A Comparative Study of Techniques for Estimation and Inference of Nonlinear Stochastic Time Series

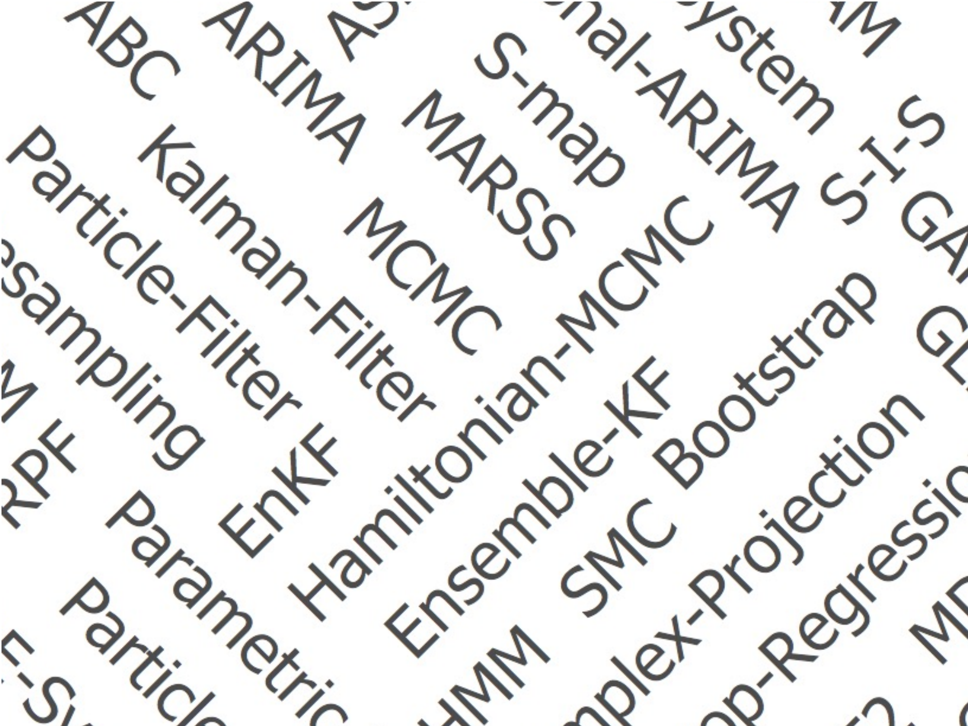
---

Dexter Barrows, B.Sc.

Masters Thesis Defence
McMaster University
April 2016

1. Framing
2. Stochastic SIR Model
3. Hamiltonian HMC
4. Iterated Filtering 2
5. Fitting Results
6. Forecasting Frameworks
7. S-maps & Seasonal Outbreaks
8. Spatiotemporal Epidemics
9. Parallelism & Future Directions

Framing


ABC
ARIMA
S-map
MARSS
Kalman-Filter
MCMC
Particle-Filter
Ensemble-MCMC
EnKF
Hamiltonian-MCMC
Ensemble-KF
Bootstrap
SMC
HMM

2 Stochastic SIR Model

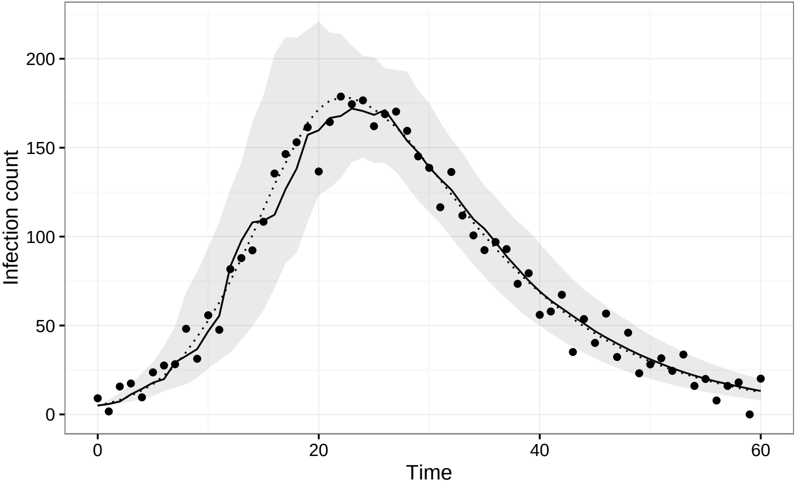
Stochastic SIR model

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

$$+$$

$$\beta_{t+1} = \exp\left[\beta_t + \eta\left(\bar{\beta} - \beta_t\right) + \mathcal{N}(0, \sigma_{proc})\right]$$

Model simulations



Solid – sample trajectory · dots – sample data · dotted – average · ribbon – overall

Hamiltonian MCMC 3

MCMC

Iteratively construct Markov chain to approximate posterior

1. Choose starting parameter set
2. Generate N samples by
   - 2.1 Propose new sample
   - 2.2 Compute acceptance ratio
   - 2.3 Accept/reject sample

HMC

Proposal via Hamiltonian dynamics

1. Choose starting parameter set
2. Generate N samples by
   - 2.1 Resample moments
   - 2.2 Simulate Hamiltonian dynamics using Leapfrog integration
   - 2.3 Compute acceptance ratio
   - 2.4 Accept/reject sample

Hamiltonian Dynamics

Energy

Kinetic $$K(r) = \frac{1}{2} r^T M^{-1} r$$

Potential $$U(\theta) = -\log(\mathcal{L}(\theta) p(\theta))$$

Hamiltonian $$H(\theta, r) = U(\theta) + K(r)$$

Dynamics simulation
$$\frac{d\theta}{dt} = M^{-1} r$$
$$\frac{dr}{dt} = -\nabla U(\theta)$$

Iterated Filtering 2

4

Basic particle filter

Iterative prediction-update cycle prunes particle cohort of poor parameter estimates

1. Initialize particles with parameter sets
2. For each data point
    2.1 Evolve particle states
    2.2 Weight via likelihood
    2.3 Resample proportional to weights

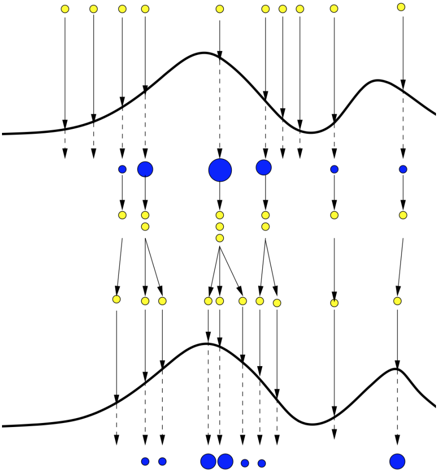


Image source [2]

Iterated Filtering 2 (IF2)

Evolution of MIF (IFI)

Multiple passes through data

Treat parameter estimates as stochastic processes

- Alleviates risk of particle collapse
- Process noise decreases with passes

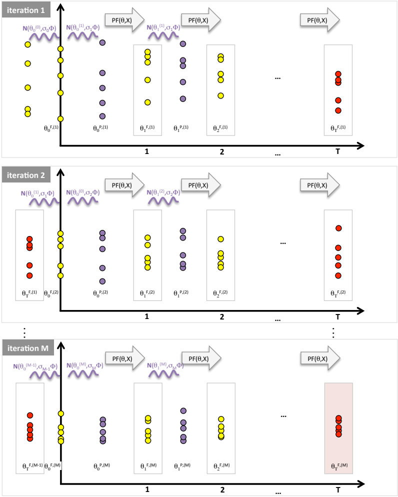
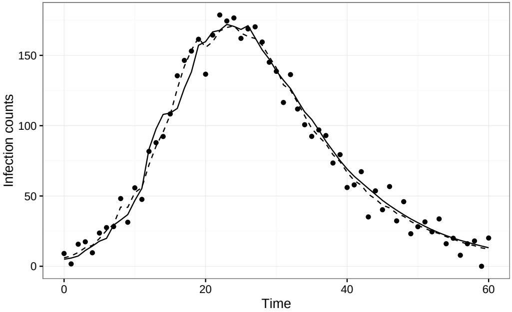


Image source [2]

5 Fitting Results

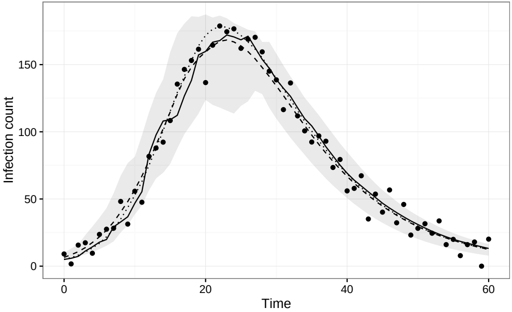
IF2

state estimates



HMC

state estimates



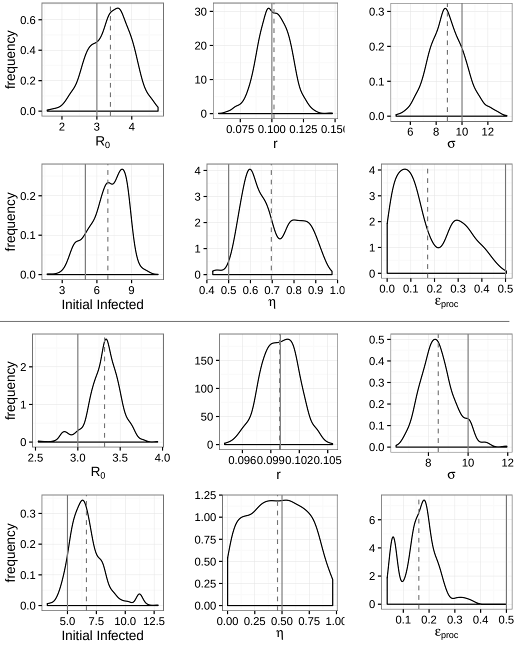
Dashed – estimates · solid – true · dots – data · ribbon – trajectories

IF2

final particle swarm samples

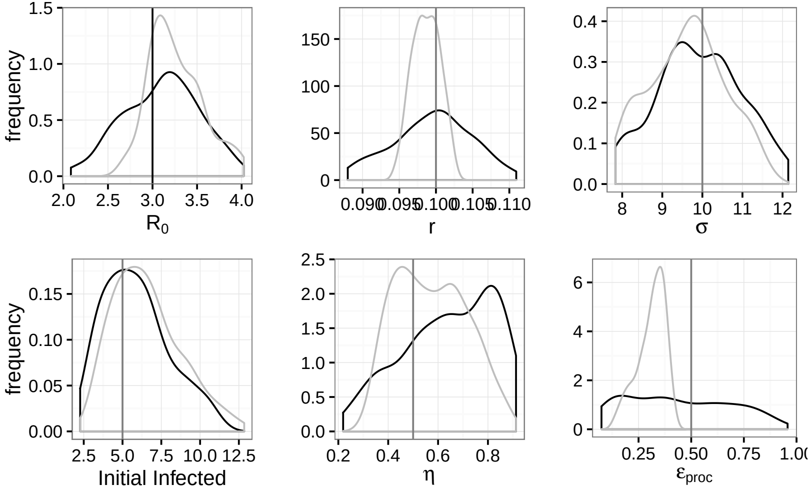
HMC

posterior distribution estimates



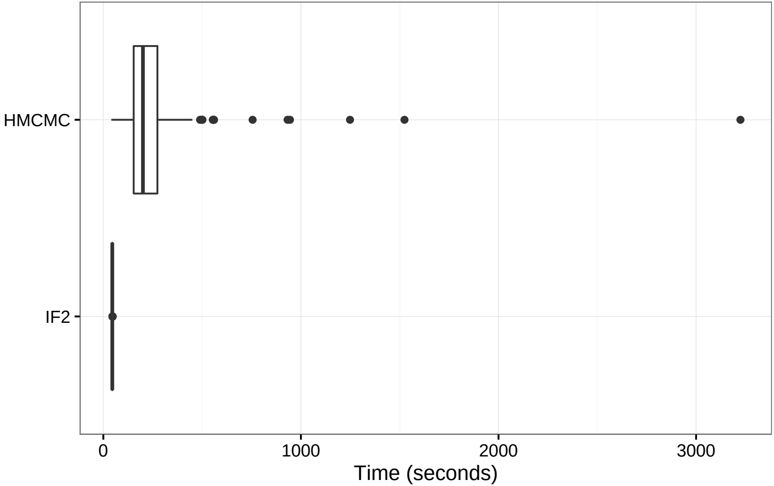
Dashed – medians · solid – true

Mean parameter estimate distributions



IF2 · HMC

Running times



IF2: 5.7x faster than HMC

6

Forecasting Frameworks

IF2

Parametric bootstrapping + forward simulation

- Provides additional samples from posterior distribution
- Forward simulation using states point estimate, posterior samples

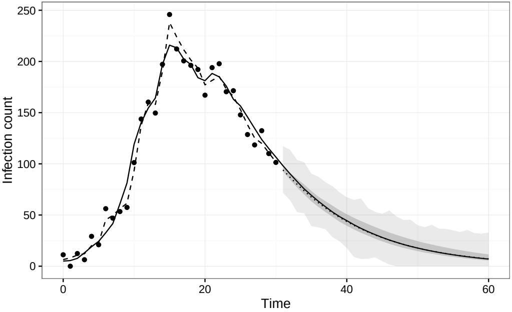
HMC

State reconstructions + forward simulation

- Reconstruct states using latent process noise samples
- Simulate forward

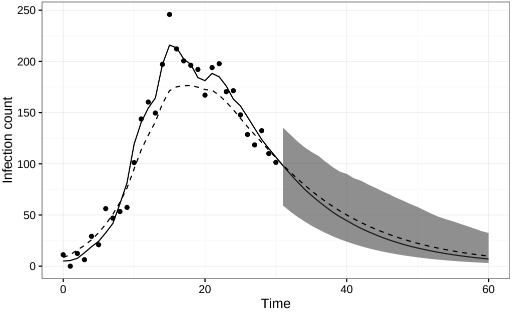
IF2

forecast



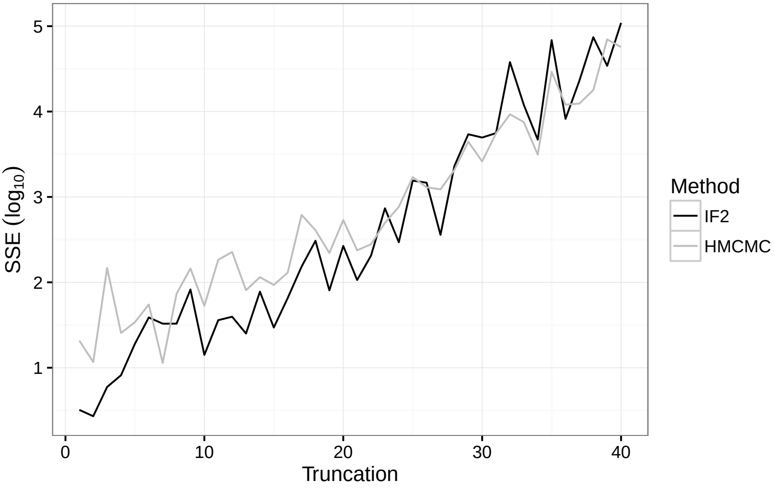
HMC

forecast



Dotted/dashed – mean · Dark ribbon - state · Light ribbon - observation

Forecast accuracy comparison



7

S-maps & Seasonal Outbreaks

Stochastic SIRS model

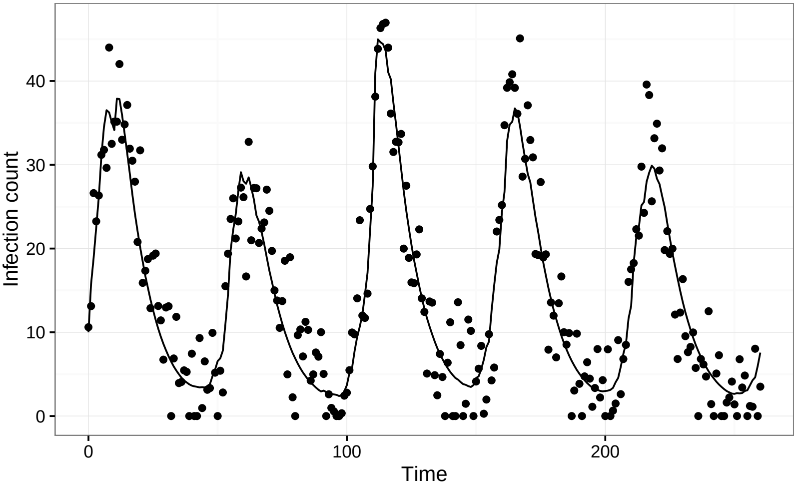
$$\frac{dS}{dt} = -\Gamma(t)\beta SI + \alpha R$$
$$\frac{dI}{dt} = \Gamma(t)\beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I - \alpha R$$

$$+$$

$$\Gamma(t) = \exp\left[2\left(\cos\left(\frac{2\pi}{365}t\right)\right)\right]$$

$$\beta_{t+1} = \exp\left[\beta_t + \eta\left(\bar{\beta} - \beta_t\right) + \mathcal{N}(0, \sigma_{proc})\right]$$

SIRS model simulation

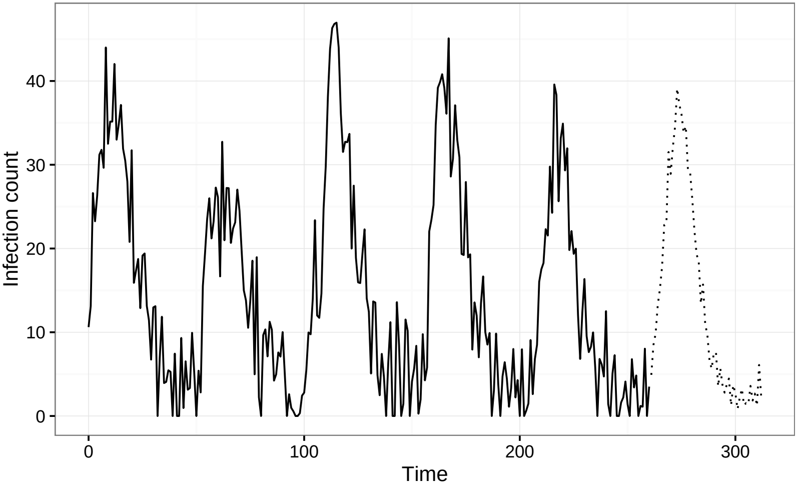


S-map

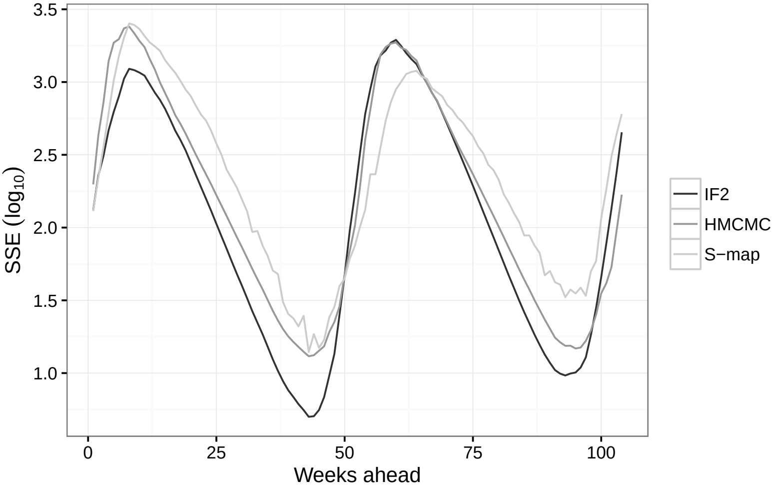
A little bit of history repeating

- Construct global weighted mapping of time-lagged vectors (the library) to future states
- Weightings are used to penalize poorly matching library vectors

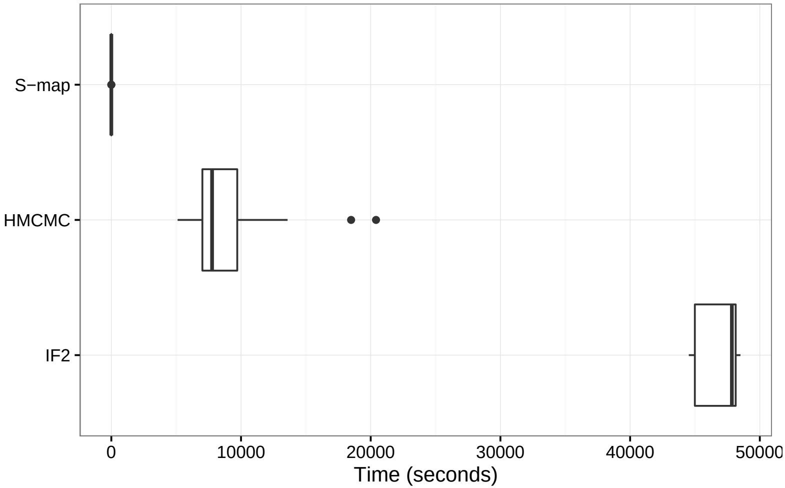
S-map forecast



SIRS model forecasting error



SIRS model forecasting runtimes



S-map: 316,000x faster than IF2, 61,800x faster than HMC

8

Spatiotemporal Epidemics

Stochastic Spatial SIR model

$$\frac{dS_i}{dt} = -\left(1 - \phi\frac{M}{M+1}\right)\beta_i S_i I_i - \frac{\phi}{M+1} S_i \sum_{j=1}^{M} \beta_{ij} I_j$$

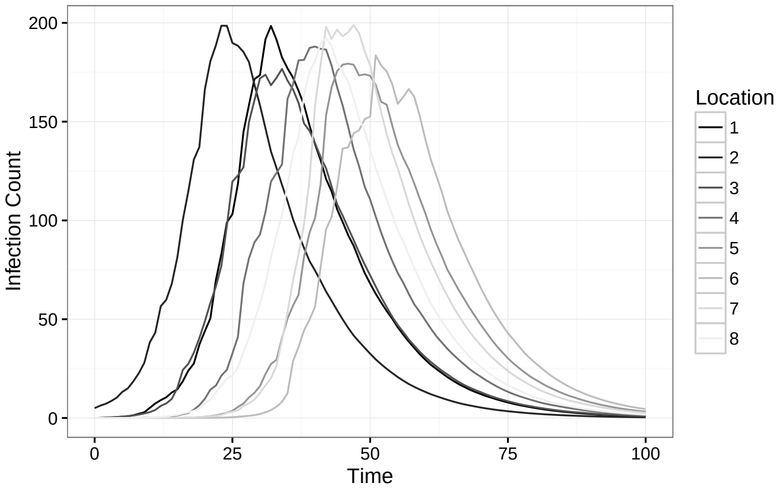
$$\frac{dI_i}{dt} = \left(1 - \phi\frac{M}{M+1}\right)\beta_i S_i I_i + \frac{\phi}{M+1} S_i \sum_{j=1}^{M} \beta_{ij} I_j - \gamma I_i$$

$$\frac{dR_i}{dt} = \gamma I_i$$

$$+$$

$$\beta_{i,t+1} = \exp\left[\beta_{i,t} + \eta\left(\bar{\beta} - \beta_{i,t}\right) + \mathcal{N}(0, \sigma_{\text{proc}})\right]$$

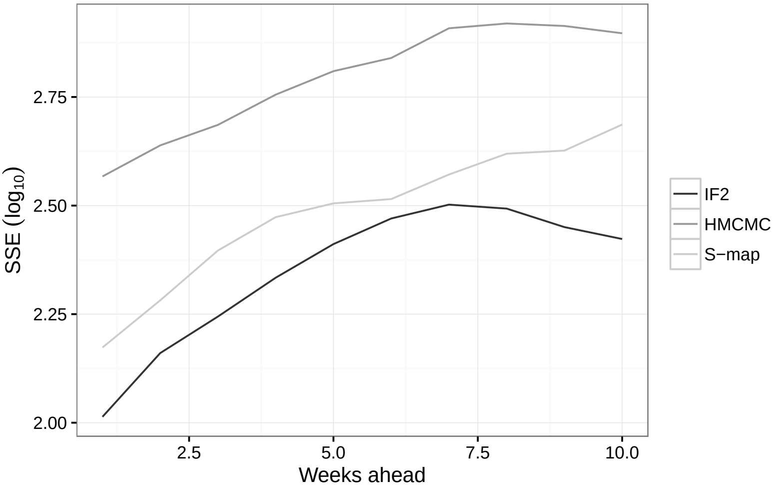
Stochastic spatial SIR model simulation (ring topology)



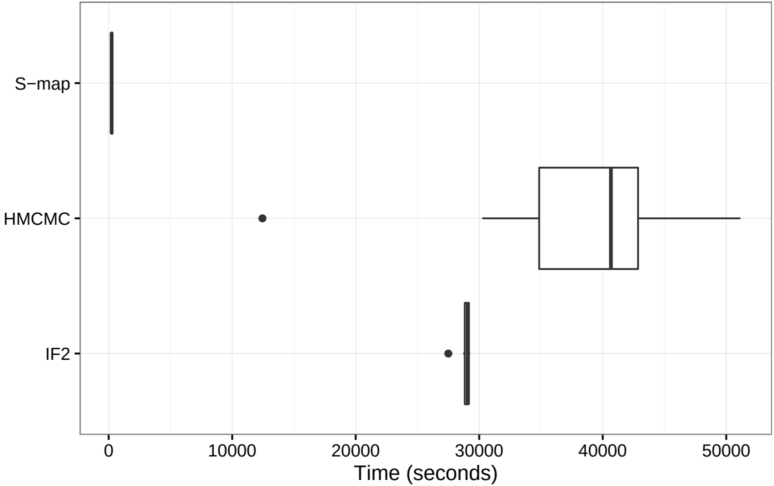
Dewdrop Regression

- “Stitching” together multiple short time series into single library
- Requires scaling

Spatial SIR model forecasting error



Spatial SIR model forecasting runtimes



S-map: 116x faster than IF2, 156x faster than HMC

9

Parallelism & Future Directions

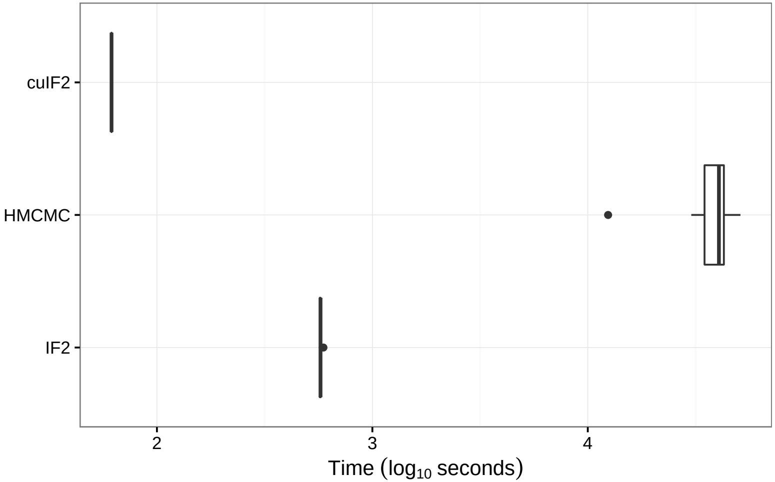
Conclusions

- IF2 produces superior forecasts in all scenarios
- S-mapping runs orders of magnitude faster than other methods

More than Moore

- Moore's Law is ceasing to hold
- Focus now on distributed computing
- MCMC-based methods are resistant to parallelization
    - Chain construction requires iterative dependence
- IF2 exhibits high parallel potential
    - Preliminary CUDA (GPU-accelerated) implementation - `cuIF2`

Spatial SIR model fitting times



`cuIF2`: 9.33x faster than IF2, 617x faster than HMC

Questions

[1] D. Barrows.
A Comparative Study of Techniques for Estimation and Inference of Nonlinear Stochastic Time Series.
Master's thesis, McMaster University, Hamilton, Canada, 2016.

[2] A. Doucet, N. de Freitas, and N. Gordon.
An Introduction to Sequential Monte Carlo Methods.
Sequential Monte Carlo Methods in Practice, pages 3–14, 2001.

THE BEST THESIS DEFENSE IS A GOOD THESIS OFFENSE.