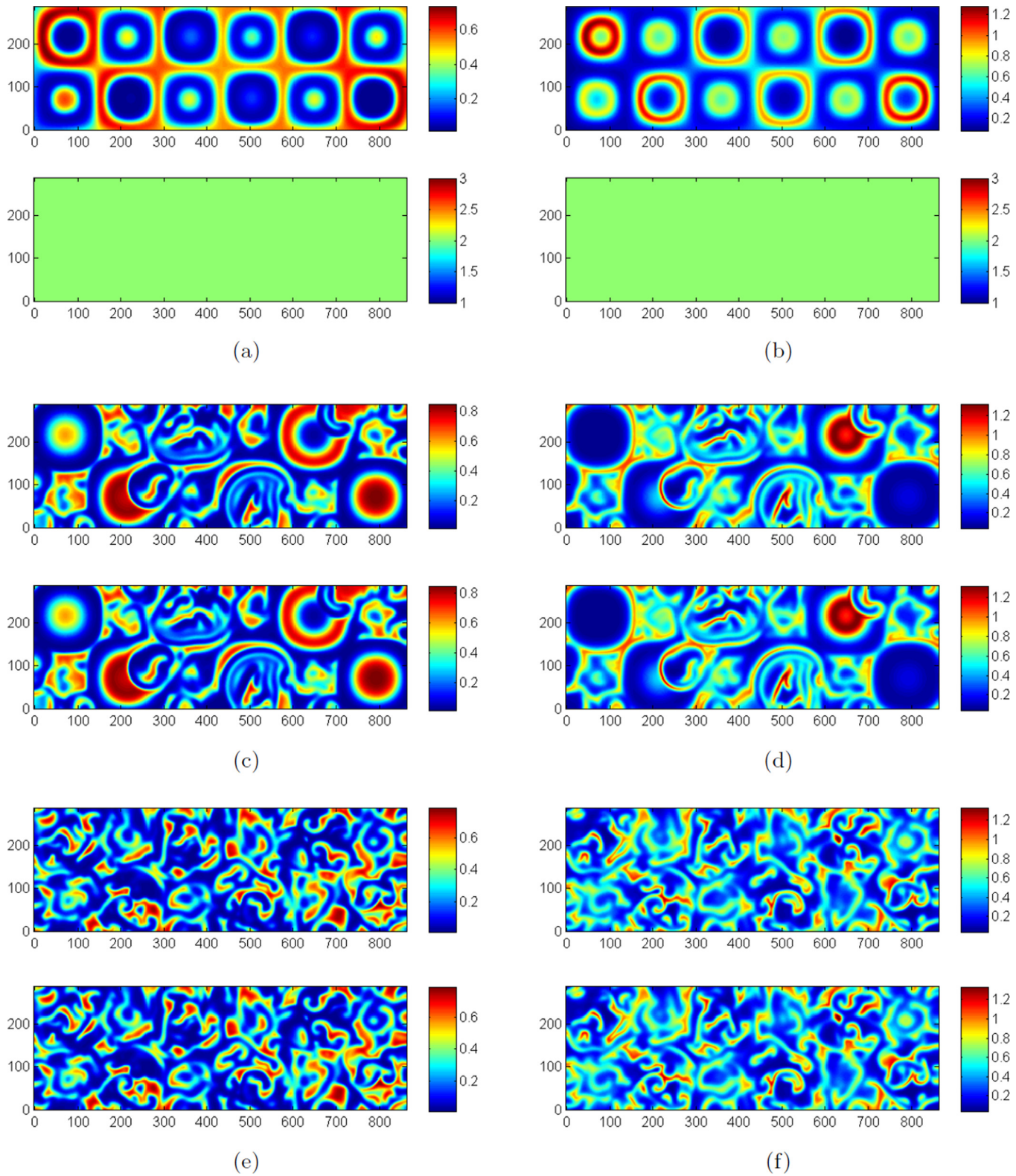

FIG. 4. Autosynchronization of species in Eqs. (8) and (9). Each figure shows drive (top) and response (bottom) pairs. $P(x, y, 0)$ and $\hat{P}(x, y, 0)$ in (a), $P(x, y, 1000)$ and $\hat{P}(x, y, 1000)$ in (c), and $P(x, y, 10\,660)$ and $\hat{P}(x, y, 10\,660)$ in (e). $Z(x, y, 0)$ and $\hat{Z}(x, y, 0)$ in (b), $Z(x, y, 1000)$ and $\hat{Z}(x, y, 1000)$ in (d), and $Z(x, y, 10\,660)$ and $\hat{Z}(x, y, 10\,660)$ in (f). Model parameters are $k_2(x, y)$ and $m_2(x, y)$.

autosynchronization for parameter estimation stems from work with ocean models for phytoplankton-zooplankton ecology. In fact, hyperspectral satellite imagery provides phytoplankton density inferences but provides no data for zooplankton. In this case, parameter estimation using the response model above will fail since we have no zooplankton observables with which to drive the response. Even given correct model parameters, it is impossible to forecast the model since zooplankton initial conditions are not supplied. Our problem of interest requires that we somehow estimate

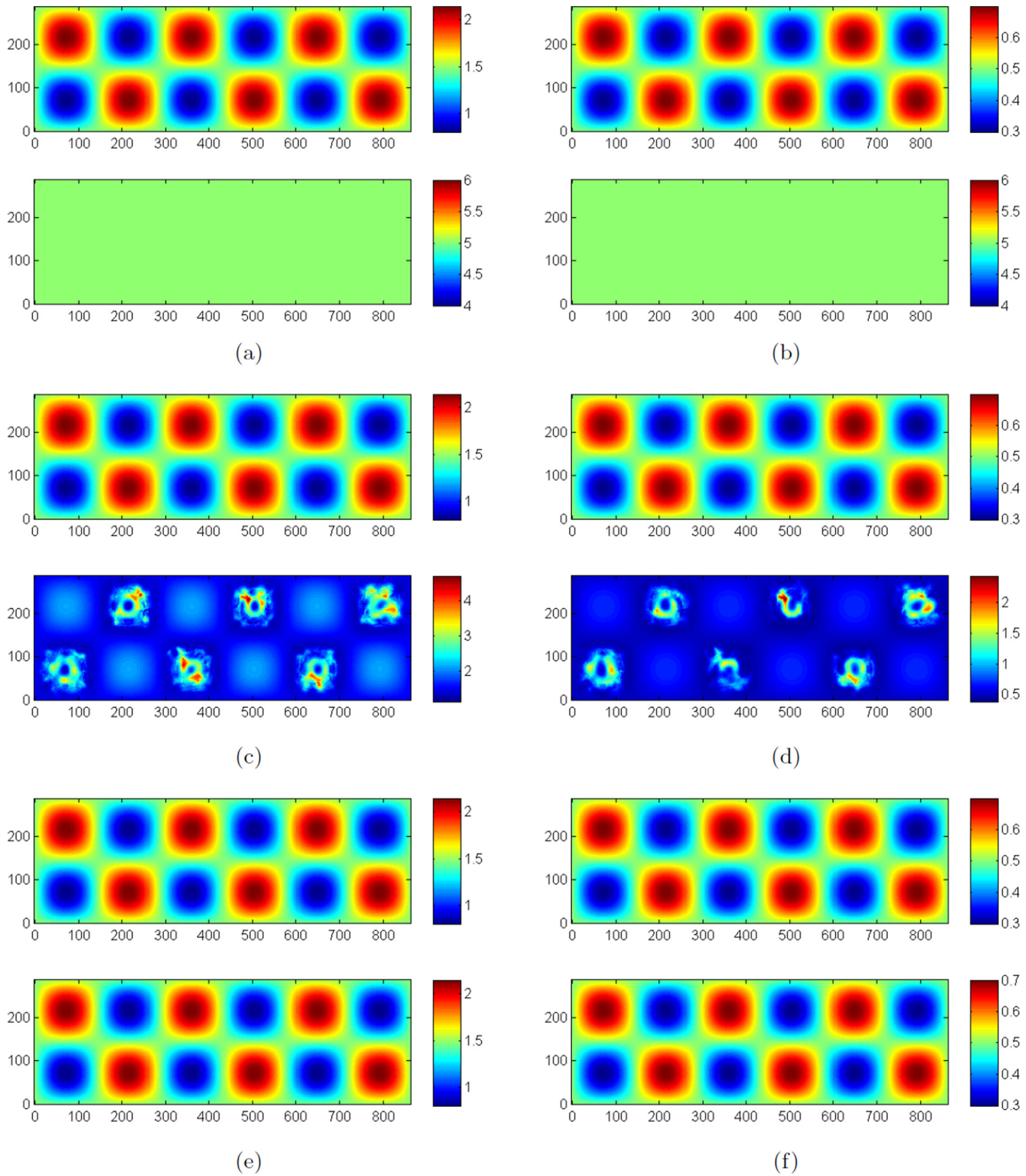

FIG. 5. Autosynchronization of response parameters in Eqs. (8) and (9). Each figure shows drive (top) and response (bottom) pairs. $k_2(x,y)$ and $\hat{k}(x,y,0)$ in (a), $k_2(x,y)$ and $\hat{k}(x,y,1000)$ in (c), and $k_2(x,y)$ and $\hat{k}(x,y,10660)$ in (e). $m_2(x,y)$ and $\hat{m}(x,y,0)$ in (b), $m_2(x,y)$ and $\hat{m}(x,y,1000)$ in (d), and $m_2(x,y)$ and $\hat{m}(x,y,10660)$ in (f).

zooplankton initial conditions based on phytoplankton observations.

We find that, by a modification of Eq. (8), it is possible to drive zooplankton density to its true state by sampling phytoplankton alone. This is a first demonstration of the possibility of simulating this system with only partial knowledge. As an added bonus we observe autosynchronization. Thus, this technique gives us a tool to estimate parameters and to initialize a model for short term forecasts. The response model that drives these results is

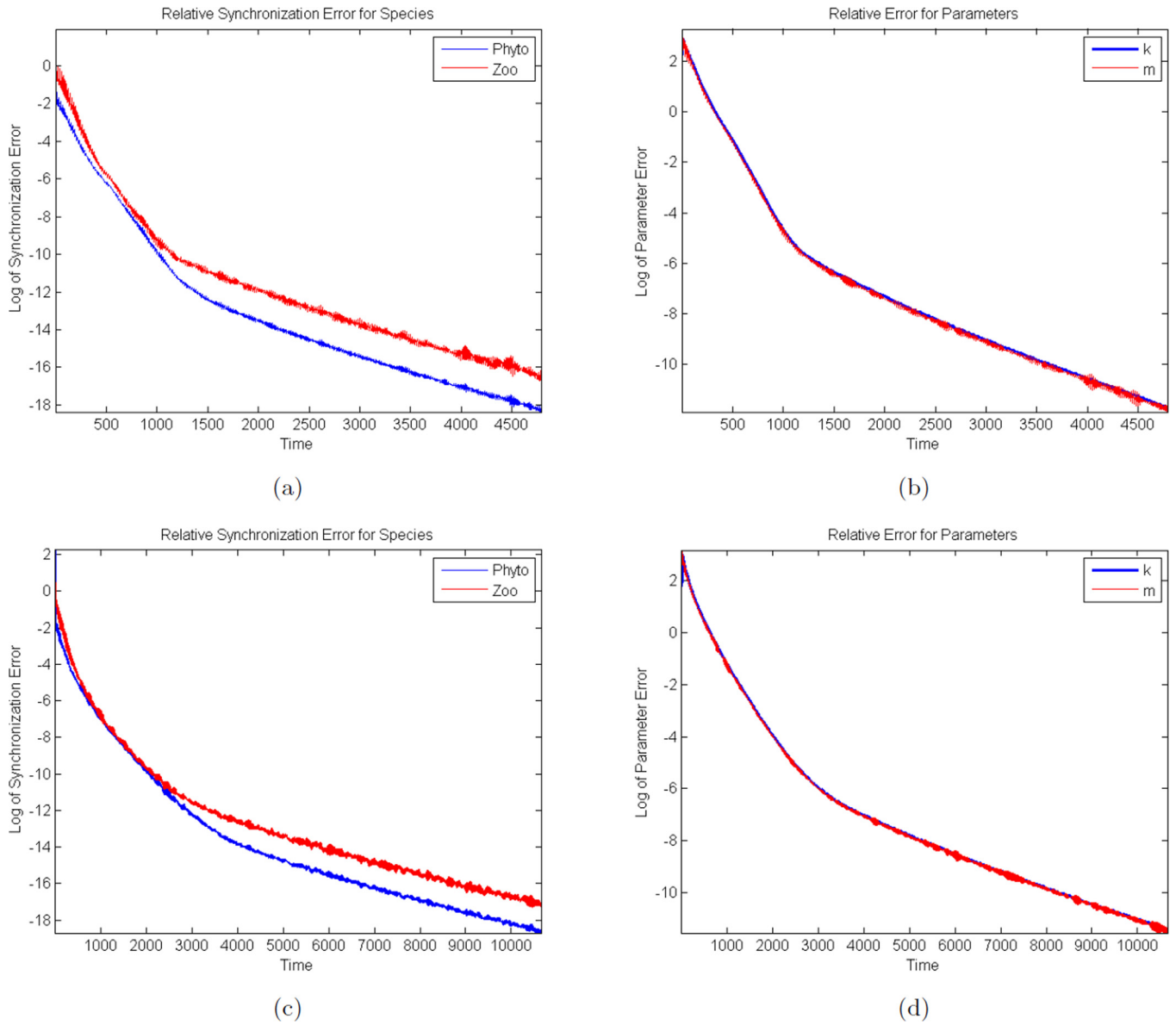

FIG. 6. Globally averaged relative synchronization error between drive and response PDE components and parameters on a log scale. Figures 6(a) and 6(b) correspond to parameters built by Eq. (12) and simulation displayed in Figures 2 and 3, respectively. Figures 6(c) and 6(d) show globally averaged relative synchronization error for species and parameters built by Eq. (13), corresponding to simulations in Figures 4 and 5, respectively.

$$\begin{aligned}\frac{\partial \hat{P}}{\partial t} &= \triangle \hat{P} + \hat{P}(1-\hat{P}) - \frac{\hat{P}\hat{Z}}{\hat{P}+h} + \kappa(P-\hat{P}),\\ \frac{\partial \hat{Z}}{\partial t} &= \triangle \hat{Z} + \hat{k}\frac{\hat{P}\hat{Z}}{P+h} - \hat{m}\hat{Z},\end{aligned} \tag{14}$$

where the absence of "hats" denotes where observation data are coupled directly into the PDE. That is, we use a combination of diffusive and complete replacement coupling in the response PDE to observe autosynchronization. Note that zooplankton density is no longer observed in Eq. (14). The parameter update equations are

$$\frac{\partial \hat{k}}{\partial t} = s_1(P-\hat{P}), \quad \frac{\partial \hat{m}}{\partial t} = s_2(P-\hat{P})\hat{P}, \tag{15}$$

with $s_1 = 0.2$, $s_2 = 0.6$, and $\kappa = 1.45$. Again, the *Ansatz* systems Eqs. (14) and (15) were chosen after testing multiple forms. As above, the parameter equations are evolved simultaneously with Eq. (14) using a forward Euler discretization. Figure 7 shows results obtained simulating Eqs. (14) and (15), driven by Eq. (7), at three time-instances, $t = 0$, $t = 1000$ and $t = 9360$. Results obtained for a spiral parameter form, which simulates perhaps more realistic mesoscale parameters, show that both $\hat{k}(x,y,t)$ and $\hat{m}(x,y,t)$ converge to their true values. Similar results are found with uniform random initial conditions defined for the response system (Figure 8).

In Figure 9, globally averaged errors are shown to diminish over time as the coupled systems evolve. Both parameter errors drop to within about 1.0×10^{-7}, the threshold at which the simulation is terminated, by $t = 9360$. Importantly, we note zooplankton synchronization error drops to within 2.0×10^{-6} of ground truth. Therefore, we need not sample zooplankton to observe autosynchronization

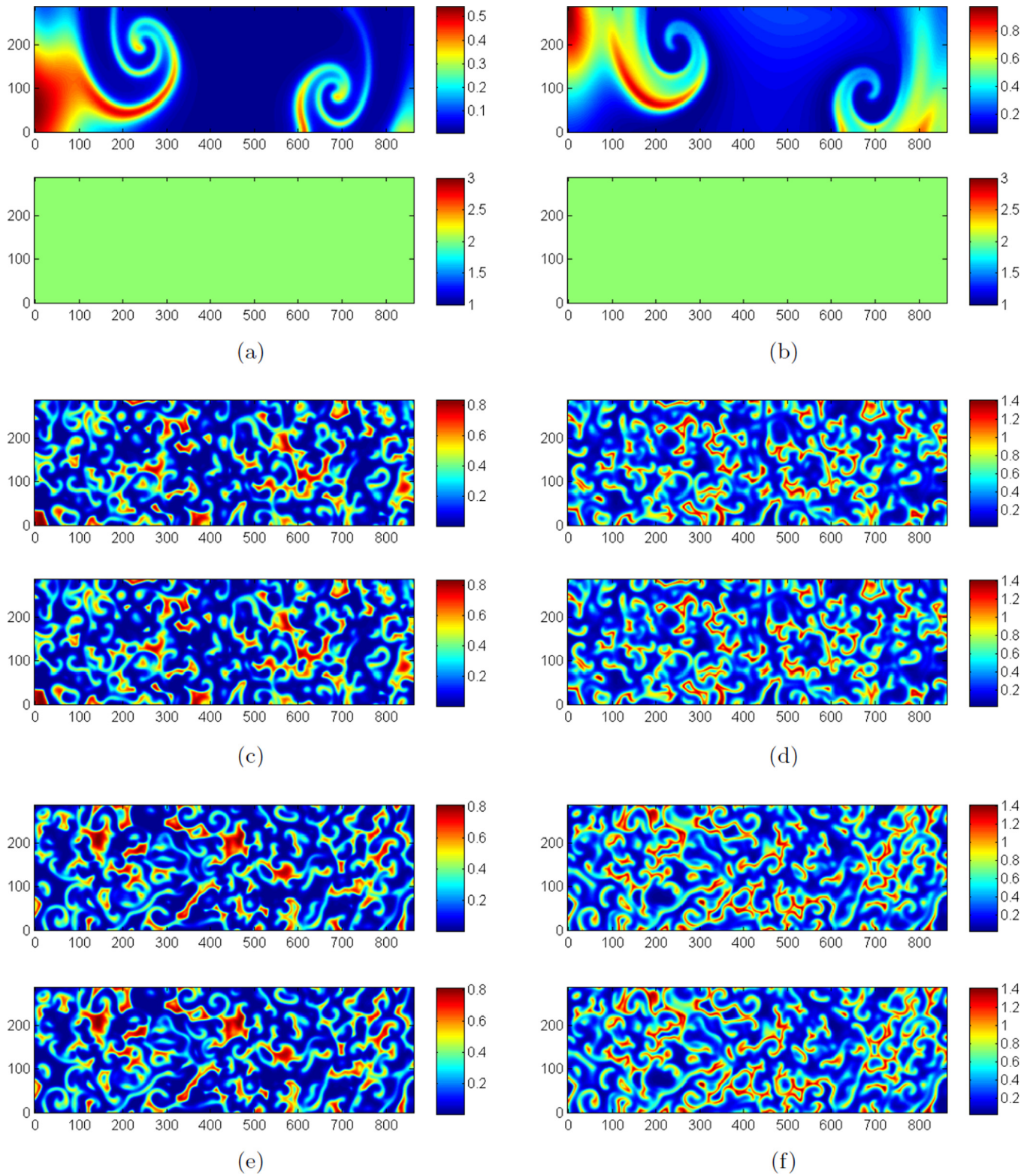

FIG. 7. Autosynchronization of species in Eqs. (14) and (15). Each figure shows drive (top) and response (bottom) pairs. $P(x, y, 0)$ and $\hat{P}(x, y, 0)$ in (a), $P(x, y, 1000)$ and $\hat{P}(x, y, 1000)$ in (c), and $P(x, y, 9360)$ and $\hat{P}(x, y, 9360)$ in (e). $Z(x, y, 0)$ and $\hat{Z}(x, y, 0)$ in (b), $Z(x, y, 1000)$ and $\hat{Z}(x, y, 1000)$ in (d), and $Z(x, y, 9360)$ and $\hat{Z}(x, y, 9360)$ in (f). Model parameters are those shown in Figures 1(e) and 1(f).

and we find the true zooplankton density profile and model parameters such that model simulations may be initialized. Furthermore, we are able to estimate swirly parameters as might be expected for spatially dependent parameters in the ocean on the mesoscale.

V. ON INCOMPLETE OBSERVATION DATA

Since we wish to apply these methods to model ocean ecology based on remote sensing, in particular hyperspectral satellite imagery, we now consider an innate complication

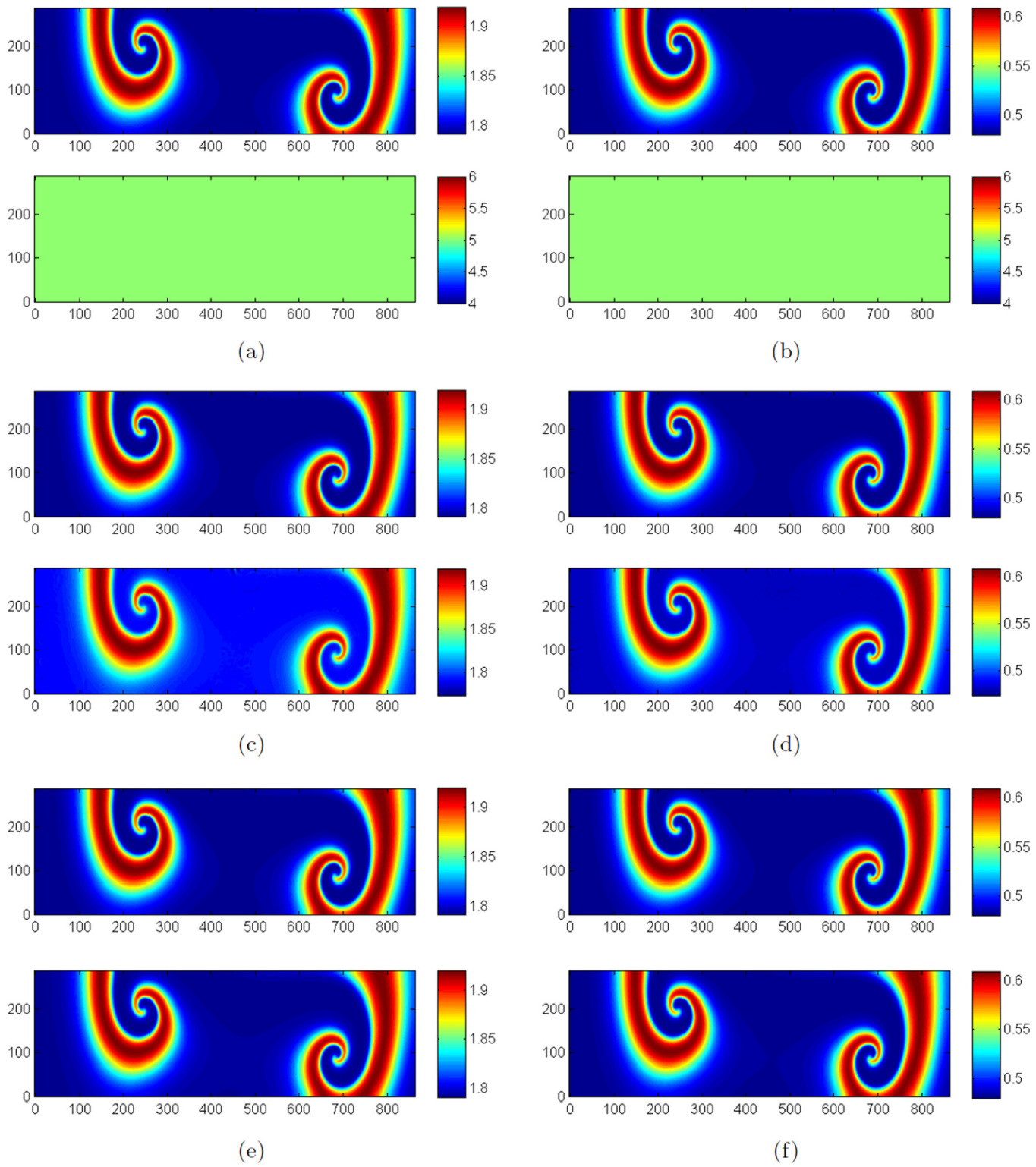

FIG. 8. Autosynchronization of parameters in Eqs. (14) and (15). Each figure shows drive (top) and response (bottom) pairs. $k_3(x,y)$ and $\hat{k}(x,y,0)$ in (a), $k_3(x,y)$ and $\hat{k}(x,y,1000)$ in (c), and $k_3(x,y)$ and $\hat{k}(x,y,9360)$ in (e). $m_3(x,y)$ and $\hat{m}(x,y,0)$ in (b), $m_3(x,y)$ and $\hat{m}(x,y,1000)$ in (d), and $m_3(x,y)$ and $\hat{m}(x,y,9360)$ in (f). Model parameters are those shown in Figures 1(e) and 1(f).

with such data. In real experimental situations, data are sampled at discrete locations, providing perhaps local averages on a coarsened subset of the domain. In order to apply the methods discussed directly one might subsample the observed data by interpolation to the desired grid. Interpolation could perform well, given the assumption of solutions in $W^2(\Omega)$. Otherwise, our method needs to be adapted for sampling by local averaging. Therefore, we now show that our techniques are robust by partially observing only phytoplankton in patches throughout the domain. We sample a coarser subset of the discretized domain and take local averages to be the driving signal, in the same way as

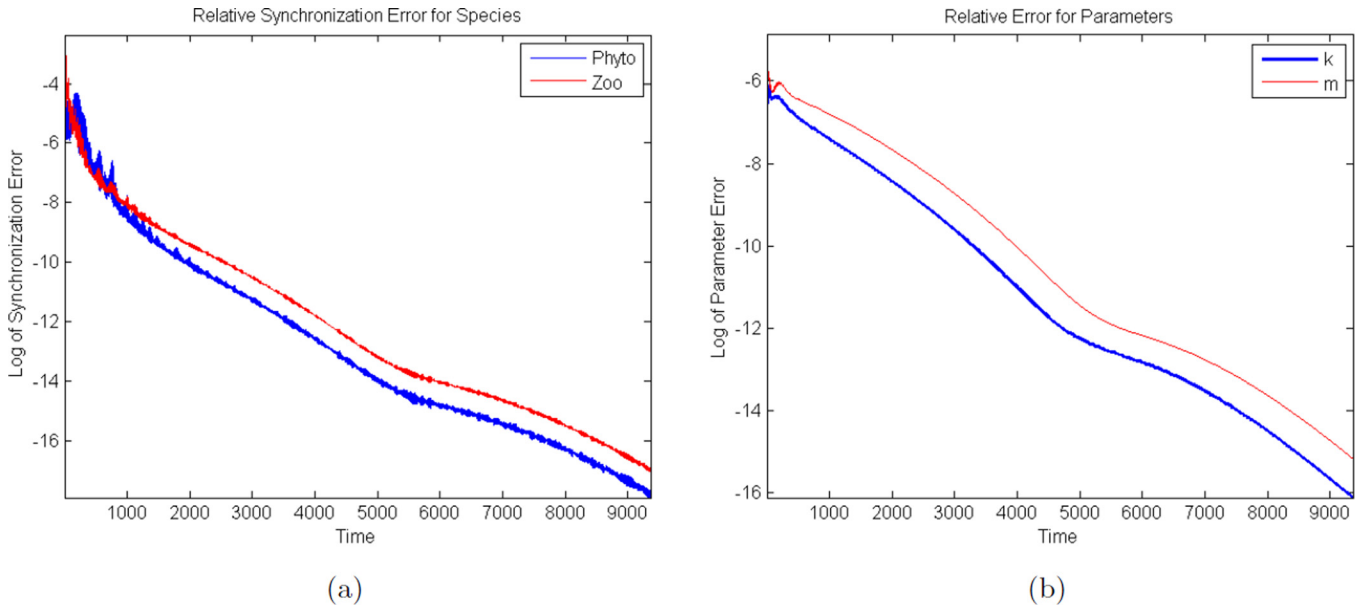

FIG. 9. Globally averaged relative synchronization error between drive and response PDE components and parameters on a log scale, estimating perhaps more realistic spiral parameters. Figures (a) and (b) correspond to parameters shown in Figures 1(e) and 1(f) and simulation displayed in Figures 7 and 8, respectively.

Ref. 16. The domain over which we sample is shown in Figure 10, where the domain is sampled in 3×3 patches with a spacing of 3 grid points between patches.

Synchronization is robust to spatial subsampling by local averaging by changing the response system to include only diffusive coupling, shown as follows:

$$\begin{aligned} \frac{\partial \hat{P}}{\partial t} &= \triangle \hat{P} + \hat{P}(1-\hat{P}) - \frac{\hat{P}\hat{Z}}{\hat{P}+h} + \kappa G_n \quad \forall x, y \in S_n, \\ \frac{\partial \hat{Z}}{\partial t} &= \triangle \hat{Z} + \hat{k}\frac{\hat{P}\hat{Z}}{\hat{P}+h} - \hat{m}\hat{Z}, \end{aligned} \tag{16}$$

where the complete replacement term is removed in the second equation. Therefore,

$$G_n(t) = \frac{1}{(dx)(dy)} \sum_{x,y \in S_n} (P(x,y,t) - \hat{P}(x,y,t)), \tag{17}$$

where S_n represents the rectangular "sensor" on the domain over which the model misfit is locally averaged. A requirement for good results is that we remove the direct replacement term in Eq. (16) since local averaging forces a slight misfit from observed data. Thus, the complete replacement term eventually works against identical synchronization and identical synchronization is not observed. Instead, we either remove the term as shown above or we allow complete replacement until the response output stops progressing toward drive observations. We then remove the complete replacement term and simply couple as shown in Eq. (16). This allows for faster synchronization from random initial conditions. Synchronization results are shown in Figure 11, wherein three different arrangements of local averaging are tested. In all three cases, the synchronization manifold is asymptotically stable, however, the rate of convergence to the manifold acts inversely with respect to sampling sparsity.

These results agree with previous results shown in Ref. 16, however, we now address autosynchronization. Thus, we modify the response system Eqs. (14) and (15)

$$\begin{aligned} \frac{\partial \hat{P}}{\partial t} &= \triangle \hat{P} + \hat{P}(1-\hat{P}) - \frac{\hat{P}\hat{Z}}{\hat{P}+h} + \kappa G_n \quad \forall x, y \in S_n, \\ \frac{\partial \hat{Z}}{\partial t} &= \triangle \hat{Z} + \hat{k}\frac{\hat{P}\hat{Z}}{\tilde{P}+h} - \hat{m}\hat{Z}, \\ \frac{\partial \hat{k}}{\partial t} &= \triangle \hat{k} + s_1(\tilde{P} - \hat{P}), \\ \frac{\partial \hat{m}}{\partial t} &= \triangle \hat{m} + s_2(\tilde{P} - \hat{P})\hat{P}, \end{aligned} \tag{18}$$

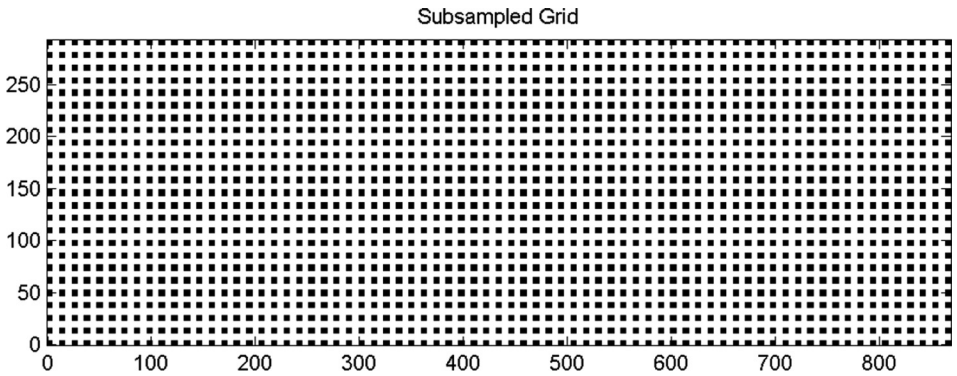

FIG. 10. Locally averaged patches over which drive system is sampled shown in black. Sampled on subset of 3×3 grid points with a distance of 3 grid points between patches.

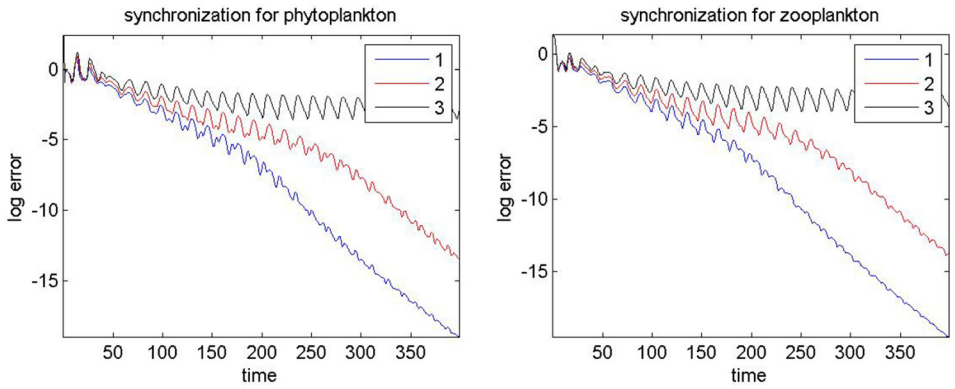

FIG. 11. Comparison of three different sampling schemes. Shown are relative synchronization errors between drive and response systems for sampling over 3×3 grid points (blue) with a distance of 3 grid points between subsequent patches, 2×2 grid points (red) with a distance of 2 grid points between subsequent patches, and 1×1 grid points (black) with a distance of 1 grid points between subsequent patches. Phytoplankton synchronization errors on left and zooplankton synchronization errors shown on right.

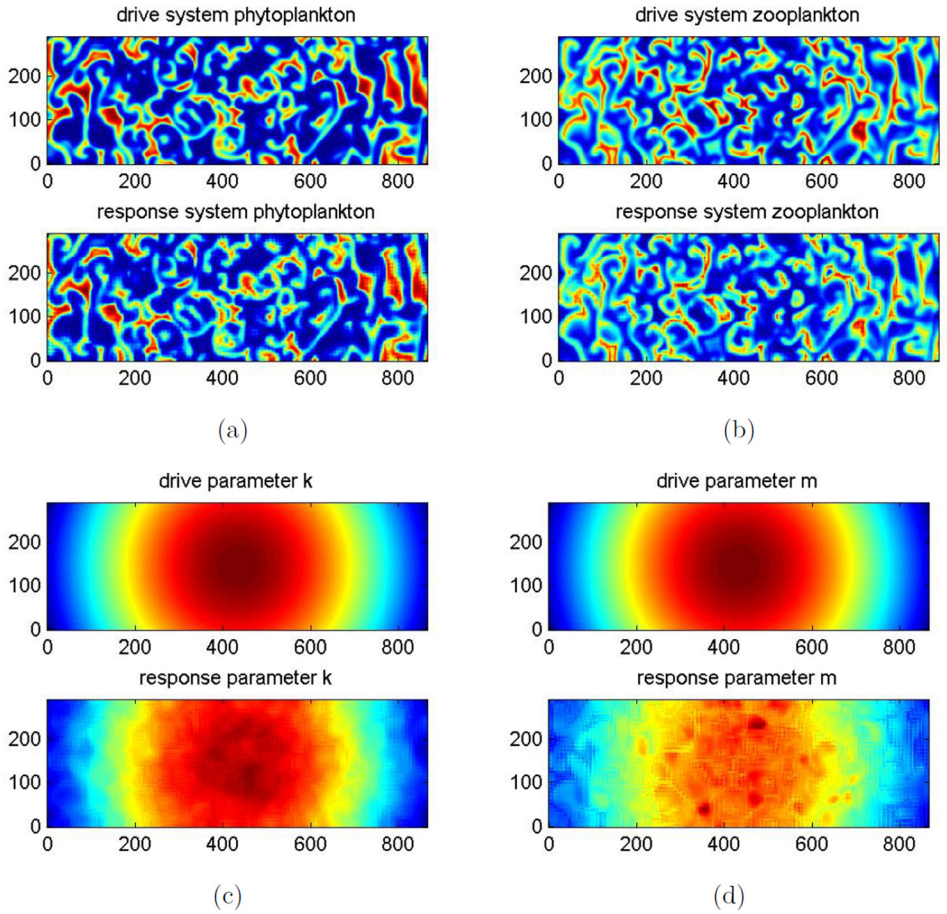

FIG. 12. Autosynchronization results shown at $t = 2000$. Both species and both parameters shown compared with drive species and true parameters. Effect of adding diffusion to parameter equations is clearly visible in estimated parameters.

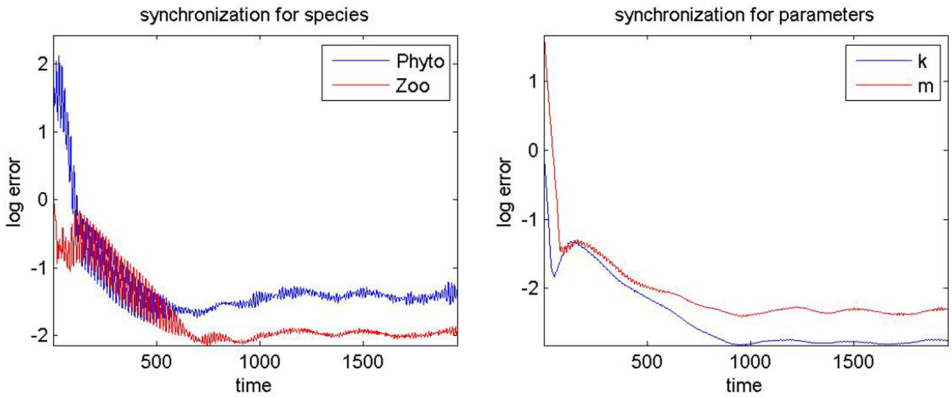

FIG. 13. Globally averaged relative synchronization errors shown for species and parameters. Local sampling destroys stability of the identical synchronization manifold, however, spatial characteristics of parameters are still observed.

where $\tilde{P}$ represents locally averaged observations from the drive system. Note, diffusion is added to the parameter equations in Eq. (18) in order that data from the driven regions, S_n will diffuse into the occluded regions. Varying sensor sizes were examined and in Figure 12 results are shown assuming a 2×2 subsampling of the domain Ω, with 1 grid point between subsequent patches.

In Figure 12, we see that despite locally averaged data, the response system is driven toward identical synchronization. Results are clearly not as good as sampling at every grid point in the domain, nor as good as the results for synchronization alone. Current work is aimed at finding how much data can be occluded from the observable set before autosynchronization completely fails. We also aim to address the considerable problem of cloud coverage over a region observed by satellite imagery. Clearly, results obtained are not as accurate when the locally averaged sampling is sparse. We note, however, that our satellite data are sufficiently fine such that interpolation or local averaging for simulation on a finer grid is unnecessary. That is, for our application to remote sensing, we evolve ecology over the same grid on which the data are observed, as described in Sec. IV. Nevertheless, an important consideration for application of these methods is missing observables due to cloud cover and sparsity in observables by local averaging (Figure 13).

VI. CONCLUSION

In this paper, we have shown that it is possible to derive an autosynchronization scheme for a system of PDEs. We emphasize here the improvements we have made upon past synchronization methods in that we use autosynchronization as a means of parameter estimation of parameters that exist in a function space. We assume prior knowledge of the model form of the observed system, but have no prior knowledge of the parameters. By sampling at every time step, we observed identical synchronization between the response and drive systems as described in Ref. 16. As a first attempt, we have given a model form for adaptive parameters in the response system such that we observe identical synchronization between response model parameters and true parameters, or autosynchronization. Our techniques were implemented on a benchmark model and estimates converge to ground truth. Thus, autosynchronization is observed for PDEs with scalar parameters.

Next, we considered the same system of PDEs wherein the parameters were spatially dependent. We provided a scheme with which we observe autosynchronization of spatially dependent parameters. We tested our results against several different functional forms for parameters and found the method to be robust.

We markedly improved upon these results once more with an autosynchronization scheme that requires sampling of only one species (phytoplankton). We noted that in order to evolve a system of PDEs for forecasting, we need initial conditions for both species; this is a serious problem when dealing with remote sensing data with which we can only observe one of the species. This concern was addressed by providing a response system that autosynchronizes parameters and synchronizes zooplankton using only phytoplankton data. These methods are plausible for use in remote sensing problems.

As discussed above, synchronization schemes can be proven to work for a given range of coupling parameters using, for example, Lyapunov functions. It remains to be shown why this scheme works on this system, and to perhaps derive autosynchronization model forms for a wider class of reaction-diffusion PDEs.

A drawback of this technique with application to hyperspectral satellite data is that data may be noisy; this is where filtering techniques have a built-in advantage. Data may also be occluded because of cloud cover. It has been shown in Ref. 16 that synchronization is possible even when subsampling the domain. Therein, samples are taken to be local averages representing a subset of the domain. By driving only on coarse subset of the original domain, we show that synchronization for our system is robust to spatial subsampling as long as subsampling is not too coarse. We further show that autosynchronization is somewhat robust to subsampling by local averaging, but not as robust as synchronization. Sampling only one species, however, required that we add diffusion into the parameter update equations to transfer information across unobservable regions of the domain. To this end, further work is to be done to improve upon these results, including perhaps interpolation of observations. By interpolating observations, diffusion could be removed again from the parameter equations. However, we have not had the same success for autosynchronization with large amounts of simply connected occluded data. Since there is no hope to synchronize PDEs without large quantities of data relative to the domain, we require techniques to fill in that which is missing, for example, inpainting. We address this problem in future work.

A problem with applying these techniques to satellite imagery is temporal data resolution. There may be several days between successive images and autosynchronization requires ample data observations. The need for frequent observables is perhaps the main drawback to this method. However, autosynchronization may be advantageous for parameter estimation or model building in situations where spatiotemporal data are abundant, and especially where parameters are expected to vary spatially.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Grant No. N00014-09-1-0647. The authors would also like to thank the anonymous referees for helpful comments and suggestions on the manuscript.

[1]S. Schiff and T. Sauer, "Kalman filter control of a model of spatiotemporal cortical dynamics," BMC Neurosci. **9**(Suppl 1), O1 (2008).
[2]J. D. Annan, J. C. Hargreaves, N. R. Edwards, and R. Marsh, "Parameter estimation in an intermediate complexity earth system model using an ensemble kalman filter," Ocean Modell. **8**(1), 135–154 (2005).
[3]E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000* (IEEE, 2000), pp. 153–158.

[4]T. G. Müller and J. Timmer, "Parameter identification techniques for partial differential equations," Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 2053–2060 (2004).
[5]T. G. Müller and J. Timmer, "Fitting parameters in partial differential equations from partially observed noisy data," Physica D **171**(1–2), 1–7 (2002).
[6]I. M. Navon, "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography," Dyn. Atmos. Oceans **27**(1), 55–79 (1998).
[7]L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," Phys. Rev. Lett. **64**(8), 821–824 (1990).
[8]U. Parlitz, "Estimating model parameters from time series by autosynchronization," Phys. Rev. Lett. **76**(8), 1232–1235 (1996).
[9]U. Parlitz, L. Junge, and L. Kocarev, "Synchronization-based parameter estimation from time series," Phys. Rev. E **54**, 6255 (1996).
[10]T. Stojanovski, L. Kocarev, and U. Parlitz, "A simple method to reveal the parameters of the lorenz system," Int. J. Bifurcation Chaos **6**(12B), 2645–2652 (1996).
[11]W. Yu, G. Chen, J. Cao, J. Lu, and U. Parlitz, "Parameter identification of dynamical systems from time series," Phys. Rev. E **75**, 067201 (2007).
[12]D. Yu and U. Parlitz, "Estimating parameters by autosynchronization with dynamics restrictions," Phys. Rev. E **77**(066221), 066221-1–7 (2008).
[13]F. Sorrentino and E. Ott, "Using synchronization of chaos to identify the dynamics of unknown systems," e-print arXiv:0909.2926.
[14]J. C. Quinn, P. H. Bryant, D. R. Creveling, S. R. Klein, and H. D. I. Abarbanel, "Parameter and state estimation of experimental chaotic systems using synchronization," Phys. Rev. E **80**(1), 016201 (2009).
[15]J. Schumann-Bischoff, J. Schröder-Schetelig, S. Luther, and U. Parlitz, "Estimating parameters and hidden variables of cardiac cell models from time series," *Proceedings of Biosignal 2010*, Berlin, Germany, Vol. 1, July 14–16, 2010.
[16]S. Berg, S. Luther, and U. Parlitz, "Synchronization based system identification of an extended excitable system," Chaos **21**(3), 033104 (2011).
[17]J. Schumann-Bischoff and U. Parlitz, "State and parameter estimation using unconstrained optimization," Phys. Rev. E **84**(5), 056214 (2011).
[18]H. D. I. Abarbanel, D. R. Creveling, R. Farsian, and M. Kostuk, "Dynamical state and parameter estimation," SIAM J. Appl. Dyn. Syst. **8**(4), 1341–1381 (2009).
[19]L. Kocarev, Z. Tasev, and U. Parlitz, "Synchronization spatiotemporal chaos of partial differential equations," Phys. Rev. Lett. **79**(1), 51–54 (1997).
[20]A. Medvinski, S. Petrovskii, I. Tikhonova, H. Malchow, and B. Li, "Spatiotemporal complexity of plankton and fish dynamics," SIAM Rev. **44**(3), 311–370 (2002).
[21]H. Malchow, B. Radtke, M. Kallache, A. B. Medvinsky, D. A. Tikhonov, and S. V. Petrovskii, "Spatio-temporal pattern formation in coupled models of plankton dynamics and fish school motion," Nonlinear Anal.: Real World Appl. **1**(1), 53–67 (2000).
[22]A. Mukhopadhyay, P. K. Tapaswi, and J. Chattopadhyay, "A space-time state-space model of phytoplankton allelopathy," Nonlinear Anal.: Real World Appl. **4**(3), 437–456 (2003).
[23]H. Malchow, F. M. Hilker, R. R. Sarkar, and K. Brauer, "Spatiotemporal patterns in an excitable plankton system with lysogenic viral infection," Math. Comput. Modell. **42**(9), 1035–1048 (2005).
[24]R. Kumar Upadhyay, N. Kumari, and V. Rai, "Exploring dynamical complexity in diffusion driven predator–prey systems: Effect of toxin producing phytoplankton and spatial heterogeneities," Chaos, Solitons Fractals **42**(1), 584–594 (2009).
[25]F. Rao, W. Wang, and Z. Li, "Spatiotemporal complexity of a predator–prey system with the effect of noise and external forcing," Chaos, Solitons Fractals **41**(4), 1634–1644 (2009).
[26]C. S. Holling, "Some characteristics of simple types of predation and parasitism," Can. Entomol. **91**(7), 385–398 (1959).