The theory and practice of iterated filtering

SAMSI workshop on Sequential Monte Carlo methods

September 9, 2008

Ed Ionides

The University of Michigan, Department of Statistics

Inference for static parameters in state space models
(i.e., unknown model parameters that do not vary in time)

- Numerical issues have led to a considerable literature on this topic.
- **Iterated filtering** maximizes the likelihood function via taking an average of filtered “local” parameter estimates obtained by adding noise to the static parameters (which regularizes the numerical issues). This process is recursively repeated while reducing the added noise.
- Iterated filtering, implemented by basic SMC, has a **plug-and-play** property: it requires simulation from the state process but not transition densities.

Plug-and-play methods for state space models

- Statistical methods are **plug-and-play** if they require simulation from the dynamic model but not explicit likelihood ratios.

- Bayesian plug-and-play:

  1. Artificial parameter evolution (Liu and West, 2001)

  2. Approximate Bayesian computation (“sequential Monte Carlo without likelihoods,” Sisson et al, *PNAS*, 2007)

- Non-Bayesian plug-and-play:

  3. Simulation-based prediction rules (Kendall et al, *Ecology*, 1999)

  4. Maximum likelihood via iterated filtering (Ionides et al, *PNAS*, 2006)

Plug-and-play is a VERY USEFUL PROPERTY for investigating scientific models.

The cost of plug-and-play

- Approximate Bayesian methods and simulated moment methods lead to a loss of statistical efficiency.
- In contrast, iterated filtering enables (almost) exact likelihood-based inference.
- Improvements in numerical efficiency may be possible when analytic properties are available (at the expense of plug-and-play). But many interesting dynamic models are analytically intractable—for example, it is standard to investigate systems of ordinary differential equations numerically.

Artificial parameter evolution

- Write θ for the vector of static parameters. Set $\theta = \theta_t$ to be a random walk with

$$E[\theta_t|\theta_{t-1}] = \theta_{t-1} \qquad \mathrm{Var}(\theta_t|\theta_{t-1}) = \sigma^2$$

- We show how to carry out recursive filtering with a sequence $\sigma \to 0$ to maximize the likelihood.

An iterated filtering algorithm

Select $\hat{\theta}^{(1)}$, σ_1, c, α and N.

For n in $1, \ldots, N$

(i) set $\sigma = \sigma_1 \alpha^{n-1}$ and initialize $E[\theta_0^{(n)}] = \hat{\theta}^{(n)}$, $\mathrm{Var}(\theta_0^{(n)}) = c\sigma^2$.

(ii) For $t = 1, \ldots, T$, evaluate **filtering means** $\hat{\theta}_t^{(n)} = E[\theta_t^{(n)} | y_{1:t}]$ and **prediction variances** $V_{t,n} = \mathrm{Var}(\theta_t^{(n)} | y_{1:t-1})$.

(iii) $\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + V_{1,n} \sum_{t=1}^{T} V_{t,n}^{-1} (\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)})$

- **Each iteration updates $\hat{\theta}_n$ using an average of the filtering means with weights determined by the prediction variances.**
- **$\hat{\theta}^{(N+1)}$ converges to a (local) maximum of the likelihood as $N \to \infty$, under regularity conditions.**

Theorem 1. (Ionides, Bretó & King, *PNAS*, 2006)

Suppose $\hat{\theta}_0$, C *and* $y_{1:T}$ *are fixed and define*

$$\begin{aligned} \hat{\theta}_t &= \hat{\theta}_t(\sigma) = E[\theta_t | y_{1:t}] \\ V_t &= V_t(\sigma) = \mathrm{Var}(\theta_t | y_{1:t-1}) \end{aligned}$$

Assuming sufficient regularity conditions for a Taylor series expansion,

$$\lim_{\sigma\to 0} \sum_{t=1}^{T} V_t^{-1}(\hat{\theta}_t - \hat{\theta}_{t-1}) = (\partial/\partial\theta) \log f(y_{1:T}|\theta, \sigma{=}0)\Big|_{\theta=\hat{\theta}_0}$$

The limit of an appropriately weighted average of local filtered parameter estimates is the derivative of the log likelihood.

Theorem 2. (Ionides, Bretó & King, *PNAS*, 2006)

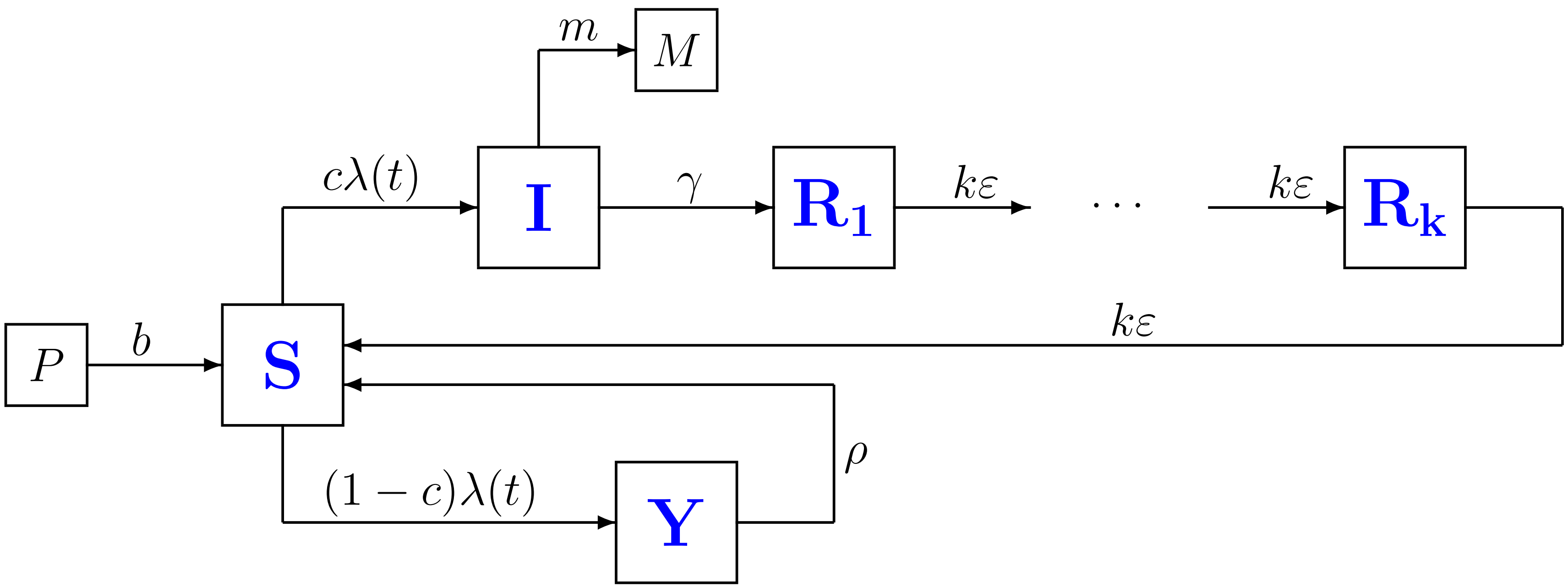
Set $\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + \sigma_n^2 M(\nabla\ell(\hat{\theta}^{(n)}) + \eta_n)$, where M is a positive definite symmetric matrix. Suppose the following:

1. $\ell(\theta)$ is twice continuously differentiable and uniformly convex.
2. $\lim_n \sigma_n^2 n^{1-\alpha} > 0$ for some $\alpha \in (0,1)$.
3. $\{\eta_n\}$ has $E[\eta_n] = o(1)$, $\mathrm{Var}(\sigma_n^2 \eta_n) = o(1)$, $\mathrm{Cov}(\eta_m, \eta_n) = 0$ for $m \neq n$.

If there is a $\hat{\theta}$ with $\nabla\ell(\hat{\theta}) = 0$ then $\hat{\theta}^{(n)}$ converges in probability to $\hat{\theta}$.

With appropriate assumptions, iterated filtering does converge to a local maximum if "cooled" sufficiently slowly.

Example: cholera (bacterial diarrhea caused by *Vibrio cholerae*)



A compartment model for cholera. Each individual is Susceptible (S), Infected and infectious (I), Asymptomatically infected (Y), or recovering ($R_1, \ldots, R_k$). Susceptibles enter at birth (from P), and infection may lead to death (M).

parameter	symbol
force of infection	$\lambda(t)$
probability of severe infection	c
recovery rate	γ
disease death rate	m
mean long-term immune period	$1/\varepsilon$
CV of long-term immune period	$1/\sqrt{k}$
mean short-term immune period	$1/\rho$

Cholera model: SDE driven by Gaussian noise $\xi(t)$

$$\frac{d}{dt}S(t) = k\epsilon\, R_k + \rho\, Y + \frac{d}{dt}P(t) + \delta\, P(t) - (\lambda(t) + \delta)\, S$$

$$\frac{d}{dt}I(t) = c\,\lambda(t)\, S - (m + \gamma + \delta)\, I(t)$$

$$\frac{d}{dt}Y(t) = (1 - c)\,\lambda(t)\, S - (\rho + \delta)\, Y$$

$$\frac{d}{dt}R_1(t) = \gamma\, I - (k\,\epsilon + \delta)\, R_1$$

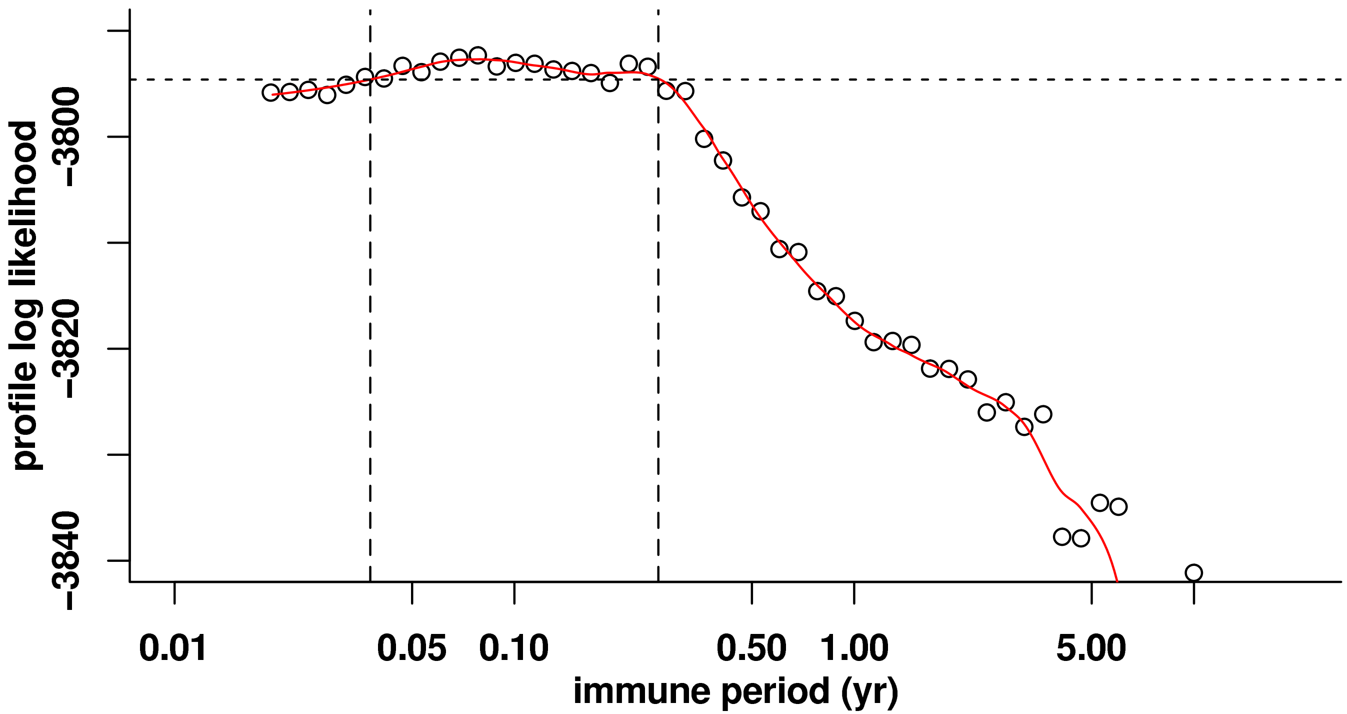
$$\vdots$$

$$\frac{d}{dt}R_k(t) = k\,\epsilon\, R_{k-1} - (k\,\epsilon + \delta)\, R_k$$

Stochastic force of infection, with periodic cubic spline $\beta_{seas}(t)$:

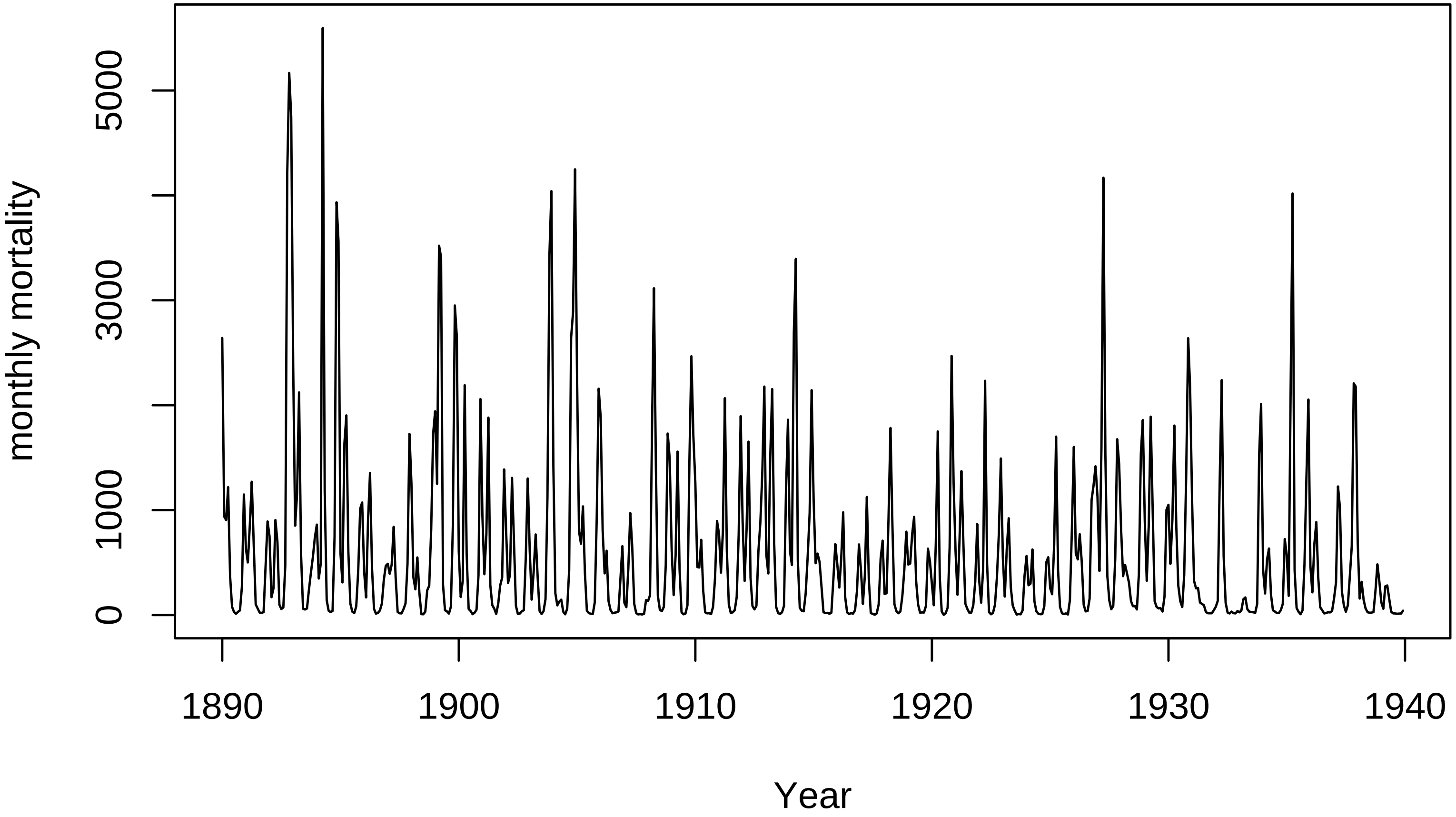
$$\lambda(t) = \left(e^{\beta_{trend}\, t}\, \beta_{seas}(t) + \xi(t)\right) \frac{I(t)}{P(t)} + \omega$$

Duration of immunity

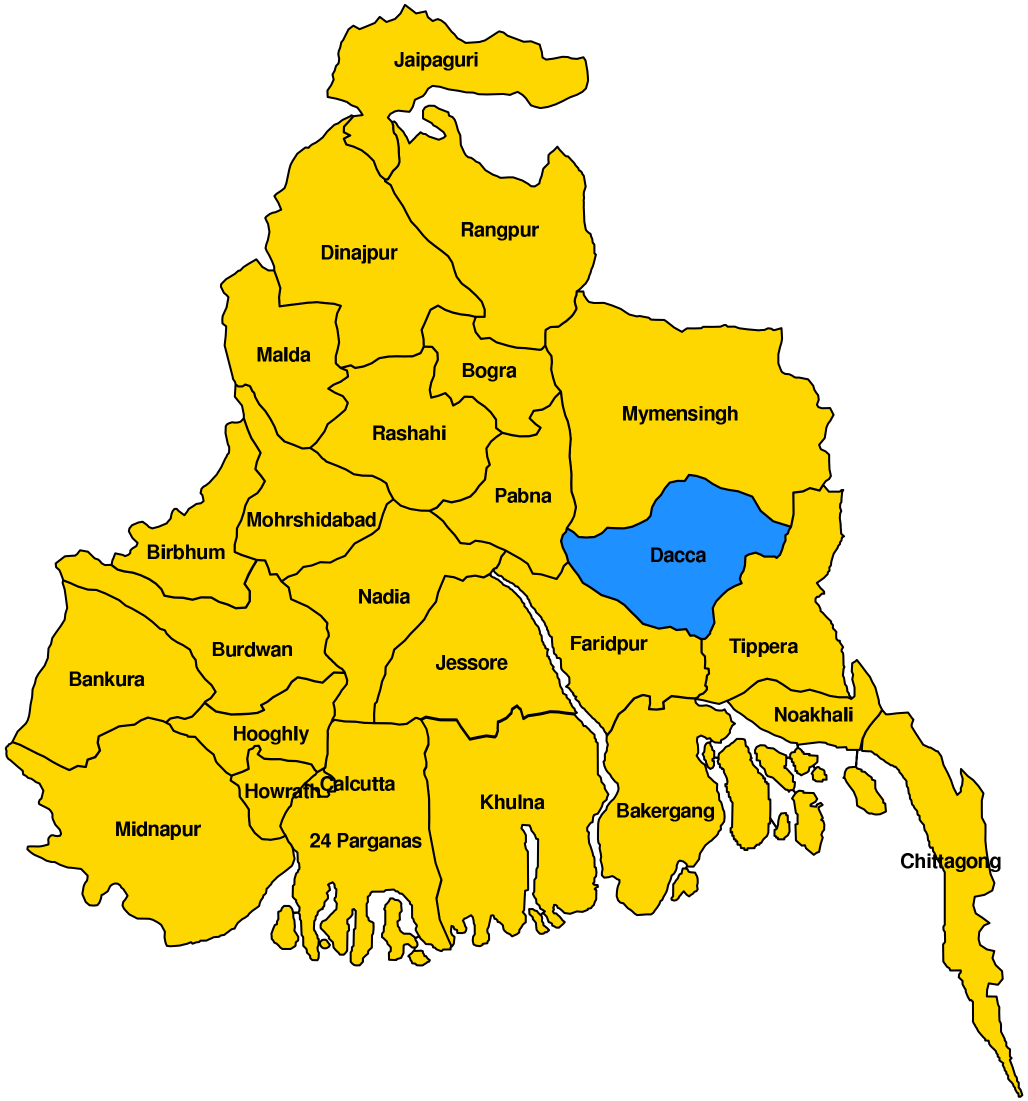


- Profile likelihood of immune period for historical time series data in Dacca, Bangladesh (similar results for other districts).
- **Conclusion: asymptomatic infections have short-term immunity with epidemiological consequences** (King et al, *Nature*, 2008).

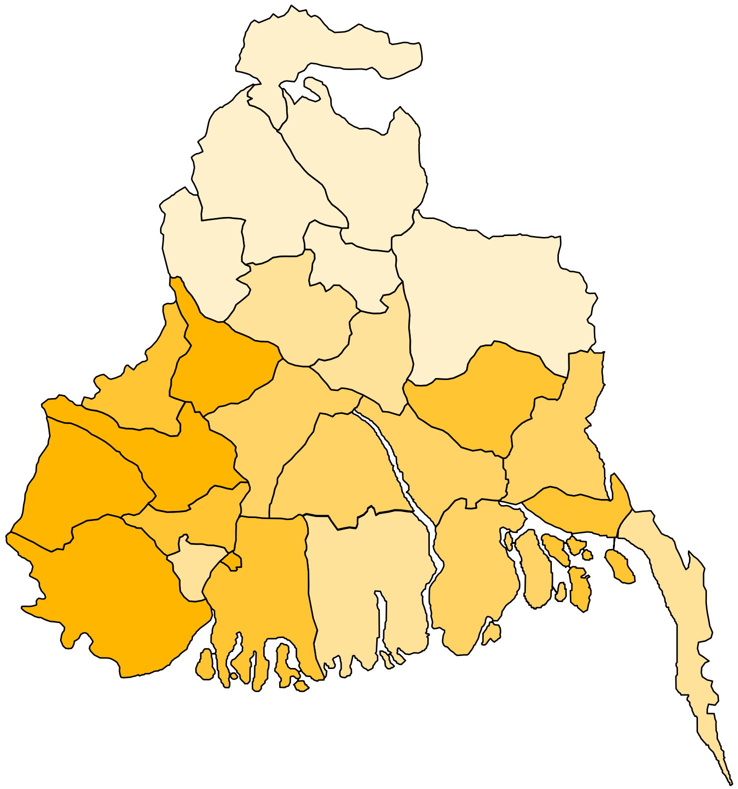
Monthly reported cholera mortality for Dacca, 1890–1940



Historic Bengal (now Bangladesh and the Indian state of West Bengal)



Parameter estimates vary smoothly across space



Example: estimated effect of environmental cholera (larger environmental reservoirs are shown as darker orange).

Iterated filtering bibliography

- Ionides, Breó & King (2006) *PNAS*. The basic theory and algorithms for iterated filtering.
- King, Ionides, Pascual & Bouma (2008) *Nature*. A case study of cholera transmission. The supporting online material gives practical advice for using iterated filtering.
- Bretó, He, Ionides & King (2008) *To appear in Annals of Applied Statistics*. A discussion of the plug-and-play concept and other related ideas. Development of continuous-time discrete-population models.

Conclusions

- Plug-and-play statistical methodology permits likelihood-based analysis of flexible classes of stochastic dynamic models.

- **It is increasingly possible to carry out data analysis via nonlinear mechanistic stochastic dynamic models**. This should help to build a link between the mathematical modeling community (for whom models are typically conceptual and qualitative) and quantitative applications (testing hypotheses about mechanisms, forecasting, evaluating the consequences of interventions).

- General-purpose statistical software for partially observed Markov processes is available in the **pomp** package for R (on CRAN).

Thank you!

These slides are available at

`www.stat.lsa.umich.edu/~ionides/pubs/samsi08.pdf`

References

Bretó, C., He, D., Ionides, E. L., and King, A. A. (2008). Time series analysis via mechanistic models. *Ann. Appl. Stat.*, to appear; pre-published at arxiv.org/abs/0802.0021.

Ionides, E. L., Bretó, C., and King, A. A. (2006). Inference for nonlinear dynamical systems. *Proc. Natl. Acad. Sci. USA*, 103:18438–18443.

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

King, A. A., Ionides, E. L., Pascual, M., and Bouma, M. J. (2008). Inapparent infections and cholera dynamics. *Nature*, 454:877–880.

Liu, J. and West, M. (2001). Combining parameter and state estimation in simulation-based filtering. In Doucet, A., de Freitas, N., and Gordon, N. J., editors, *Sequential Monte Carlo Methods in Practice*, pages 197–224. Springer, New York.

Sisson, S. A., Fan, Y., and Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA*, 104(6):1760–1765.