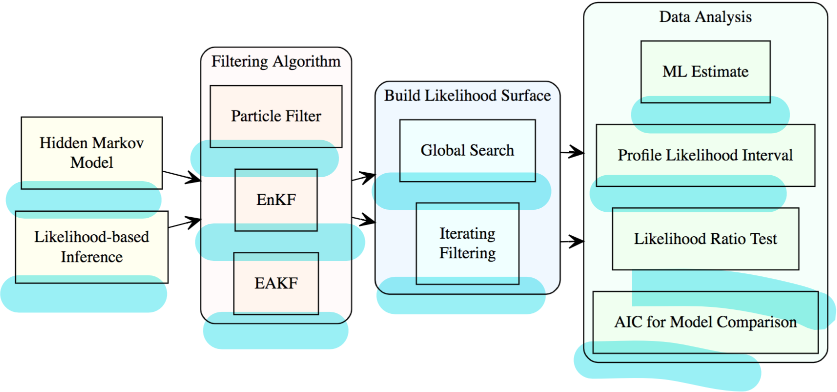
A Comparison of Kalman-type Filters and Particle Filter

Linke Li, Felicia Magpantay

Queen's University at Kingston

4 August 2020

Overview



Hidden Markov Model
Likelihood-based Inference
Filtering Algorithm
Particle Filter
EnKF
EAKF
Build Likelihood Surface
Global Search
Iterating Filtering
Data Analysis
ML Estimate
Profile Likelihood Interval
Likelihood Ratio Test
AIC for Model Comparison

Hidden Markov Model

Process Model

- $x_{k+1} = f_{x_{k+1}|x_k}(x_k)$

Measurement Model

- $z_k = h_{z_k|x_k}(x_k)$

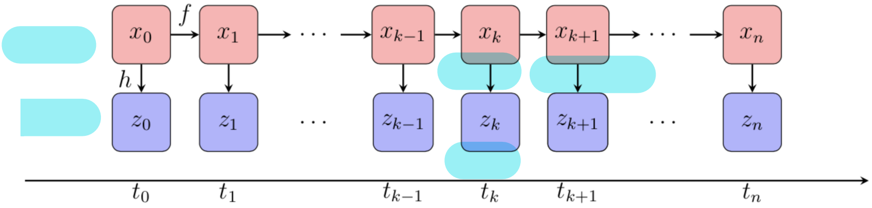
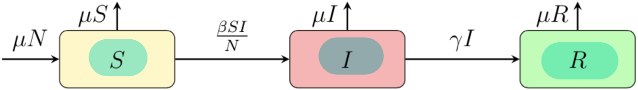


Figure 1: Diagram of an HMM.

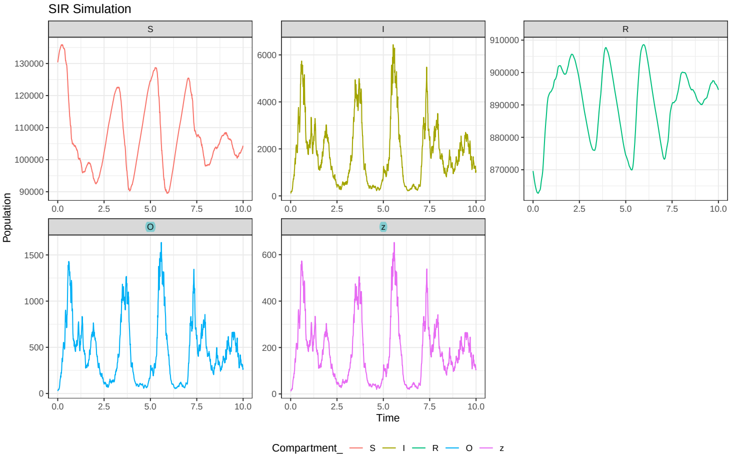
Data

- $z^*_{1:n}$, which are realizations of Random Variable $z_{1:n}$

SIR Model



μN
μS
S
$\frac{\beta SI}{N}$
μI
I
γI
μR
R



SIR Simulation
S
I
R
O
z
Population
Time
Compartment_
S
I
R
O
z

Likelihood Function of Hidden Markov Model

- Data $z^*_{1:n}$ can be seen as the random draw from distribution $f_{z_{1:n}}(z_{1:n};\theta)$, thus the likelihood function for HMM is:

$$\mathcal{L}(\theta) = f_{z_{1:n}}(z^*_{1:n};\theta)$$

  In general, $\mathcal{L}(\theta)$ does not have the closed form and can not be written explicitly.

- We can factorize the likelihood function such that:

$$\mathcal{L}(\theta) = f_{z_{1:n}}(z^*_{1:n};\theta) = \prod_{k=1}^{n} f_{z_k|z_{1:k-1}}(z^*_k|z^*_{1:k-1};\theta)$$
$$= \prod_{k=1}^{n} \int h_{z_k|x_k}(z^*_k|x_k) f_{x_k|z_{1:k-1}}(x_k|z^*_{1:k-1};\theta)\,dx_k$$

Likelihood Function of Hidden Markov Model

- Filtering algorithms work by iteratively updating the **prediction distribution** $f_{x_k|z_{1:k-1}}(x_k|z^*_{1:k-1};\theta)$ using the **update distribution** $f_{x_{k-1}|z_{1:k-1}}(x_{k-1}|z^*_{1:k-1};\theta)$ at time t_{k-1}, we can compute the **prediction distribution**, $f_{x_k|z_{1:k-1}}(x_k|z^*_{1:k-1};\theta)$, at time t_k by:

$$f_{x_k|z_{1:k-1}}(x_k|z^*_{1:k-1};\theta) = \int f_{x_k|x_{k-1}}(x_k|x_{k-1}) f_{x_{k-1}|z_{1:k-1}}(x_{k-1}|z^*_{1:k-1};\theta)dx_{k-1}$$

- The **update distribution** at time t_{k-1}, $f_{x_{k-1}|z_{1:k-1}}(x_{k-1}|z^*_{1:k-1};\theta)$, can be computed by fusing the prediction distribution $f_{x_{k-1}|z_{1:k-2}}(x_{k-1}|z^*_{1:k-2};\theta)$ at time t_{k-1} with the data z^*_{k-1} using the Bayes Rule:

$$\begin{aligned} f_{x_{k-1}|z_{1:k-1}}(x_{k-1}|z^*_{1:k-1};\theta) &= f_{x_{k-1}|z_{1:k-2},z_{k-1}}(x_{k-1}|z^*_{1:k-2}, z^*_{k-1};\theta) \\ &= \frac{f_{x_{k-1}|z_{1:k-2}}(x_{k-1}|z^*_{1:k-2};\theta) h_{z_{k-1}|x_{k-1}}(z^*_{k-1}|x_{k-1};\theta)}{\int f_{x_{k-1}|z_{1:k-2}}(x_{k-1}|z^*_{1:k-2};\theta) h_{z_{k-1}|x_{k-1}}(z^*_{k-1}|x_{k-1};\theta)dx_{k-1}} \end{aligned}$$

Filtering Algorithms

- **Particle Filter (PF)** works by iteratively approximate the prediction distribution and update distribution using particles.
- If the process model is linear and the measurement model has additive normal noise, prediction distribution and update distribution are always normal. Therefore, we only need to compute the mean and variance of the prediction and update distributions, which is exactly the **Kalman Filter (KF)**.
  In this case, the likelihood can be computed analytically.

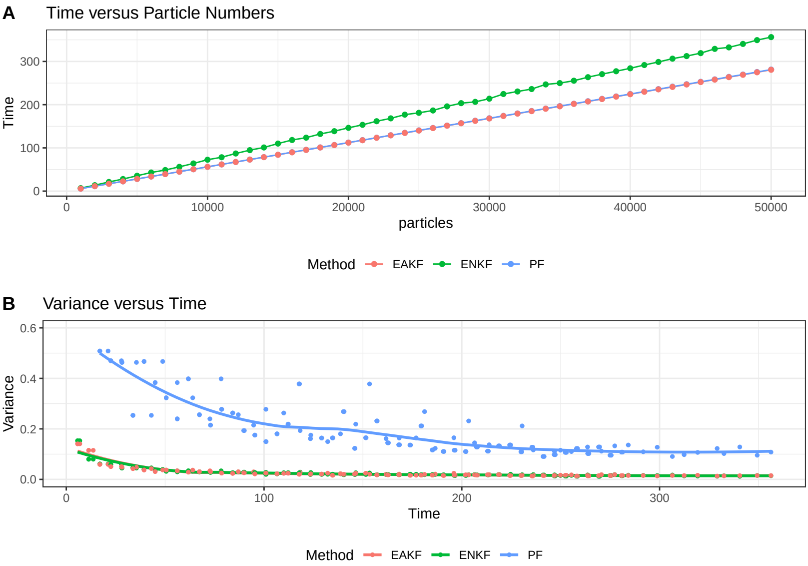
Filtering Algorithms

- If the process model is not strictly linear, we can use the idea of PF to modify KF. We start from a number of particles, using the formula of the mean of the KF to update each particles, and approximate the variance of the distributions by the variance of the particles.
- To solve the inconsistent between the particle numbers and data at t_k, we can add normal noise to the data, this is the idea of **Ensemble Kalman Filter (EnKF)** algorithm.
- Instead of introducing artificial noise, we can also deterministically update the particles, which leads to **Ensemble Adjustment Kalman Filter (EAKF)** algorithm.

Filtering Algorithms

- Kalman-type Filters (EnKF and EAKF) are only unbiased when the process model is linear and measurement model has normal additive noise, but they has smaller Monte Carlo error.
- PF can provide an unbiased estimate of the likelihood with greater Monte Carlo error, so it usually needs larger particle number J.

Computation Time and Monte Carlo Error



SIR Model

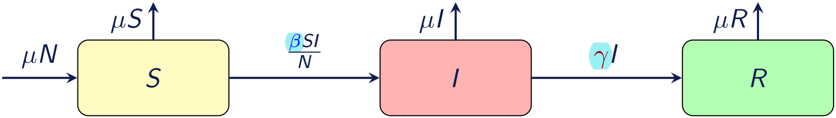
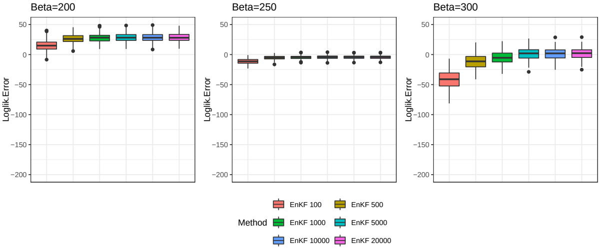
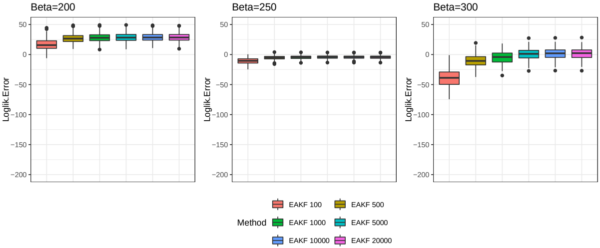


Figure 2: Diagram of an SIR model

Kalman-type Filter can be Biased





Likelihood of (β, γ)

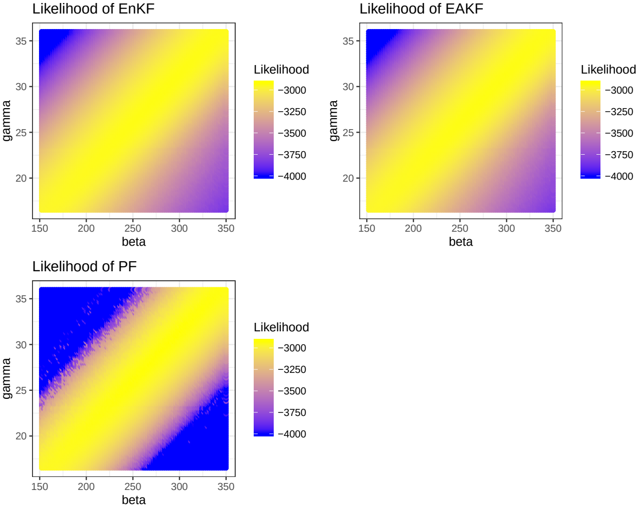


Figure 3: Likelihood for (β, γ)

Profile Likelihood and Wilks' Theorem

- The profile log likelihood function of β is defined to be

$$\ell^{\text{Profile}}(\beta) = \max_{\gamma} \ell(\beta, \gamma)$$

- The ML estimate of β is

$$\hat{\beta} = \arg\max_{\beta} \ell^{\text{Profile}}(\beta)$$

- Using Wilks' theorem, An approximate 95% confidence interval for β is given by

$$\{\beta : \ell^{\text{Profile}}(\hat{\beta}) - \ell^{\text{Profile}}(\beta) < 1.92\}$$

  1.92 comes from 0.95 quantile for a chi-squared distribution with one degree of freedom divide by 2.

Profile Likelihood of β

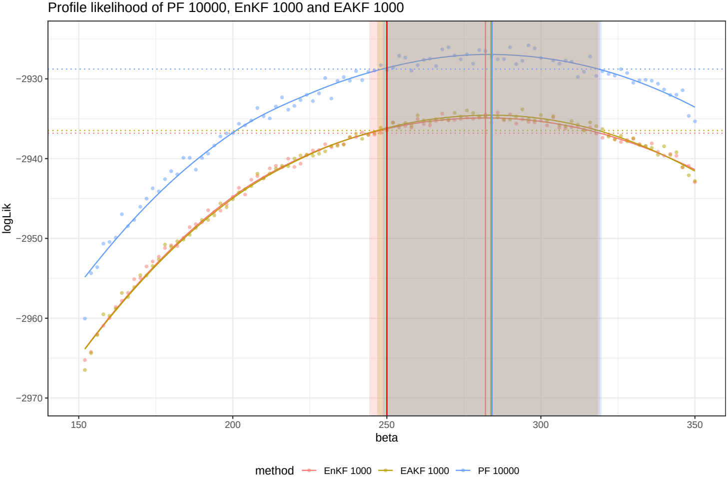


Figure 4: Profile Likelihood for infection rate (β)

Likelihood of β

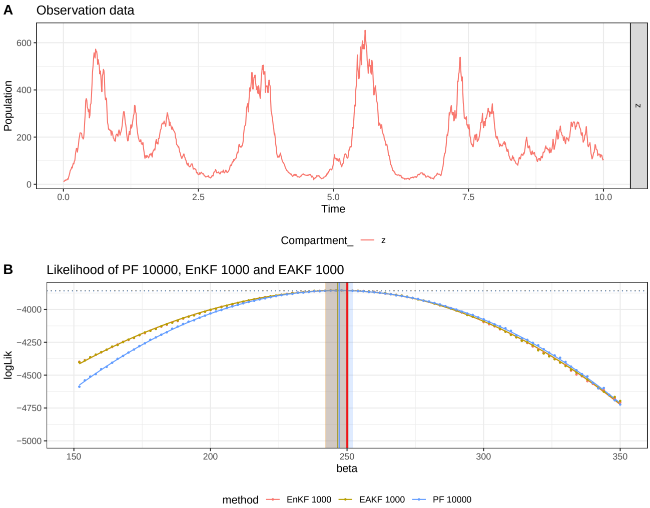
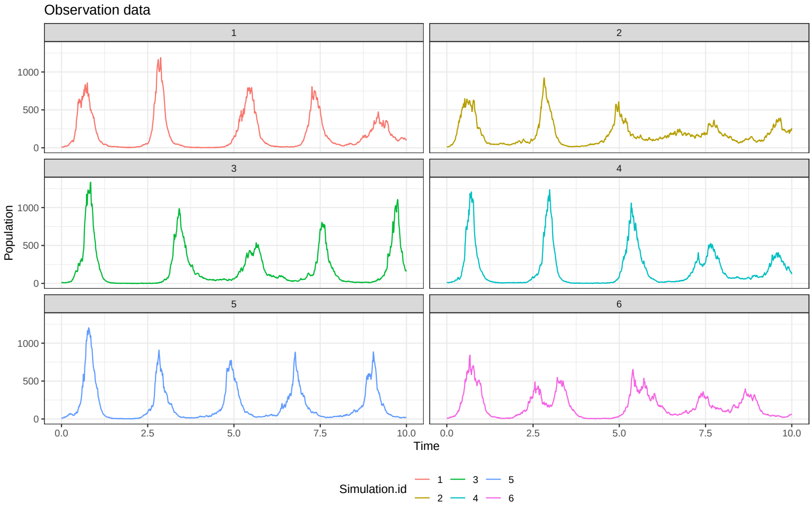
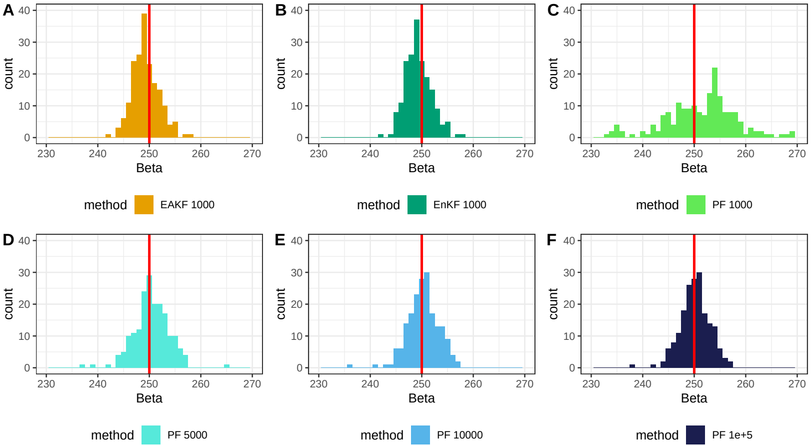


Figure 5: Likelihood for infection rate (β)

Simulation Samples of Observations $z^*_{1:n}$



Coverage Rate of β



	EnKF 10^3	EAKF 10^3	PF 10^3	PF 5×10^3	PF 10^4	PF 10^5
In 0.95 CI	0.941	0.941	0.602	0.871	0.892	0.914

Iterating Filtering

- Using filtering algorithms, we can only compute the stochastic estimate of the likelihood of HMM model. It is hard to find the ML estimate using such stochastic estimate by traditional optimizer (e.g. Newton-Raphson method)
- Iterated filtering method (Ionides et al., 2006, 2015) is the only currently available optimizer for the likelihood function of the HMM models.

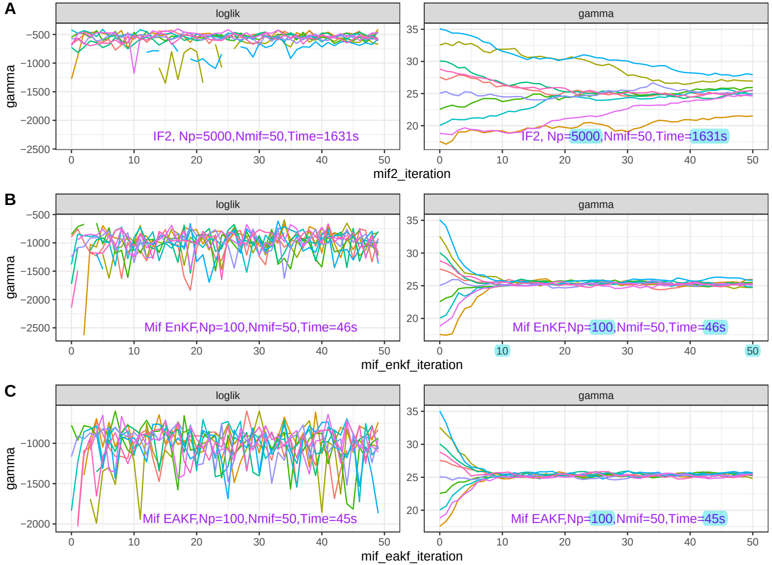
IF2 Algorithm

In the IF2 algorithm of Ionides et al. (2015):

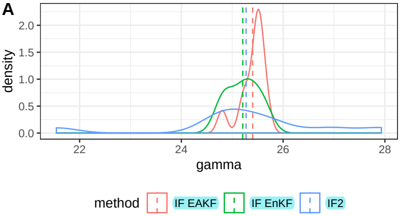
- Each iteration consists of a **PF**, carried out with the parameter vector, for each particle, doing a random walk.
- At the end of the time series, the collection of parameter vectors is recycled as starting parameters for the next iteration.
- The random-walk variance decreases at each iteration.

In theory, this procedure converges toward the region of parameter space maximizing the maximum likelihood. We implement the IF EnKF and IF EAKF algorithms by replacing the **PF** in the IF2 algorithm with **EnKF** and **EAKF** algorithms.

Convergence Plot of γ Given by Iterating Filtering



Statistics Given by Iterating Filtering



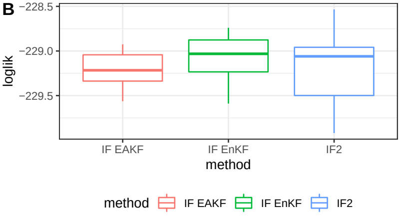


Table 1: Statistics of γ computed by IF2, IF EnKF and IF EAKF.

	γ	$\hat{\ell}$	$\hat{\ell}$ s.e.	$\tilde{\ell}$
Truth	26	−229.05	0.41	
IF2 MLE	21.54	−228.53	0.40	−229.12
IF EnKF MLE	24.87	−228.74	0.36	−229.06
IF EAKF MLE	25.49	−228.92	0.29	−229.18

Conclusion & Future Work

1. Conclusion
   - Maximum likelihood inference given by Kalman-type filters are comparable to PF in both the linear and non-linear model examples.
   - Iterated Kalman-type filtering algorithms outperform IF2 with fast convergence speed and promising results.
2. Future Work
   - A proof of the convergence of the iterated Kalman-type filtering is necessary.
   - The role that measurement noise plays in the iterated Kalman-type filtering is worth further investigation.