Applied Stochastic Analysis
Final project: Inference for nonlinear dynamical systems

Author:	Jaime Cascante-Vega
Email:	jc12343@nyu.edu
University ID:	N19711762

Introduction

Simulation of life processes across scientific domains is done using dynamical systems. The dynamical systems are complex simulations to describe a process of interest and are typically stochastic. Simulations span domain of science including chemical systems, economics, epidemiology, climate, population genetics and cosmology. These domains both scales of life, from the molecular to the population level as well as time scales. These mechanistic models are paired with inference methods to estimate the model parameters given a partial observation of the state space referred to as inverse problems.

In inverse problems the goal is to find the parameters given the probability density (likelihood) of a measured nature observation at a given time given the current model state. That frameworks allow then use of Bayesian inference in where broadly two methods exists: The first kind of methods seek to find the parameters that maximize this likelihood of observing the measured data. The second goal is rather to find the set of parameters that maximize probability density over parameter space given the data, or posterior parameter density. Both problems are challenging as typically just a few state variables are being observed. This partial-observation lead to inference procedure in where re-estimation of the state space happens when new data is available. Additionally, while the maximum likelihood approach typically suffer from mathematically intractable likelihoods and thus Monte Carlo methods are used to numerically approximate it.

Additionally, the stochastic process model satisfy the Markov property. The parameter inference procedure can then be treated by embedding the vector of parameters in the state space and treating it as a forward problem [1]. In this setting the inference of the state space will then update both the state variables representing the mechanistic model of the simulated phenomena and the re-estimation of parameters when observations are available.

In this work I implement such iterated filtering algorithm [1, 2] and compare the performance of two popular filtering methods: the Ensemble Adjustment Kalman Filter [3] and the Particle Filter or Sequential Monte Carlo (SMC) [4]. I used the Lorenz63 system as a system of study to investigate the methods proposed. In the Materials and Methods I introduce the notation and theory beyond parameter inference of a Partially Observed Markov Process (POMP), the Iterated Filtering (IF) and both SMC and EAKF. I present the results of parameter inferences as well as study the convergence of the mean and convergence of the posterior distribution. Lastly I discuss the differences between both filters.

Material and Methods

Overview

In this work I used the Lorenz-63 stochastic dynamical systems (see Lorenz-63) to investigate performance of one parameter inference framework the Iterated Filtering (IF). We introduce the notation as well as the mathematical framing of general simulation based inference methods (see Simulation-Based Inference). We then delineate the method used for parameter inference (see Iterated Filtering) [1, 2]. This method use Monte Carlo filtering algorithms to estimate the state and parameter space based on data. We investigate and compare the performance of two popular filtering methods the Ensemble Adjustment Kalman Filter (see EAKF) and the particle filter or sequential Monte Carlo (see SMC). The system studied is both stochastic by nature and is the famous Lorenz attractor in which small perturbations of trajectories results in divergent dynamics. To investigate the performance of the methods we compare convergence of the mean and in distribution and error of the inferences of each methods (see Comparing the filters).

Lorenz-63

The Stochastic Differential Equation (SDE) describing the evolution of the Lorenz-63 is presented below. u, v and w are independent Brownian motions. We use the Euler-Maruyama scheme to integrate the model as presented in Eq 8.18 of the Chapter 8 resource with $P(\xi = 1) = P(\xi = -1) = 1/2$. The integration step as well as simulation horizon are expected to impact the quality of the inference. We choose a fine integration step of $\Delta t = 0.001$ and simulate for a maximum time of 10.

$$\begin{aligned} dx_t &= \alpha(y_t - x_t)dt + dU_t \\ dy_t &= (x_t(\rho - z_t) - y_t)\,dt + dV_t \\ dz_t &= (x_t y_t - \beta z_t)\,dt + dW_t \end{aligned}$$

Simulation-Based Inference

Inverse problems in simulation-based inference assumes there exists an stochastic model that satisfies the Markov property. The simulator inputs a vector of parameters $\theta \in \mathbb{R}^p$ and sample latent variables at a specified time t over a time interval Δt according to a model $x_i^{t+\Delta t} \sim p_i\left(x_i^{t+\Delta t} | \theta, x_i^t\right)$. These latent variables can be represented in a vector named state space $\mathbf{z}_t \in \mathbb{R}^n$. I assume the simulated data $\mathbf{y}_i \in \mathbb{R}^k$ is a transformation applied to the state space $\mathbf{y}_t \sim g(\mathbf{y}_t | \mathbf{x}_t, \theta)$. The likelihood is then the product of the updates of the observations over the times where the data was generated, or data assimilation times. We assume data are generated at discrete times $\tau_1, \tau_2, \ldots, \tau_T$ and does not necessarily match with the integration steps of the process model. The likelihood of observing the vector of data

$\mathbf{y} = [\mathbf{y}_{\tau_1}, \mathbf{y}_{\tau_2}, \ldots, \mathbf{y}_{\tau_T}]$ is presented below.

$$g(\mathbf{z}|\mathbf{x}_t, \theta) = \prod_{t \in \{\tau_1, \ldots, \tau_T\}} \mathbf{g}(z_t | z_{t<t})$$

The goal of the filtering is then to infer the state and parameter space distributions given data $p(x_t|z_t)$ and $p(\theta|z_t)$ respectively. By applying the Bayes rule and using Markov one can write $p(x_t|z_t)$ and $p(\theta|z_t)$ as a product of the likelihood of observing the data, computed from the Monte Carlo samples $g(\mathbf{z}|\mathbf{x}_t, \theta)$ and $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ as shown below. The term $p(z_t|x_t)$ is then the likelihood that in case of the EAKF () is assumed to be Gaussian and in the PF () is approximated applying iteratively importance sampling. The denominator is omitted but should be $p(z_t)$, that is also a modeling choice to impose observational error. The same exact expression can be used for the parameter density.

$$p(x_t|z_t) \approx p(z_t|x_t) \cdot p(x_t|x_{t-1}, z_{t_1}) \tag{1}$$

Iterated Filtering

The IF modify the state space by embedding the parameters to be estimated in it $\mathbf{x}_t = [\mathbf{x}_t, \theta_t]$. The densities of model transitions are thus dependent on time-varying parameters $x_t \sim p(x_t|x_{t-1}, \theta_t)$, and so does the observational density $y_t \sim p(y_t|x_t, \theta_t)$. A completely specified model should also specify the conditional density of the initial conditions $p(x_0|\theta_0)$. The evolution of the state variables is determined by the stochastic process model. To simulate a evolution of the parameter space the IF assume a random walk within assimilation steps $\theta_{t+1} = \theta_t + \sigma_\theta \xi$, where ξ is assumed to be a standard normal random variable ($\xi \sim \mathcal{N}(0,1)$). Thus the conditional covariance $Var(\theta_t|\theta_{t-1}) = \sigma_\theta^2$. Within IF iteration the magnitude of the random walk is diminished $\sigma_\theta \to 0$ providing and estimate of θ. The pseudo-algorithm is presented in the appendix (see Algorithm 1).

Ensemble Adjustment Kalman Filter (EAKF)

The prior of the Markov process consists of m model simulations collectively referred to as an ensemble. Given a current measure of simulated variables z_τ at data assimilation time τ, we estimate the random error of observations assuming a normal observation with error variance proportional to the measured variable $z_\tau \sim \mathcal{N}(z_\tau, \sigma_y)$, as shown below in Equation 2. The model simulations of the full ensemble provide a distribution of simulated data, $\mathbf{y}_\tau = [y_\tau^1, y_\tau^2, y_\tau^3, \ldots, y_\tau^{300}]$, each $y_\tau^i \in \mathbb{R}^k$ where k is the size of the measured variable. The mean and co-variance of the priors can be calculated assuming normality, i.e. $y_\tau^i \sim \mathcal{N}(\mu_{prior}, \sigma_{prior}^2)$. Measured data z_τ is also assumed to be normally distributed, $z_t \sim \mathcal{N}(z_\tau, \sigma_{z,\tau}^2)$. Using these moments, observations and the EAKF, we compute the posterior variance and mean as shown below [5]. The pseudo-algorithm is presented in the appendix (see Algorithm 2).

$$\sigma_z^2 = \sigma_{z,\tau} = 1 + (p \cdot c_\tau)^2 \tag{2}$$

$$\begin{aligned}\sigma^2_{\text{post}} =& \sigma^2_{\text{prior}} \frac{\sigma^2_z}{\sigma^2_{\text{prior}} + \sigma^2_z} \\ \mu_{\text{post}} =& \sigma^2_{\text{post}} \left(\frac{\mu_{\text{prior}}}{\sigma^2_{\text{prior}}} + \frac{c_t}{\sigma^2_z} \right)\end{aligned} \tag{3}$$

The Kalman gain dy of each ensemble member $\mathbf{y}_\tau$ is given by

$$\mathbf{dy}_\tau = (\mu_{\text{post}} - \mathbf{y}_\tau) + \sqrt{\frac{\sigma^2_y}{\sigma^2_y + \sigma^2_{\text{prior}}}} (\mathbf{y}_\tau - \mu_{\text{prior}}) \tag{4}$$

The EAKF then uses the covariance between the parameters and the observations to compute the Kalman gain of the parameters $d\theta$ of each ensemble member for each set of parameters $\theta = [\theta_1, \theta_2, \ldots, \theta_{300}]$.

$$\mathbf{d}\theta = \frac{\text{cov}(\theta, \mathbf{y}_\tau)}{\sigma^2_{\text{prior}}} \times \mathbf{dy} \tag{5}$$

Here θ_i is the tuple of ensemble members $\theta_i = [\rho, \beta]$. The posterior observed state variable $\mathbf{y}_{\text{post}}$ and parameters θ_{post} are then given by

$$\begin{aligned}\mathbf{y}_\tau{}^{\text{post}} &= \mathbf{y}_\tau + \mathbf{dy} \\ \boldsymbol{\theta}_{\text{post}} &= \boldsymbol{\theta} + \boldsymbol{d\theta}\end{aligned} \tag{6}$$

Sequential Monte Carlo

The Particle Filter (PF) or Sequential Monte Carlo (SMC) use Monte Carlo samples of the stochastic process model to approximate and unknown likelihood function using importance sampling (IS). This methodology applied to a time series results in sequential use of IS once observations of a dynamical system are available. Each Monte Carlo sample is named as a particle. At the start of the data assimilation process all the particles are equally good or likely as no data has been observed. Given a importance density to measure the quality of the simulated in order to quantify the relative importance of each particle. The algorithm advances by assigning a relative weight to the particles according to the importance density and filtering the state and parameter space with probability equal to the density given by the distribution of weights. By including resampling (with replacement) a proportion of the best particles in each SMC iteration is maintained. The mathematical presentation as well as the pseudo-algorithm are presented below. Let $q(\mathbf{x})$ denote the importance density and assume that the prior of the state space $p(\mathbf{x})$ is approximated using Monte Carlo. The importance of each particle w_i is approximated by $w_i \approx \frac{p(\mathbf{x})}{q(\mathbf{x})}$. Once observations are available the density can be written in a conditional form $p(\mathbf{x}_t) = p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{z_t})$. The update of the weights can be updated using Bayes. Also note that in this form the state space $\mathbf{x}_t$ has embedded also the parameter space. The additional step in the particle filter is to resample the same number of particle with replacement with probability w^i_t. The pseudo-algorithm is presented in the appendix (see Algorithm 3).

$$w^i_t = w^i_{t-1} \frac{p(z_t | x^i_t) p(x^i_t | x^i_{t-1})}{q(x^i_t | x^i_{t-1}, z_t)} \tag{7}$$

Experiments comparing the filters

Given the stochastic nature and that the system is the famous Lorenz attractor perturbations near the divergence point of the attractor will results in very different trajectories. To overcome this issue we present inferences over a 100 of simulated trajectories. We studied the converge of the mean and of the distribution of the IF-EAKF and the IF-SMC. To study this we present the convergence across iterations of 100 of parameter inferences. The inferences were conducted on a random simulation of the SDE. I present the average of the statistics, mean and quantiles, of the 100 inferences and compare them between the EAKF and SMC filters. I compute the convergence in distribution using the Continuous Ranked Probability Score (CRPS) [6]. The expression for the CRPS is presented below. F is the cumulative distribution at a specific time and $\mathbf{1}$ is the indicator function such that $\mathbf{1}(x \geq y) = 1$ if $x \geq y$ or 0 otherwise. $x \in \mathbb{R}^{k \times m}$ is observation Monte Carlo samples and $y \in \mathbb{R}^k$ the measured observations. In results I present the average CRPS over the inferred time series, distribution of CRPS was computed using the parameters inferences over a 100 different trajectories.

$$CRPS(F,x) = \int_{-\infty}^{\infty} (F(x) - \mathbf{1}(x \geq y))\, dx$$

Results

I investigate the convergence of the mean across IF iterations for each filter, I present the mean across IF iterations in Figure 1. The y-axis is limited to match the flat uninformed priors used for the algorithm. Both filters mean distribution tend towards the truth value (red lines). The convergence of both algorithms is similar; both reach the truth by iteration 10. However, the mean of the EAKF appears to be more consistent that the one using the SMC filter. To investigate the convergence in distribution I present the average across a 100 inferences of the mean, and 3 quantiles (see Figure 2). The EAKF provides a estimate of the covariance while the SMC does not. This estimate is used within IF iterations and thus results in the IF-EAKF having an uncertainty estimate of all the parameters. This is seen as the truth value is being captured, falling between, the average credible intervals in the IF-EAKF. In the SMC however the uncertainty parameter estimate is absent resulting in a very narrow credible intervals (see Figure 2 right column). Another qualitative difference between the IF-EAKF and the IF-SMC is that the uncertainty estimate and the mean appear unchanged after convergence to a steady value in the IF-EAKF. The IF-SMC however exhibit variation and only until iteration 25 appears to convergence. These results suggest the IF-EAKF might have better convergence than the IF-SMC, and it's possible that this results from the covariance estimate in the EAKF.

To visually inspect the co-variance information as well as the marginal estimates of each parameter we present a corner plot and the marginal distributions in Figure 3 and Figure 4. As expected the EAKF provides a better visualization of the covariance between parameters than the PF. That is, the non-diagonal plots in Figure 3 compared with those in Figure 4

show correlations between the parameters. The IF-EAKF shows a correlation between the estimates of ρ and α and of ρ and β. It also suggest there is no covariance between β and ρ as the estimated distribution does not show changes of β as ρ change.

Lastly, we quantified the convergence of the posterior distribution by computing the CRPS and present it for the 3 parameters in Figure 5. In Figure 6 I present the average mean absolute error (MAE) across a 100 inferences and the 95% credible interval. Note that if the posterior distribution was a point both scoring rules would be the same. The first results is that the IF-EAKF have lower uncertainty as the 95% CI is narrower than those obtained by the IF-SMC. The IF-EAKF also obtained a better point estimate for ρ and β but a slightly worse for α. It shows the same pattern for the *CRPS* suggesting that in distribution most estimates are better for those two parameters for the IF-EAKF than the IF-PF.

Discussion

I presented a Bayesian inference framework for estimating parameters of non-linear dynamical systems, or solving complex inverse problems. The framework solve the inverse problem by treating the parameter space as part of the state space and using filters to update/filter this increased state space. The stochastic process model provide prior distribution of the state variables and the parameter space distribution follow a normal distribution with specified variance for each parameter. The variance that explore the parameter space decays across iteration of the algorithm and eventually reach a steady distribution. The Iterated Filtering framework was proposed using a Sequential Monte Carlo approach to maximize a user likelihood once a observation of the state space is available. We compared the performance of this SMC with an Ensemble Adjustment Kalman Filter that use Monte Carlo samples to estimate the mean and covariance of the dynamical system. By assuming normality of the state space and of the observations the posterior is also normally distributed. It's mean and co-variance can be updated deterministically. The EAKF thus provide covariance information of the filtered quantity. That covariance information appears to improve the consistency of the convergence by reaching a steady state in the posterior estimate in fewer iterations. It also provide a direct manner to quantify parameter estimate while the SMC does not. Other approaches such as likelihood profiling have been proposed to compute parameter uncertainty in SMC filters [7]. However, such approaches require more steps while the EAKF provides a direct way to estimate confidence in the parameter estimates. Lastly, I investigated the convergence of the mean point estimate across a 100 simulations as well as the convergence in distribution. I found that such the uncertainty quantified by the EAKF results in narrower errors of both the point and the distribution estimates for the IF-EAKF. It also showed smaller mean estimate for 2 of the 3 parameters.

Pseudo-algorithms

Algorithm 1 The Iterated Filtering (IF2)

Number of iterations N_{IF}

Number of ensembles m

Current model state space $\mathbf{x}_t \in \mathbb{R}^{N \times M}$

Current parameter space $\boldsymbol{\theta_n} \in \mathbb{R}^{P \times M}$

Process model $\mathscr{M}$

Cooling sequence σ_n^{IF}, $n \in [1, N_{IF}]$

Normal distribution of parameters to sample ensemble members $g \sim \mathscr{N}(\boldsymbol{\theta}_n, \sigma_n)$

$\sigma^2_{\theta_i}$ is the variance of the noise for parametrizing perturbations.

Vector of measured data y_t, $t \in [1, T]$

$\boldsymbol{\theta_1}$:= Initial guess of parameter space.

for $n = 2$ to N_{IF} **do**

  $\boldsymbol{x_0}$:= Initial guess of the state space.

  $\boldsymbol{\theta_n} = g(\boldsymbol{\theta}_{n-1}, \sigma_n)$

  for $\tau = 1$ to T **do**

    $\boldsymbol{x_\tau}, \boldsymbol{y_\tau} = \mathscr{M}(\boldsymbol{x}_{\tau-1}, \boldsymbol{\theta}_{n-1})$

    $\boldsymbol{\theta}_{\boldsymbol{\tau}}^{n-1} = \boldsymbol{\theta}_{\boldsymbol{\tau}}^{n-1} + \mathscr{N}(0, \mathrm{diag}([\sigma^2_{\theta_\tau^1}, \sigma^2_{\theta_\tau^2}, \ldots, \sigma^2_{\theta_\tau^p}])) \cdot \sigma_n$

    $\boldsymbol{x}_\tau, \boldsymbol{\theta}_\tau^{n-1}, \boldsymbol{y}_\tau = \mathbf{Filter}(\boldsymbol{\theta}_\tau^{n-1}, \mathbf{y}_\tau, \mathbf{z}_\tau, \sigma_z)$

  end for

end for

Output: Final parameter swarm $\boldsymbol{\theta_n}$.

Algorithm 2 Ensemble Adjustment Kalman Filter (EAKF)

Parameter space across m ensemble members $\theta \in \mathbb{R}^{p \times m}$

Simulated data at assimilation times τ in all ensembles $\mathbf{y}_\tau \in \mathbb{R}^{k \times m}$

Observed data at assimilation time τ, $\mathbf{z}_\tau \in \mathbb{R}^k$

Observed error variance at data assimilation time of each observation $\sigma_z = \sigma_{z,\tau} \in \mathbb{R}^k$.

Compute mean and variance of simulated data

$\mu_{\text{prior}} = \text{mean}(\mathbf{y}_\tau)$

$\sigma^2_{\text{prior}} = \text{var}(\mathbf{y}_\tau)$

Compute posterior mean and variance

$$\sigma^2_{\text{post}} = \sigma^2_{\text{prior}} \frac{\sigma^2_z}{\sigma^2_{\text{prior}} + \sigma^2_z}$$

$$\mu_{\text{post}} = \sigma^2_{\text{post}} \left(\frac{\mu_{\text{prior}}}{\sigma^2_{\text{prior}}} + \frac{c_\tau}{\sigma^2_z} \right)$$

Compute Kalman gain of observation and parameters space

$$\mathbf{dy} = (\mu_{\text{post}} - \mathbf{y}_\tau) + \sqrt{\frac{\sigma^2_z}{\sigma^2_z + \sigma^2_{\text{prior}}}} (\mathbf{y}_\tau - \mu_{\text{prior}})$$

$$\boldsymbol{d\theta} = \frac{cov(\boldsymbol{\theta}, \mathbf{y}_\tau)}{\sigma^2_{\text{prior}}} \times \mathbf{dy}$$

Compute posterior observations and parameter space

$\mathbf{y}^{\mathbf{post}}_\tau = \mathbf{y}_\tau + \mathbf{dy}$

$\boldsymbol{\theta}_{\mathbf{post}} = \boldsymbol{\theta} + \mathbf{dy}$

Output: Posterior number of data $\mathbf{y}^{\mathbf{post}}_\tau$ and posterior parameter state $\boldsymbol{\theta}_{\mathbf{post}}$.

Algorithm 3 Particle Filter

Number of particles m

Current model state space $\mathbf{x}_t \in \mathbb{R}^{N \times M}$

Importance density to quantify quality of the particles $g(x_k)$

Vector of measured data z_t, $t \in [1, T]$

$\boldsymbol{w}^i_0 = \frac{1}{m}$ Initial weights of each particle.

for $i = 1$ to m **do**

$x^i_t \sim q(x_t | x^i_{t-1}, \mathbf{z}_t)$ Draw each particle according to the proposal density $q(\cdot)$.

Update w^i_t according to Equation 7.

Normalize weights $w^i_t = \frac{w^i_t}{\sum_{i=1}^m w^i_t}$.

Resample particle x^i_t with probability w^i_t.

end for

Output: Final parameter swarm $\boldsymbol{\theta_n}$.

Figures

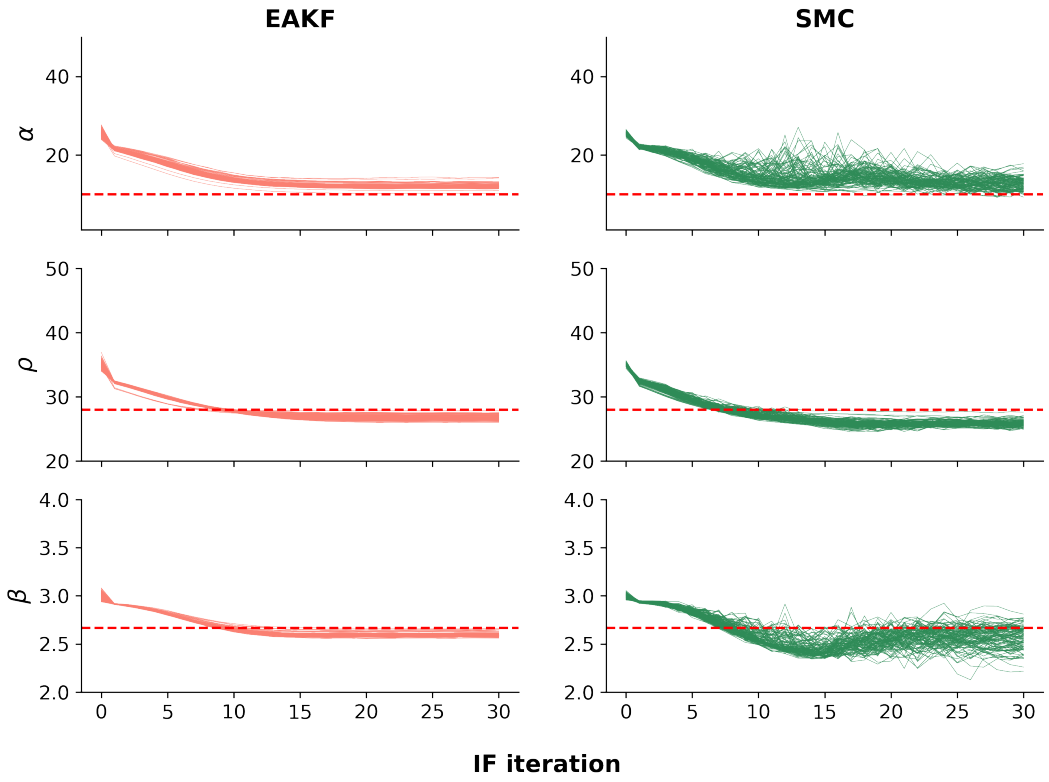

Figure 1: **Convergence of the mean.** Converge of the mean parameters estimates across Iterated Filtering iterations. Each line present the mean posterior parameter estimates of an inference on a simulated trajectory. The truth value is presented as a red dashed line.

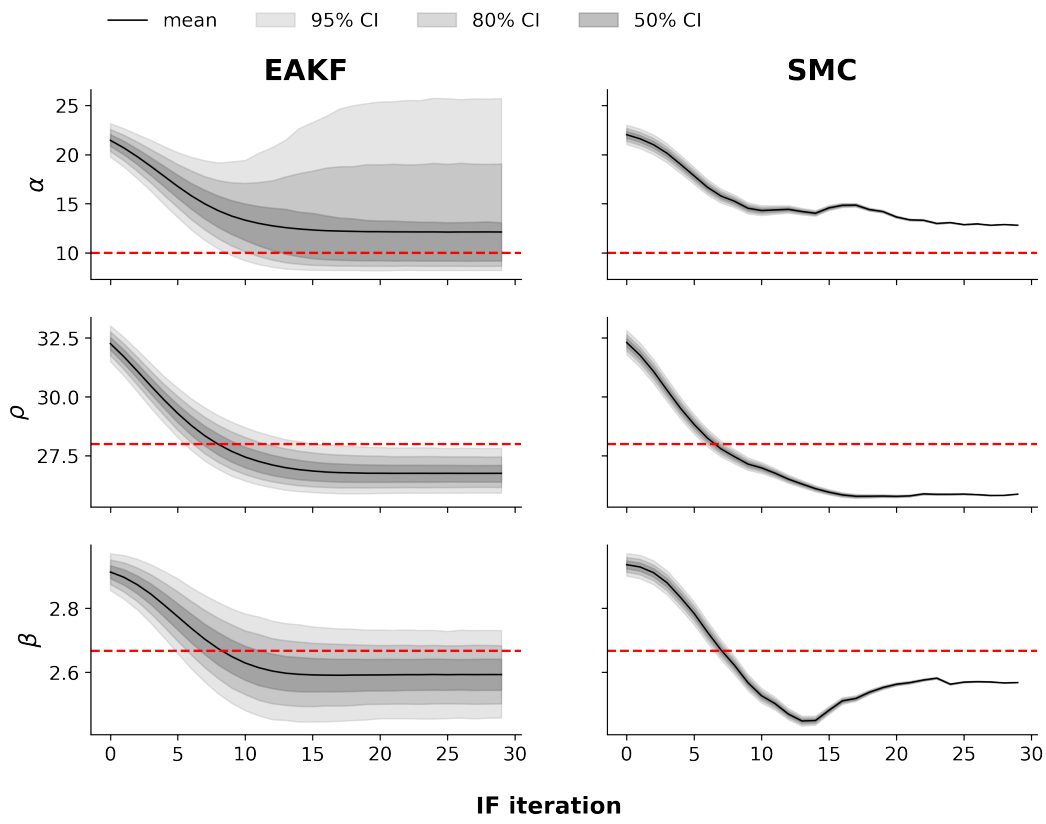

Figure 2: **Convergence of distribution.** Converge of the whole ensemble trajectories. Line present the mean across the 100 IF estimates. Ribbons show the average quantile across a 100 IF runs as presented in the legend. From light to dark ribbons I present the 95%, 80% and 50% CI respectively. The truth value is presented as a red dashed line. The left columns correspond to the IF using the EAKF (IF-EAKF) while the right one to the IF-SMC.

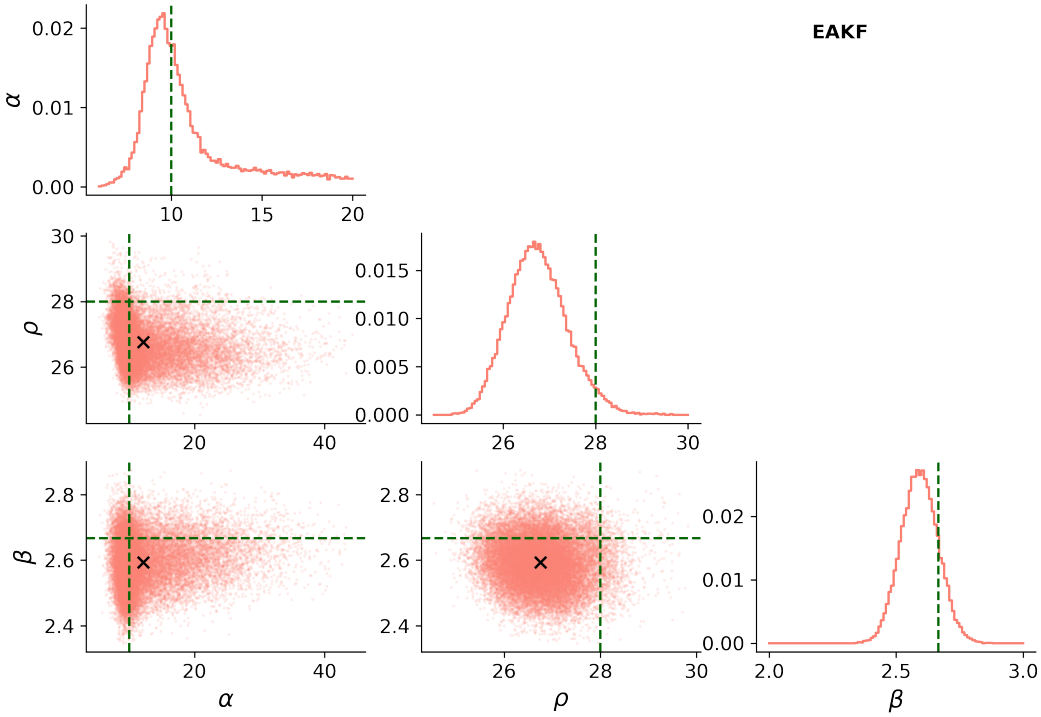

Figure 3: **Corner plot of IF-EAFK.** Marginal distribution for each parameter is presented in the diagonal. The parameter is indicated in the y-label of the first subplot in each row. Plots in the off-diagonal present the whole ensemble members or particles. The truth is presented in all sub-plots with a green dashed line.

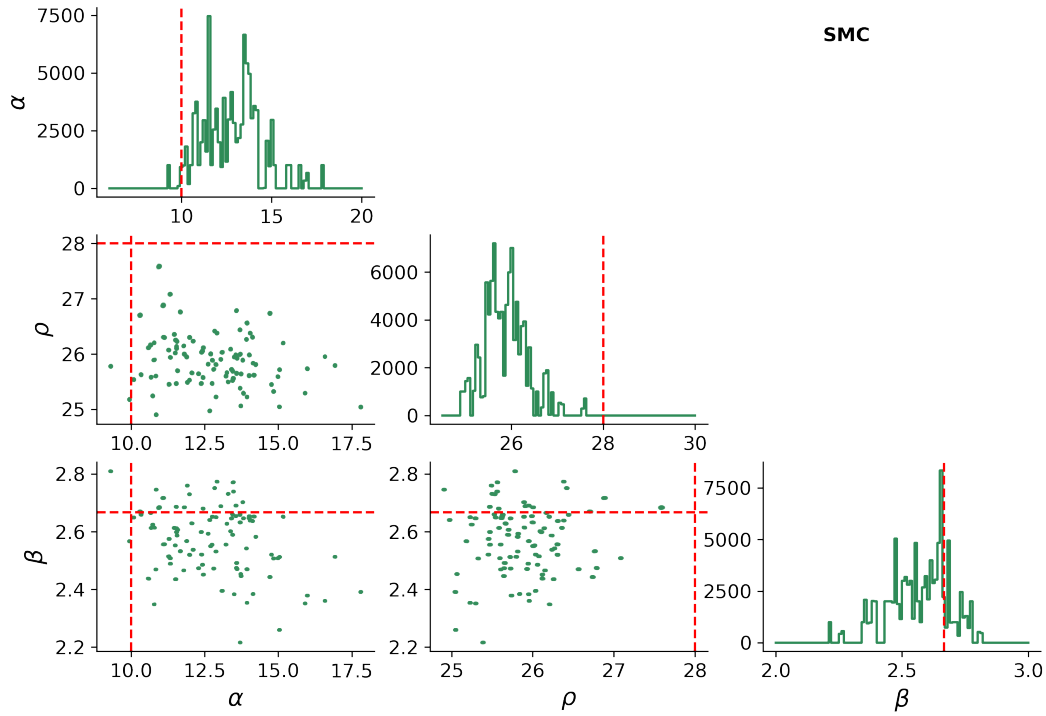

Figure 4: **Cornet plot of IF-PF.** Marginal distribution for each parameter is presented in the diagonal. The parameter is indicated in the *y*-label of the first subplot in each row. Plots in the off-diagonal present the whole ensemble members or particles. The truth is presented in all sub-plots with a red dashed line.

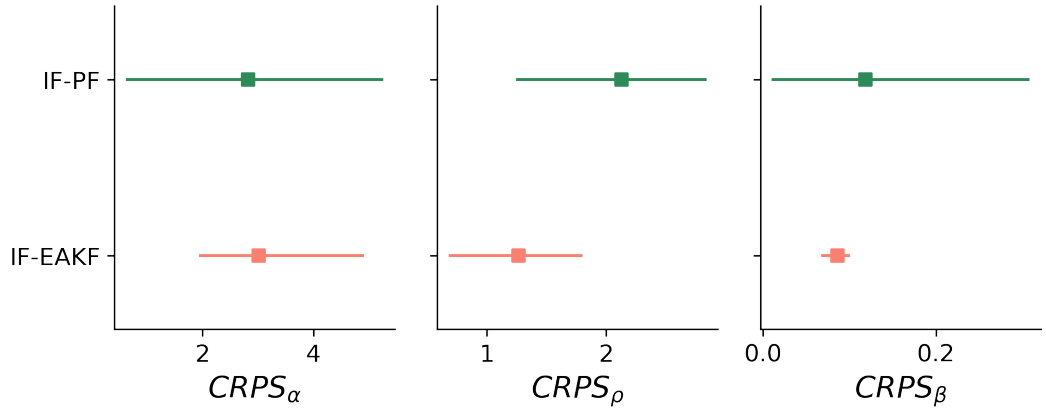

Figure 5: **CRPS of each filtering method and the 3 estimated parameters.** From left to right each subplot present a different parameter. The squares are the average CRPS across a 100 inferences while the lines present the 95%CI.

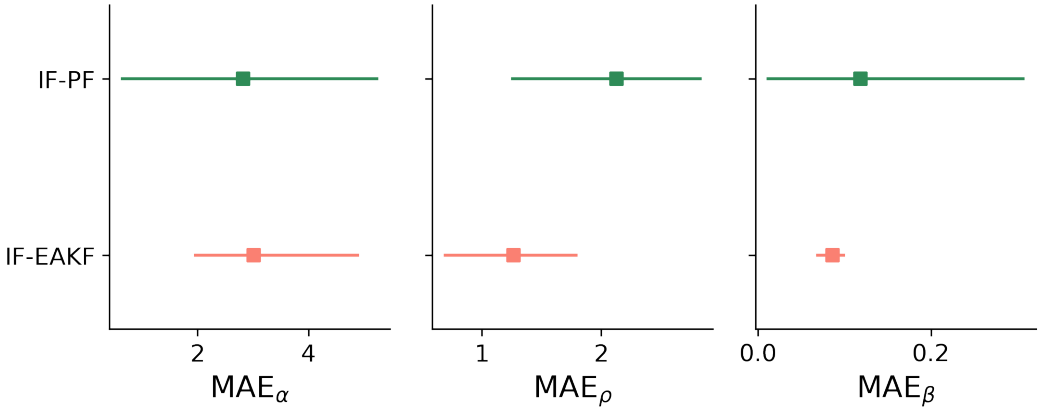

Figure 6: **MAE of each filtering method and the 3 estimated parameters.** From left to right each subplot present a different parameter. The squares are the average CRPS across a 100 inferences while the lines present the 95%CI.

References

[1] Ionides, E. L., Bretó, C. & King, A. A. Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **103**, 18438–18443 (2006). URL `https://pnas.org/doi/full/10.1073/pnas.0603181103`.

[2] Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S. & King, A. A. Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences* **112**, 719–724 (2015). URL `https://pnas.org/doi/full/10.1073/pnas.1410597112`.

[3] Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans* **99**, 10143–10162 (1994). URL `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/94JC00572`.

[4] Arulampalam, M., Maskell, S., Gordon, N. & Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing* **50**, 174–188 (2002). URL `http://ieeexplore.ieee.org/document/978374/`.

[5] Anderson, J. L. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review* **129**, 2884–2903 (2001). URL `http://journals.ametsoc.org/doi/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2`.

[6] Gneiting, T., Balabdaoui, F. & Raftery, A. E. Probabilistic Forecasts, Calibration and Sharpness. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **69**,

243–268 (2007). URL https://academic.oup.com/jrsssb/article/69/2/243/7109375.

[7] Ionides, E. L., Breto, C., Park, J., Smith, R. A. & King, A. A. Monte Carlo profile confidence intervals for dynamic systems. *Journal of The Royal Society Interface* **14**, 20170126 (2017). URL https://royalsocietypublishing.org/doi/10.1098/rsif.2017.0126.