Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data

Charles T. Perrett, Stephan B. Munch, and George Sugihara

Accurate predictions of species abundance remain one of the most vexing challenges in ecology. This observation is perhaps unsurprising, because population dynamics are often strongly forced and highly nonlinear. Recently, however, numerous statistical techniques have been proposed for fitting highly parameterized mechanistic models to complex time series, potentially providing the machinery necessary for generating useful predictions. Alternatively, there is a wide variety of comparatively simple model-free forecasting methods that could be used to predict abundance. Here we pose a rather conservative challenge and ask whether a correctly specified mechanistic model, fit with commonly used statistical techniques, can provide better forecasts than simple model-free methods for ecological systems with noisy nonlinear dynamics. Using four different control models and seven experimental time series of flour beetles, we found that Markov chain Monte Carlo procedures for fitting mechanistic models often converged on best-fit parameterizations far different from the known parameters. As a result, the correctly specified models provided inaccurate forecasts and incorrect inferences. In contrast, a model-free method based on state-space reconstruction gave the most accurate short-term forecasts, even while using only a single time series from the multivariate system. Considering the recent push for ecosystem-based management and the increasing call for ecological predictions, our results suggest that a flexible model-free approach may be the most promising way forward.

Understanding fluctuations in species abundance is a long-standing goal of ecology and is particularly important for the conservation of severely depleted populations (1, 2). Dramatic changes in species abundance are common (3) and large declines can have disastrous impacts on ecosystem users (4). Accurate forecasts could allow for improved conservation efforts and increased fishery yields, however, ecological surprises are common (5), and accurate predictions remain a major challenge (6).

As ecological time series continue to lengthen, and computer-intensive statistics become more convenient, there has been an increasing trend toward fitting complex mechanistic models that incorporate both process and observation error (i.e., state-space models) (7). A Bayesian fitting procedure for such a model typically involves computationally intensive Markov chain Monte Carlo (MCMC) sampling of the joint posterior probability distribution to find the best fitting parameters (8). Previous work concluded that misspecified mechanistic models can produce poor forecasts (9), but it has been suggested that state-space models may be more robust to misspecification due to their inclusion of a process-error term (8); however, their forecast accuracy remains largely untested.

Model-free time series analysis provides an alternative method for generating forecasts that, unlike most state-space fitting procedures, does not require intensive computations. Although there exists a vast literature on time series forecasting methods, we focus on three commonly used models: the linear autoregressive (AR) and autoregressive moving average (ARMA) models, and the nonlinear state-space reconstruction (SSR) method. AR and ARMA time series models have been used extensively in ecology; recent examples include forecasting quasi-extinction risk (10), investigating the effects of climate change on fisheries (11), and detecting critical thresholds in ecosystem dynamics (12). SSR methods have been less pervasive in ecology (13, 14), although they are ubiquitous in many other scientific disciplines for describing stochastic nonlinear systems (15).

Here, we pose a rather conservative challenge and ask whether a correctly specified mechanistic model, fit with commonly used statistical techniques, can provide better forecasts than simple model-free methods for ecological systems with noisy nonlinear dynamics. To answer this question, we used computer simulations and experimental data to fit a suite of mechanistic models using a Bayesian MCMC algorithm under ecologically realistic levels of noise. Surprisingly, we found that even using only a single time series variable for each system, an SSR forecast approach consistently outperformed both the linear time series models and the correctly specified mechanistic models.

Simulation Methods

Using four different mechanistic control models, we generated simulated time series with both process noise and observation error (see Table 1 for models and parameters). Each model was fit to a 50-y time series (referred to as the training set) and forecast accuracy was evaluated using an additional 50-y time series that contained no observation error (referred to as the test set). Because each time series contains a unique sequence of process noise, we simulated 100 replicates for each model to ensure robust results. To facilitate comparison of forecast accuracy between models, forecast error is described using the standardized rms error statistic (SRMSE), which is defined as the rms error divided by the SD of the test set. Forecast errors greater than unity indicate predictions less accurate than simply predicting the mean of the test set.

All species (or age classes) in each model were forecasted. Importantly, however, to make the results more conservative, the model-free methods used only the single time series of the focal species (or age class) to make forecasts, whereas the mechanistic models were fit to the entire multivariate set of time series. In real systems we often lack time series of important driving variables or species; thus, for this additional reason, this comparison represents a very optimistic scenario for the mechanistic modeling approach.

Author contributions: C.T.P., S.B.M., and G.S. designed research; C.T.P. performed research; C.T.P. analyzed data; and C.T.P., S.B.M., and G.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. E-mail: cperrett@ucsd.edu.
This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216076110/-/DCSupplemental.
Here is the processed text:

Control Models. We chose four control models that encompass a wide range of ecological scenarios—the logistic model, a two-species coupled logistic model (16), an age-structured model (17), and a four-patch spatial model (18). Deterministic model structures are given in Table 1 (parameter values given in Table S1).

All control models included log-normal process and observation error, and parameter values that resulted in chaotic or near-chaotic deterministic dynamics. A coefficient of variation (CV) of 0.2 was chosen for the observation error, which is consistent with that reported in fish abundance surveys (19).

To avoid nonidentifiability of error sources, the process noise variance was fixed to the correct value (giving the mechanistic approach yet another advantage), and only the observation error variance was fit. Each control model was fit to its training set using Bayesian adaptive MCMC with Geyer’s algorithm (20) in MATLAB (21) (Table S2); similar results were obtained using R and JAGS (22, 23) (Figs. S1 and S2). All priors were independent, noninformative, truncated-normal distributions with mean equal to the simulations.

The expected value of the stochastic models closely matched that of their deterministic skeleton (Fig. S4); therefore, forecasts were made using the deterministic model with parameter values corresponding to the mode of the posterior distribution (similar results are obtained using the median). We initiated the MCMC search on the known parameters because likelihood surfaces of chaotic and near-chaotic time series are notoriously rugose (26), and we wanted to ensure that the known model parameters were evaluated; this is a luxury we lack when fitting models to real time series (and represents another advantage given to the approach). Also, in real systems, the control model is never known. Thus, the use of the control model here must be considered a best-case scenario for the mechanistic modeling approach and should be the ultimate advantage in this competition among models.

SSR Time Series Model. The SSR time series model uses Takens’ theorem, which allows for the representation of a multivariate dynamical system through the method of time-delay embedding (27). Takens’ theorem provides the conditions under which a multivariate dynamical system may be reconstructed from a univariate time series by transforming the univariate time series into a set of time-delayed vectors: \( x_t = [x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-(E-1)n}] \), where \( x \) is the univariate time series, \( t \) is the time index, \( r \) is the time lag, and \( E \) is the embedding dimension. Provided that \( E \) is sufficiently large, the collection of vectors \( X_t \), for \( t = 1 + (E-1)r, \ldots, N \), is an embedding of the attracting manifold, where \( N \) is the length of the time series.

After transforming the time series into an embedding, we generated predictions using the S-Map model (13, 28, 29), which is a locally weighted autoregressive model of the embedded time series with weights given by

\[
w(d) = \exp(-\theta d / \bar{d}) .
\]

where \( \theta \geq 0, d \) is the Euclidean distance between the predictee and the neighbor vector in embedded space, and \( \bar{d} \) is the average distance between the predictee and all other vectors. The degree of local weighting (\( \theta \)) and the embedding dimension (\( E \)) were fit using cross-validation on the 50-y training set. We searched to a maximum of \( E = 9 \) lags, because the highest dimensionality of the models was \( n = 4 \), and it has been shown that any nonlinear stochastic model can be represented by a univariate model of order \( 2n + 1 \), where \( n \) is the dimension of the true model (30).

Predictions were made on the 50-y test set.

Linear Time Series Models. Last, we evaluated the forecast accuracy of two linear time series models: AR and ARMA. Linear time series models have been extensively used in ecology. The ARMA(\( p, q \)) model is given as (31)

\[
(x_t - \mu) = \sum_{i=1}^{p} \beta_i (x_{t-i} - \mu) + \sum_{j=0}^{q} \alpha_j e_{t-j},
\]

where \( p \) is the dimension of the autoregressive term, \( q \) is the dimension of the moving average term, \( \mu \) is the mean of the time series, \( \alpha \) and \( \beta \) are the coefficients of the autoregressive and moving average terms, respectively, and \( e \) is a normally distributed random variable with mean zero.

It has been shown that many stochastic nonlinear processes, and especially low-order limit cycles, are well approximated by an ARMA model (31); however, the forecast accuracy of such models...
under large observation error and strong nonlinear dynamics (e.g., chaos) remains largely untested.

Analogous to the nonlinear model, the dimension of the AR and ARMA models was chosen using leave-one-out cross-validation on the 50-y training set, where the maximum dimension tested was 2n + 1 (i.e., the maximum possible value of p and q was 2n + 1). To differentiate between the performance of the AR and ARMA models, we only considered ARMA models that included a moving average component (i.e., parameter q was always greater than zero).

Simulation Results

Surprisingly, the SSR time series method consistently provided more accurate short-term forecasts than even the control models. In contrast, the linear forecasting methods were often no better than a prediction of the mean of the test set (represented by SRMSE = 1; Figs. 1 and 2). The performance of the control model varied depending on the system, but it too was often no better than a mean predictor.

The MCMC convergence diagnostics suggested that the control models often settled on best-fit parameters that were substantially different from the true parameters (Figs. S5 and S6). Often the best-fit parameters corresponded to stable (equilibrium) dynamics even though the true dynamics were chaotic; this resulted in poor forecast accuracy despite the correct underlying model structure.

For the logistic model and the two-species model, the SSR model provided one-step-ahead forecasts that were \( \sim 40\% \) better than a mean predictor. In contrast, the linear time series methods were often no better than a mean predictor, and the control model was often even worse (Fig. 1).

Similar results were obtained for the age-structured model. For the age 0 class, the short-term predictions of the SSR model were up to 50% better than the mean predictor, whereas the linear models and the control model were no better than a mean predictor. One-step-ahead predictions of the linear models for the age 1 and age 2 classes were consistently poor, whereas the control model prediction accuracy increased to that of the SSR method (Figs. 3 and 4).

The performance of the different forecasting methods varied widely for the spatial model. The SSR method again generated the most accurate predictions, with a forecast accuracy 70% better than the mean predictor. The control model provided short-term forecasts that were \( \sim 50\% \) better than the mean predictor, whereas the linear forecasting methods were no better than forecasting the mean (Figs. 3D and 4D).

Among the linear time series models, the simpler AR model provided a slightly better fit to the training set than the ARMA model (mean Akaike information criterion of \(-5.00 \) vs. \(-4.85 \) for the AR and ARMA models, respectively). Thus, the additional moving-average terms of the ARMA model were unable to capture the unexplained dynamics of the noisy, nonlinear models, which resulted in the slightly lower forecast error of the AR model (Figs. 1 and 3).

Experimental Data Methods

The holy grail of population biology is a model that both accurately predicts changes in abundance and can be used to anticipate critical thresholds. Through a battery of perturbation experiments and analyses of long time-series, this is exactly what was achieved in an extensive study of the flour beetle *Tribolium castaneum* (32). It was demonstrated that *Tribolium* population dynamics are well represented by a discrete, age-structured model of larvae, pupae, and adults (32–34), given as

\[
L_t = bA_{t-1} \exp(-c_0A_{t-1} - c_{cpa}A_{t-1}) \\
P_t = L_{t-1}(1 - \mu_1) \\
A_t = P_{t-1} \exp(-c_{pa}A_{t-1}) + A_{t-1}(1 - \mu_2),
\]

where \( L_t \) is the abundance of feeding larvae at time t; \( P_t \) is the abundance of large larvae, nonfeeding larvae, pupae, and callow adults; and \( A_t \) is the number of sexually mature adults. The parameter \( b \) is the number of larval recruits per adult per unit of time in the absence of cannibalism, and the fractions \( \mu_1 \) and \( \mu_2 \) are the larval and adult mortality rates, respectively. The exponential terms in the larva equation describe the probability that an egg will be eaten in the presence of \( L_{t-1} \) larvae and \( A_{t-1} \).
adults, and the exponential term in the adults’ equation describes the probability that a pupae will survive to become an adult when \( A_{t+1} \) adults are present.

We used time series from the Tribolium experiments described in Dennis et al. (32) in which adult recruitment rates (for a total of seven treatments) were experimentally manipulated, producing limit cycles and chaos. For each experimental treatment, the population was censused every 2 wk for 80 wk, resulting in a 41-point time series. Each experimental treatment was replicated twice, providing three independent time series for each treatment with which to compare the forecast accuracy of the different methods. For each treatment, one time series was randomly chosen to fit the models (the training set), and the other two time series were used to evaluate forecast accuracy (the test set), providing a total of 21 evaluations.

To emulate ecologically realistic conditions, we added log-normal observation error with a CV of 0.2 to all of the training sets. The forecast accuracy was evaluated using the test-set time series without observation error. The larvae-pupae-adult (LPA) model was fit using the Bayesian MCMC routine described above with the parameter search initiated on the parameter values estimated in Dennis et al. (32) and on the correct initial conditions.

**Experimental Results**

As with the simulation results, our analysis of the experimental data shows that the SSR time series method provided the most accurate short-term forecasts of all of the methods tested. When averaged over all treatments, short-term predictions (4 wk or less) made by the SSR method were significantly more accurate than the linear methods, and more accurate than the control model in all but one instance (Fig. 5). Short-term predictions (2 wk ahead) provided by the SSR method were routinely over 40% more accurate than a prediction of the mean of the time series (Fig. 5), whereas the linear methods were only 10% more accurate than a mean predictor.

Similar to the simulation results, although the MCMC routine for the control model had converged, it often converged on parameterizations that resulted in equilibrium dynamics when the true dynamics were either limit cycle or chaotic (Fig. 5S); this resulted in poor prediction accuracy for the control model. The control model outperformed the time series methods only for the one-step predictions of the pupae stage, which is unsurprising given that the one-step-ahead transition from larvae to pupae is determined by a single mortality parameter.

For the longest prediction interval (20 wk), the forecast accuracy of all methods converged to that of the time series mean. Although this was expected for the chaotic treatments, due to sensitive dependence on initial conditions, it was notable that it also occurred for deterministic limit cycles and is likely due to the interaction of nonlinear dynamics and process noise.

For the larvae and adult time series, both the nonlinear (SSR) and linear methods (AR and ARMA) were significantly more accurate than the control model up to the 10-step-ahead prediction horizon (Fig. 5A and C). The results for the pupae time series were similar, with the exception that the half-generation predictions of the control model were most accurate (Fig. 5B).

**Discussion**

Despite the absence of mechanistic information about the underlying ecological processes, the relatively simple SSR method consistently outperformed the control models over near-term prediction horizons. This result was robust across all simulations and all life stages of the experimental data. Moreover, the SSR model achieved this feat using only a single time series, whereas the control model used all time series simultaneously (it is an ecologically unrealistic scenario to assume we know the model and have time series for all of the relevant variables). Other analyses have shown that multivariate SSR methods (35) improve with additional information (14, 36), suggesting that the performance of the SSR model tested here represents a lower bound on forecast accuracy attainable with this general approach.

Overall, the control model forecast accuracy was similar to the SSR model for the linear components of the models. For example, for the linear (age 1 and age 2) components of the age-structured model, the one-step predictions of the pupae stage, which is unsurprising given that the one-step-ahead transition from larvae to pupae is determined by a single mortality parameter.

For the longest prediction interval (20 wk), the forecast accuracy of all methods converged to that of the time series mean. Although this was expected for the chaotic treatments, due to sensitive dependence on initial conditions, it was notable that it also occurred for deterministic limit cycles and is likely due to the interaction of nonlinear dynamics and process noise.

For the larvae and adult time series, both the nonlinear (SSR) and linear methods (AR and ARMA) were significantly more accurate than the control model up to the 10-step-ahead prediction horizon (Fig. 5A and C). The results for the pupae time series were similar, with the exception that the half-generation predictions of the control model were most accurate (Fig. 5B).

**Discussion**

Despite the absence of mechanistic information about the underlying ecological processes, the relatively simple SSR method consistently outperformed the control models over near-term prediction horizons. This result was robust across all simulations and all life stages of the experimental data. Moreover, the SSR model achieved this feat using only a single time series, whereas the control model used all time series simultaneously (it is an ecologically unrealistic scenario to assume we know the model and have time series for all of the relevant variables). Other analyses have shown that multivariate SSR methods (35) improve with additional information (14, 36), suggesting that the performance of the SSR model tested here represents a lower bound on forecast accuracy attainable with this general approach.

Overall, the control model forecast accuracy was similar to the SSR model for the linear components of the models. For example, for the linear (age 1 and age 2) components of the age-structured model, the one-step predictions of the pupae stage, which is unsurprising given that the one-step-ahead transition from larvae to pupae is determined by a single mortality parameter.

For the longest prediction interval (20 wk), the forecast accuracy of all methods converged to that of the time series mean. Although this was expected for the chaotic treatments, due to sensitive dependence on initial conditions, it was notable that it also occurred for deterministic limit cycles and is likely due to the interaction of nonlinear dynamics and process noise.

For the larvae and adult time series, both the nonlinear (SSR) and linear methods (AR and ARMA) were significantly more accurate than the control model up to the 10-step-ahead prediction horizon (Fig. 5A and C). The results for the pupae time series were similar, with the exception that the half-generation predictions of the control model were most accurate (Fig. 5B).

**Discussion**

Despite the absence of mechanistic information about the underlying ecological processes, the relatively simple SSR method consistently outperformed the control models over near-term prediction horizons. This result was robust across all simulations and all life stages of the experimental data. Moreover, the SSR model achieved this feat using only a single time series, whereas the control model used all time series simultaneously (it is an ecologically unrealistic scenario to assume we know the model and have time series for all of the relevant variables). Other analyses have shown that multivariate SSR methods (35) improve with additional information (14, 36), suggesting that the performance of the SSR model tested here represents a lower bound on forecast accuracy attainable with this general approach.

Overall, the control model forecast accuracy was similar to the SSR model for the linear components of the models. For example, for the linear (age 1 and age 2) components of the age-structured model, the one-step predictions of the pupae stage, which is unsurprising given that the one-step-ahead transition from larvae to pupae is determined by a single mortality parameter.
model, the short-term accuracy of the control model was similar to that of the SSR model, whereas the forecast accuracy of the control model for the nonlinear (age 0) component was significantly worse than the SSR model. Similar results were obtained for the experimental data because the control model provided the best forecasts for only the one-step-ahead predictions of the pupae stage, which is governed by a single linear mortality term. This finding suggests that the optimal forecasting model for age-structured populations may be one that combines a mechanistic model for the upper age classes with a time series model for the recruit dynamics.

For the cases considered here, we have shown that commonly used statistical procedures for fitting mechanistic models to chaotic time series will often converge on parameterizations that yield decadently nonchaotic dynamics. In these cases, models with the wrong parameters will likely yield wrong inferences, and measures such as extinction risk, maximum sustainable yield, or population growth rate (Fig. S5) can be severely biased. This observation has particularly important implications for fishery management where mechanistic model-fitting is often a critical element of the assessment process.

An important question remains: How chaotic is too chaotic for traditional Bayesian methods? To address this question, we performed additional simulations for the logistic model under both chaotic and nonchaotic parameterizations (r = 2.5–3.95). The stability of the model strongly influenced parameter estimation error. Error grew rapidly as the model became unstable (Lyapunov exponent near zero; Fig. S1), and estimates of the Lyapunov exponent were severely biased toward stability (Fig. S2). This bias was present even when the true model was marginally stable, suggesting that our results are relevant even for nonchaotic systems in the presence of process error; we speculate that this may explain the apparent paucity of chaos in ecological time series. For a more detailed treatment of this question, see Judd (37).

Alternative model-fitting procedures have been proposed that estimate goodness of fit using summary statistics rather than directly estimating parameter likelihood (38, 39). However, those methods lack general recommendations for designing informative statistics for arbitrary dynamic systems, and therefore have not been widely adopted in ecology. Another promising method involves fitting semiparametric functions to first-differenced data (40); however, that method is considerably more challenging to implement than the model-free methods described here, and it has not been widely tested or adopted.

As ecologists, our ultimate goal is to uncover the processes that govern natural systems. The ideal route to such discovery is through hypothesis testing via controlled experiments. Unfortunately, population-level hypotheses are impossible to test through such experiments, and moreover, ecosystem dynamics are often sufficiently complex that perturbation of any single variable is unlikely to yield great insight. Forecasting provides a way forward, because population-level hypotheses can be evaluated based on our ability to make accurate predictions. Importantly, however, even correct models with the correct parameters will fail to provide accurate long-term forecasts due to sensitive dependence on initial conditions and nonlinearity. Therefore, we would be foolish to forego the future work to determine how to use short-term forecasts, which can be validated empirically, to ensure sustainability in a changing environment.

ACKNOWLEDGMENTS. We thank Hao Ye, Ethan Deyle, and two anonymous reviewers for helpful comments on an earlier version of the manuscript. Funding support was provided by National Marine Fisheries Service/Sea Grant Population Dynamics Fellowship EIDP-9 and CAMEO Program Grant NA08OAR4320894, and National Science Foundation Grant DEB-1020372.


PNAS Early Edition | S of S

Perretti et al.
Supporting Information

Perretti et al. 10.1073/pnas.1216076110

Fig. S1. The logistic model was fit with values for parameter $r$ ranging from 2.5 to 3.9. For each value of $r$, we generated a 50-y time series using the observation and process noise levels described in the main text. The Markov chain Monte Carlo procedure was performed in JAGS (1) via R (2). Convergence was checked using batch-mean plots and the Gelman–Rubin statistic ($R < 1.2$). As the true model becomes more unstable (represented by a higher Lyapunov exponent), the estimation error increases. Estimation error for parameter $r$ is defined as $|\hat{r} - r_{true}|$.


Fig. S2. Estimated Lyapunov exponent vs. true Lyapunov exponent for the logistic model fit with parameter $r$ ranging from 2.5 to 3.9. For each value of $r$, we generated 50 points using the observation and process noise levels described in the main text. The red line is one to one. The failure of the Bayesian fitting routine is demonstrated for unstable values of $r$. The estimated model is always stable despite the true model being unstable. The Lyapunov exponent is always poorly estimated when the true Lyapunov exponent is positive.
Fig. S3. Example trace plot (A) and normal–normal plot (B) of 500-point batch means for the LPA model illustrating that the Markov chain Monte Carlo routine converged. The red points denote the range of the data and the end of the first and third quantiles.

Fig. S4. Time series expected values (100 y, 100 replicates) for the stochastic vs. deterministic version of each model illustrating the similarity between the expected value of the stochastic model and deterministic models. A–D show results for the logistic, two-species, age-structured, and spatial models, respectively.
Fig. S5. Example histograms of accepted Markov chain Monte Carlo draws (15,000 iterations) and true parameter values (red dots) for each parameter in the logistic model. $X_0$ is the initial condition, and $V_{obs}$ is the variance of the observation error.

Fig. S6. Example trace plots (Left) and normal–normal plots of 500-point batch means (Right) for the logistic (A and B), two-species (C and D), age-structured (E and F), and spatial (G and H) illustrating the convergence of the Markov chain Monte Carlo procedure. Despite initialization on the correct parameters, the models often converged on incorrect best-fit parameters.
Table S1. Model structure and model parameter values used in the simulations

<table>
<thead>
<tr>
<th>Model name</th>
<th>Process model structure*</th>
<th>Parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>$x_{t+1} = x_t \left(1 - \frac{x_t}{K}\right) \exp(\epsilon_t)$</td>
<td>$x_0 = 0.5, r = 3.7, K = 1$</td>
</tr>
<tr>
<td>Two-species</td>
<td>$x_{t+1} = x_t \left(1 - \frac{K_1}{x_t}\right) - c y_t \exp(\epsilon_t)$ [y_{t+1} = y_t \left(1 - \frac{K_2}{y_t}\right) + c x_t \exp(\epsilon_t)]</td>
<td>$x_0 = 0.4, y_0 = 0.4, r_1 = 3.8, r_2 = 3.7$ $K_1 = 1, K_2 = 1, c = 0.1$</td>
</tr>
<tr>
<td>Age-structured</td>
<td>$x_{0,t} = (x_{2,t} + x_{3,t})f$ [x_{1,t+1} = x_{0,t} \exp(\epsilon - x_{0,t} + \epsilon_t)] [x_{2,t+1} = s_1 x_{1,t}] [x_{3,t+1} = s_2(x_{2,t} + x_{3,t})]</td>
<td>$x_{1,0} = 0.5, x_{2,0} = 0.1, x_{3,0} = 0.1, f = 20$ $r = 3.0, s_1 = 0.05, s_2 = 0.04$</td>
</tr>
<tr>
<td>Spatial</td>
<td>$x_{t+1} = x_t \exp(r - x_t + \epsilon_t)$ [x_{t+1} = \sum_{j=1}^{4} d_{ij} x_{j,t}]</td>
<td>$R_0 = 5.1$ $d = \begin{bmatrix} 0.015 &amp; 0.019 &amp; 0.055 &amp; 0.041 \ 0.042 &amp; 0.021 &amp; 0.012 &amp; 0.026 \ 0.038 &amp; 0.066 &amp; 0.052 &amp; 0.031 \ 0.037 &amp; 0.047 &amp; 0.002 &amp; 0.069 \end{bmatrix}$</td>
</tr>
</tbody>
</table>

*Process noise $\epsilon_t$ is normally distributed with $\mu = 0$ and $\sigma = 0.005$.

Table S2. Parameter values and descriptions of nonmodel parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>100</td>
<td>Total number of time-steps for model simulations</td>
</tr>
<tr>
<td>$L$</td>
<td>50,000</td>
<td>Number of iterations in the Markov chain Monte Carlo (MCMC) routine</td>
</tr>
<tr>
<td>$B$</td>
<td>5,000</td>
<td>Number of iterations in the MCMC burn-in</td>
</tr>
<tr>
<td>$C$</td>
<td>3</td>
<td>Number of MCMC chains in Geyer’s algorithm</td>
</tr>
<tr>
<td>$CV_J$</td>
<td>0.005</td>
<td>Initial coefficient of variation of the adaptive MCMC proposal distribution</td>
</tr>
<tr>
<td>$I$</td>
<td>500</td>
<td>Number of iterations between adjustment of proposal coefficient of variation and covariance</td>
</tr>
<tr>
<td>$CV_P$</td>
<td>0.5</td>
<td>Coefficient of variation of the prior probability distribution for each parameter (independent truncated normal distribution with mean equal to the true parameter value)</td>
</tr>
</tbody>
</table>