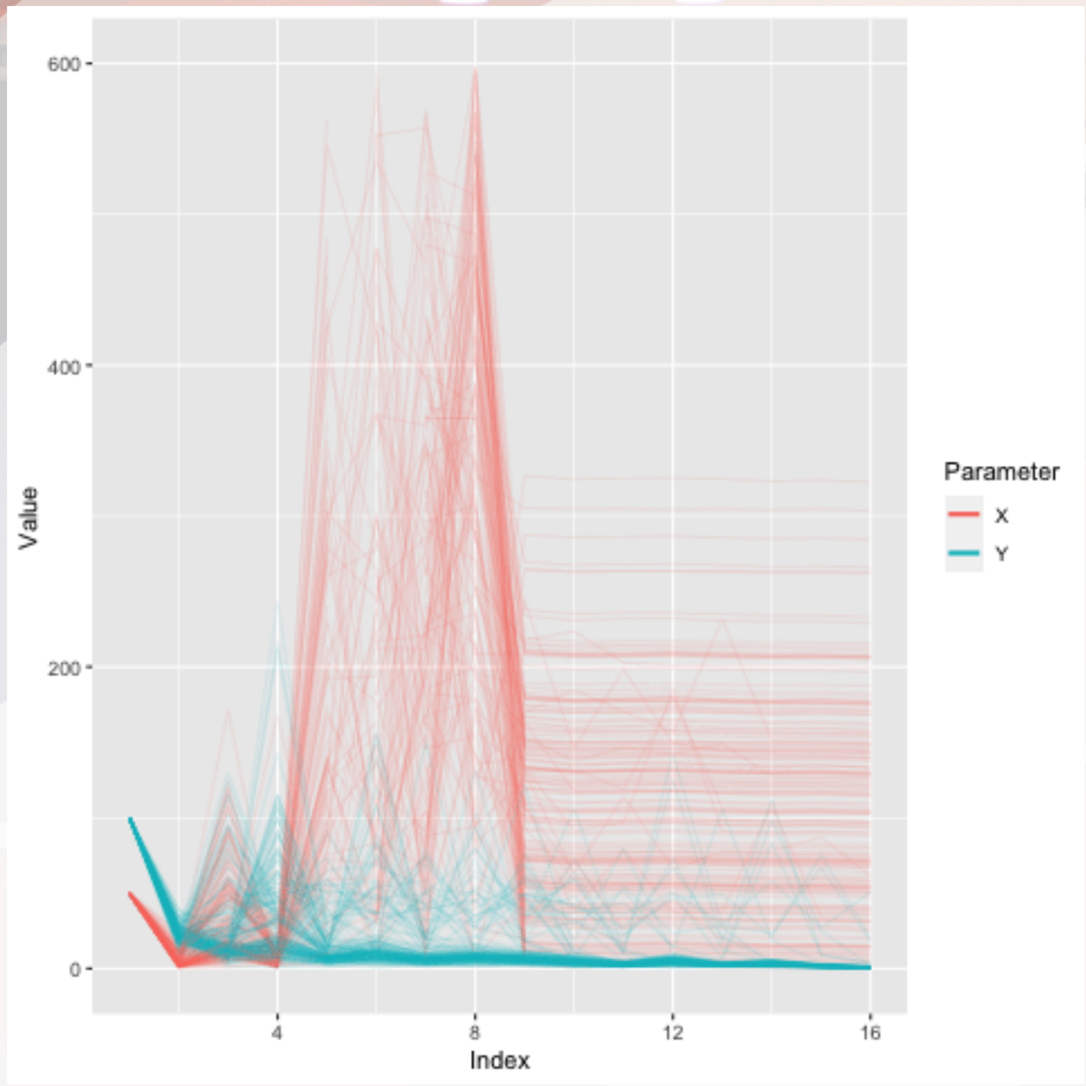
Ensemble Kalman inversion ABC



Richard Everitt

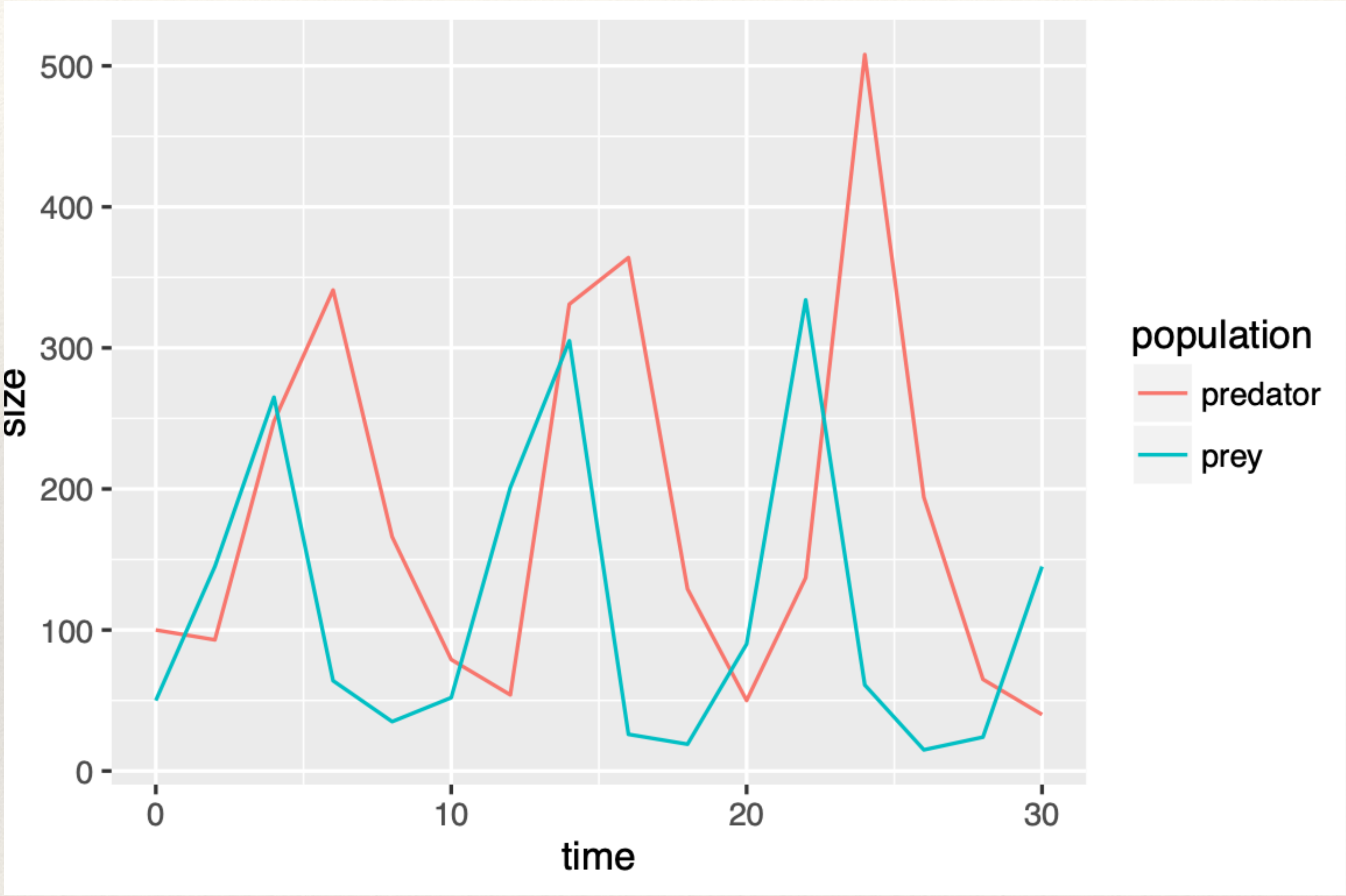
Warwick Statistics

THE UNIVERSITY OF WARWICK

UKRI Natural Environment Research Council

UKRI Engineering and Physical Sciences Research Council

Animal populations

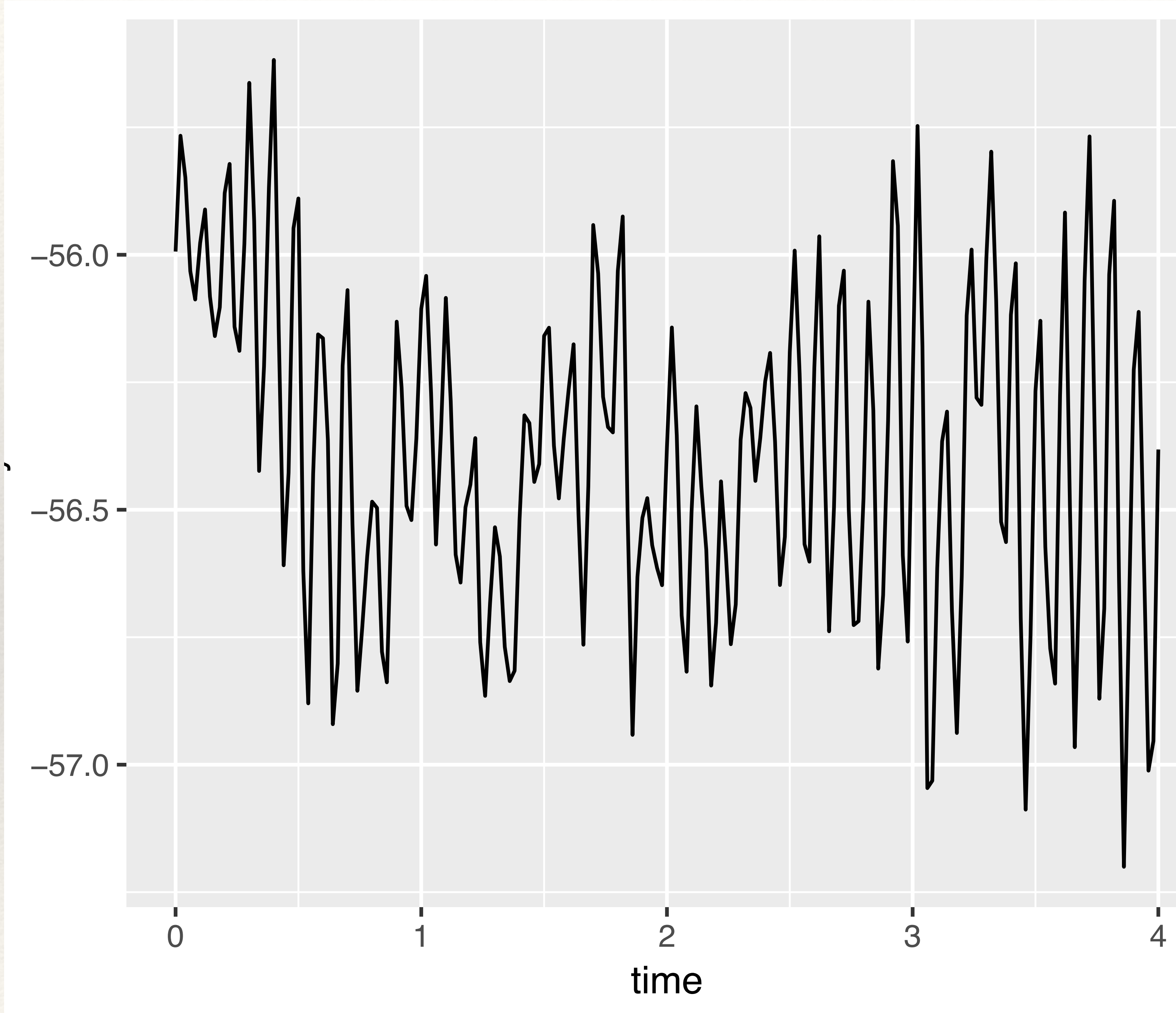


Natural Environment Research Council

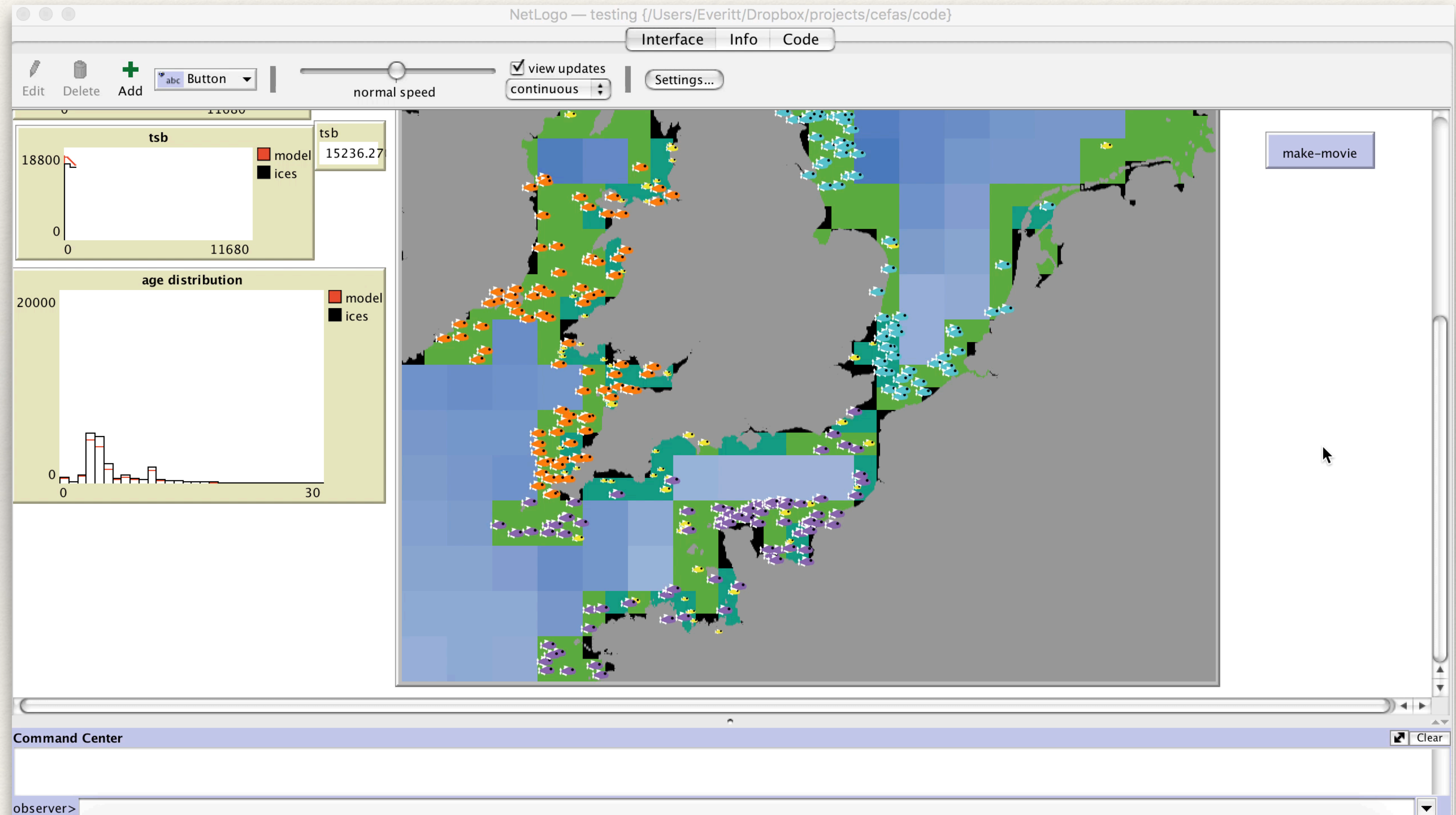
Warwick Statistics

Electroencephalogram (EEG) time-series

Thuglas at English Wikipedia
[Public domain]



Individual-based models (IBMs)



ABC-MCMC

- **Approximate Bayesian computation**
  - For the case where f cannot be evaluated pointwise at θ, use the ABC likelihood

$$l_{\text{ABC}}\left(y|\theta\right) = \int_y f\left(y|\theta\right) K_\epsilon\left(S\left(y_{\text{obs}}\right)|S\left(y\right)\right) dy$$

  - l_{ABC} is estimated, $\hat{l}_{\text{ABC}}\left(y|\theta\right)$, through using simulations $y \sim f\left(\cdot \mid \theta\right)$.

- **ABC-MCMC.** At iteration i:
  - $\theta^* \sim q\left(\cdot \mid \theta_i\right)$, and for $j = 1 : M$, $y^*_{i+1,j} \sim f\left(\cdot \mid \theta^*\right)$
  - With probability

$$1 \wedge \frac{p\left(\theta^*\right) \frac{1}{M} \sum_{j=1}^{M} K_\epsilon\left(S\left(y\right) \mid S\left(y^*_{i+1,j}\right)\right)}{p\left(\theta_i\right) \frac{1}{M} \sum_{j=1}^{M} K_\epsilon\left(S\left(y\right) \mid S\left(y_{i,j}\right)\right)}$$

  - let $\theta_{i+1} = \theta^*$ and for $j = 1 : M$, $y_{i+1,j} = y^*_{i+1,j}$.
  - otherwise let let $\theta_{i+1} = \theta_i$ and for $j = 1 : M$, $y_{i+1,j} = y_{i,j}$.

ABC-MCMC properties

- Efficiency of ABC-MCMC depends on controlling the variance of the likelihood estimator.
- Theory suggests to ensure $\mathbb{V}\left[\log \hat{l}_{\text{ABC}}\left(y \mid \theta\right)\right] \approx 3$.
- To achieve this, with $d_S = \dim S\left(y\right)$ and small ϵ, we need the number of simulations M to be exponential in d_S.

$$\boxed{l_{\text{ABC}}\left(y \mid \theta\right) = \int_y f\left(y \mid \theta\right) K_\epsilon\left(S\left(y_{\text{obs}}\right) \mid S\left(y\right)\right) dy}$$

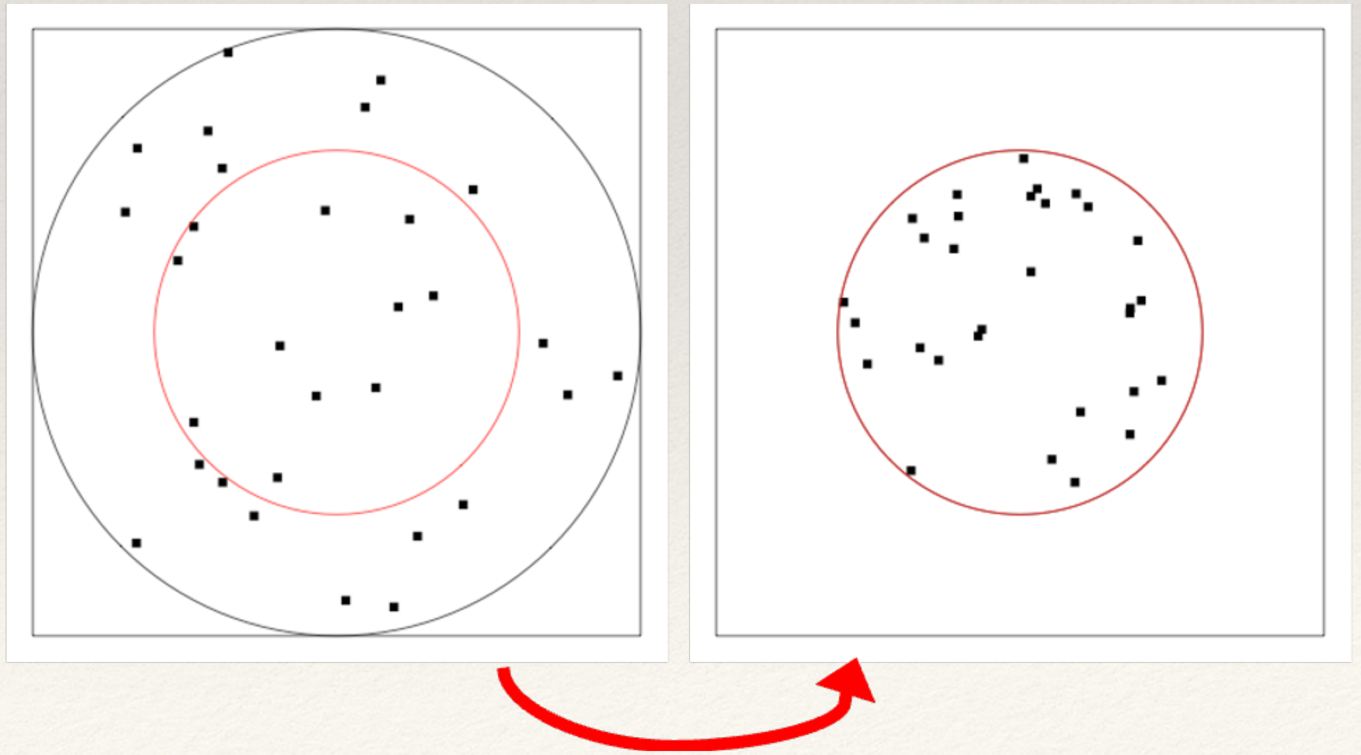
Prangle, Everitt and Kypraios, A rare event approach to high dimensional Approximate Bayesian computation, 2018.

Sherlock, Thiery, Roberts and Rosenthal, On the efficiency of pseudo-marginal random walk Metropolis algorithms, 2015.

Rare event ABC

- **Rare event ABC** (Prangle, Everitt and Kypraios, 2018)
  - reparameterise the simulator
    - instead of using $y \sim f(\cdot \mid \theta)$
    - call $u \sim \phi$ then take $y = G(\theta, u)$, where we can evaluate ϕ pointwise and G is a deterministic function.
  - Estimate $l_{\text{ABC}}(y|\theta)$ with an SMC sampler with a decreasing sequence of tolerances $\infty = \epsilon_0 > \ldots > \epsilon_T = \epsilon$.

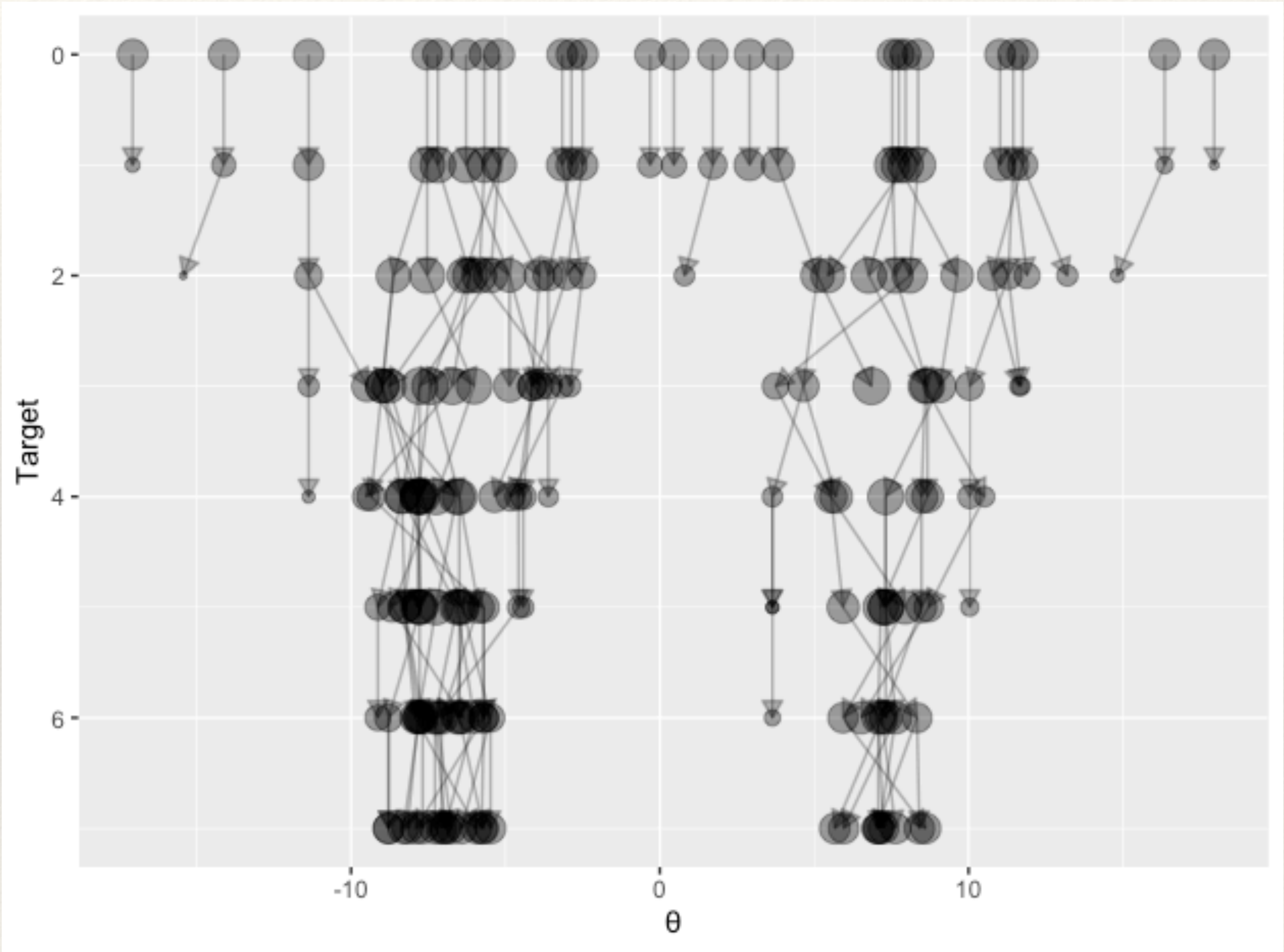
  - Cost of stabilising the variance of the likelihood estimator changes to $O\left(d_s^2\right)$.
  - Requires MCMC moves for exploring u-space.
  - Not possible for IBMs.



$$l_{\text{ABC}}(y|\theta) = \int_y f(y|\theta)\, K_\epsilon\left(S\left(y_{\text{obs}}\right) | S(y)\right) dy$$

SMC samplers recap

- **SMC sampler with annealing**:
  - p is a prior and l a likelihood
  - iterate from $t = 0 : T$ with target distribution at iteration t
    $$\pi_t(x) \propto p(x)\, l^{\alpha_t}(x)$$
    with $0 = \alpha_0 < \ldots < \alpha_T = 1$
  - a collection of importance points is iteratively reweighted, resampled and moved, as α_t changes.



Iterative ensemble Kalman inversion

- Assume that

$$l(x) = (2\pi)^{-d_x/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}\left(y_{\text{obs}} - H(x)\right)^T \Sigma^{-1} \left(y_{\text{obs}} - H(x)\right)\right)$$

- Use the sequence of targets

$$\pi_t(x) = p(x)\, l^{\alpha_t}(x).$$

- We have

$$l^{\alpha_t}(x) = (2\pi)^{-\alpha_t d/2} \det(\Sigma)^{-\alpha_t/2} \exp\left(-\frac{1}{2}\left(y_{\text{obs}} - H(x)\right)^T \left(\Sigma/\alpha_t\right)^{-1} \left(y_{\text{obs}} - H(x)\right)\right)$$

Iterative ensemble Kalman inversion

- Like an SMC sampler, iterative EnKI performs iterative updates to move from target t to $t+1$. We use

$$\pi_{t+1}(x) \propto \pi_t(x)\, l^{\alpha_{t+1}-\alpha_t}(x)$$

- At each iteration an update is performed using the likelihood

$$l^{\alpha_{t+1}-\alpha_t}(x) \propto \exp\left(-\frac{1}{2}\left(y_{\text{obs}} - H(x)\right)^T \left(\Sigma/\left(\alpha_{t+1}-\alpha_t\right)\right)^{-1} \left(y_{\text{obs}} - H(x)\right)\right)$$

- The “ensemble Kalman” idea is to:
  - suppose that $\pi_t(x)$ is Gaussian;
  - suppose that $l(x)$ (and hence that $l^{\alpha_{t+1}-\alpha_t}(x)$) is linear-Gaussian;
  - use conjugate Bayesian analysis to work out what the mean and covariance of $\pi_{t+1}(x)$ would be when the Gaussian $\pi_t(x)$ is updated with the linear-Gaussian $l^{\alpha_{t+1}-\alpha_t}(x)$;
  - represent $\pi_t(x)$ and $\pi_{t+1}(x)$ with Monte Carlo points, and construct updates such that their sample means and covariances match the analytic versions.

Iterative ensemble Kalman inversion

- Begin by simulating points from the prior. For $j = 1 : M$

$$x^{(j)} \sim p.$$

- Perform an ensemble Kalman update of these points at each iteration.
- At iteration $t+1$, for $j = 1 : M$

$$x_{t+1}^{(j)} = x_t^{(j)} + \hat{K}_{t+1}\left(y_{\text{obs}} - \tilde{y}_{t+1}^{(j)}\right)$$

with

$$\tilde{y}_{t+1}^{(j)} \sim \mathcal{N}\left(h_t^{(j)}, \left(\alpha_{t+1} - \alpha_t\right)^{-1}\Sigma\right)$$

where $h_t^{(j)} = H\left(x_t^{(j)}\right)$ and

$$\hat{K}_{t+1} = \hat{C}_t^{x_t h_t}\left(\hat{C}_t^{h_t h_t} + \left(\alpha_{t+1} - \alpha_t\right)^{-1}\Sigma\right)^{-1}$$

$$\hat{C}_t^{y_t h_t} = \frac{1}{M-1}\sum_{j=1}^{M}\left(x_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} x_t^{(j)}\right)\left(h_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} h_t^{(j)}\right)^T \qquad \hat{C}_t^{h_t h_t}\frac{1}{M-1}\sum_{j=1}^{M}\left(h_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} h_t^{(j)}\right)\left(h_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} h_t^{(j)}\right)^T$$

Normalising constants from IEnKI

- To obtain an estimate of $Z = \int_x p(x)\, l(x)\, dx$ we could use

$$\hat{Z} = \prod_{t=0}^{T-1} \widehat{\frac{Z_{t+1}}{Z_t}} \qquad \text{where} \qquad \frac{Z_{t+1}}{Z_t} = \int_x p(x)\, l^{(\alpha_{t+1}-\alpha_t)}(x)\, dx$$

- **Observation:** the sequence of distributions used in IEnKI is not

$$p(x)\left(\mathcal{N}\left(y_{\text{obs}} \mid x, \Sigma\right)\right)^{\alpha_t}$$

  it is

$$p(x)\left(\mathcal{N}\left(y_{\text{obs}} \mid x, \alpha_t \Sigma\right)\right).$$

- **Approach:**
  - use an ensemble-Kalman approximation of the ratio of normalising constants for the latter sequence of targets;
  - apply a correction to obtain an estimate of Z_{t+1}/Z_t.

Use in ABC

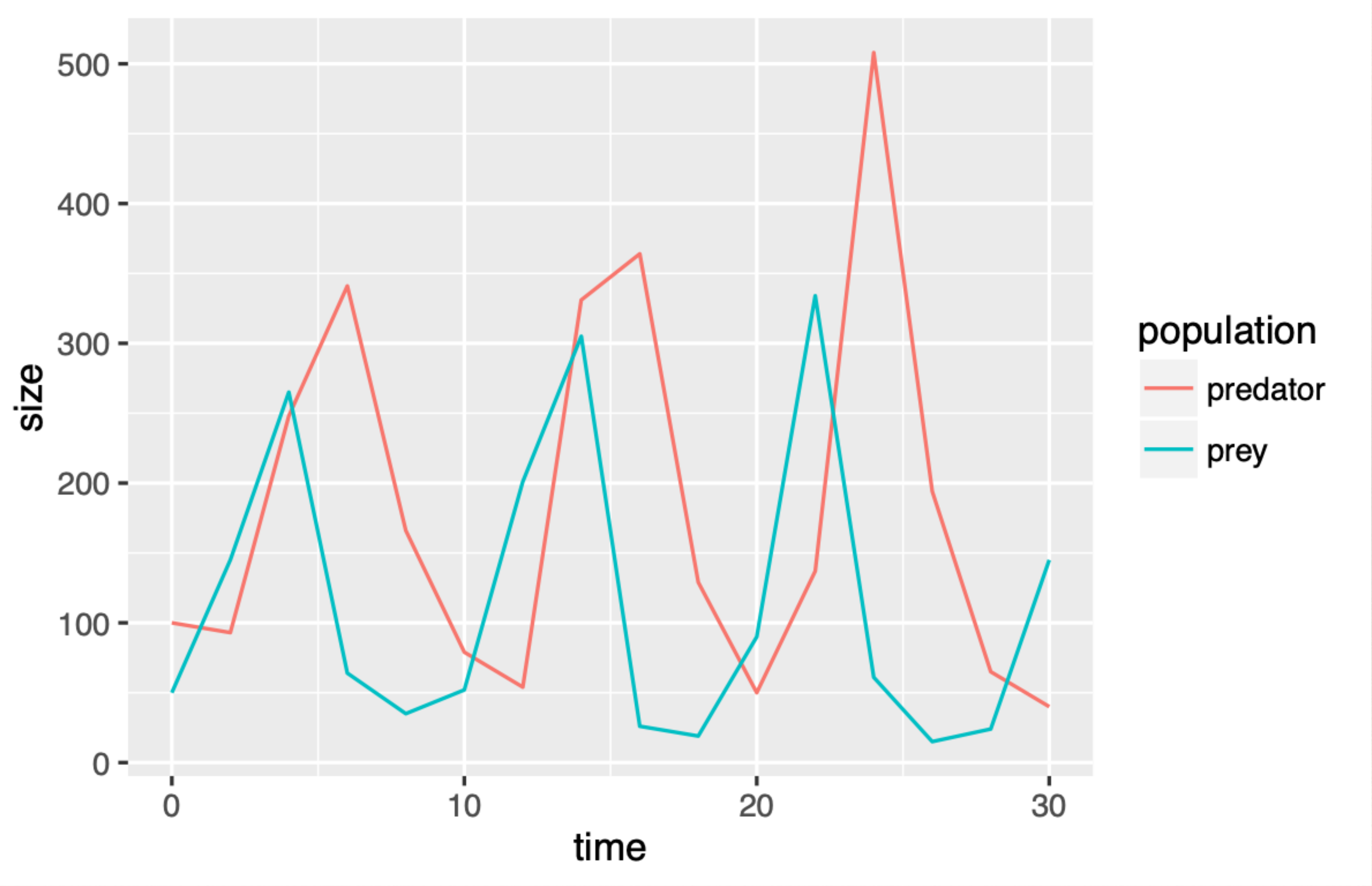
$$\boxed{l_{\text{ABC}}\left(y\,|\,\theta\right)=\int_{y}f\left(y\,|\,\theta\right)K_{\epsilon}\left(S(y_{\text{obs}})\,|\,S(y)\right)dy}$$

- Let
$$K_{\epsilon}\left(S(y_{\text{obs}})\,|\,S(y)\right)=\mathcal{N}\left(S(y_{\text{obs}})\mid S(y),\epsilon\mathbf{I}_{d_S}\right).$$
- Recall we set up IEnKI to estimate
- $$Z=\int_{x}p\left(x\right)l\left(x\right)dx$$
- To estimate $l_{\text{ABC}}\left(y\,|\,\theta\right)$, we can use IEnKI:
  - simulate M points $f\left(y\,|\,\theta\right)$, then find $S\left(y\right)$ for each;
  - perform IEnKI steps on each $S\left(y\right)$ until we reach the desired target.
- Note:
  - we only need to simulate from f the same number of times as in standard ABC;
  - the IEnKI moves on the summary statistics.

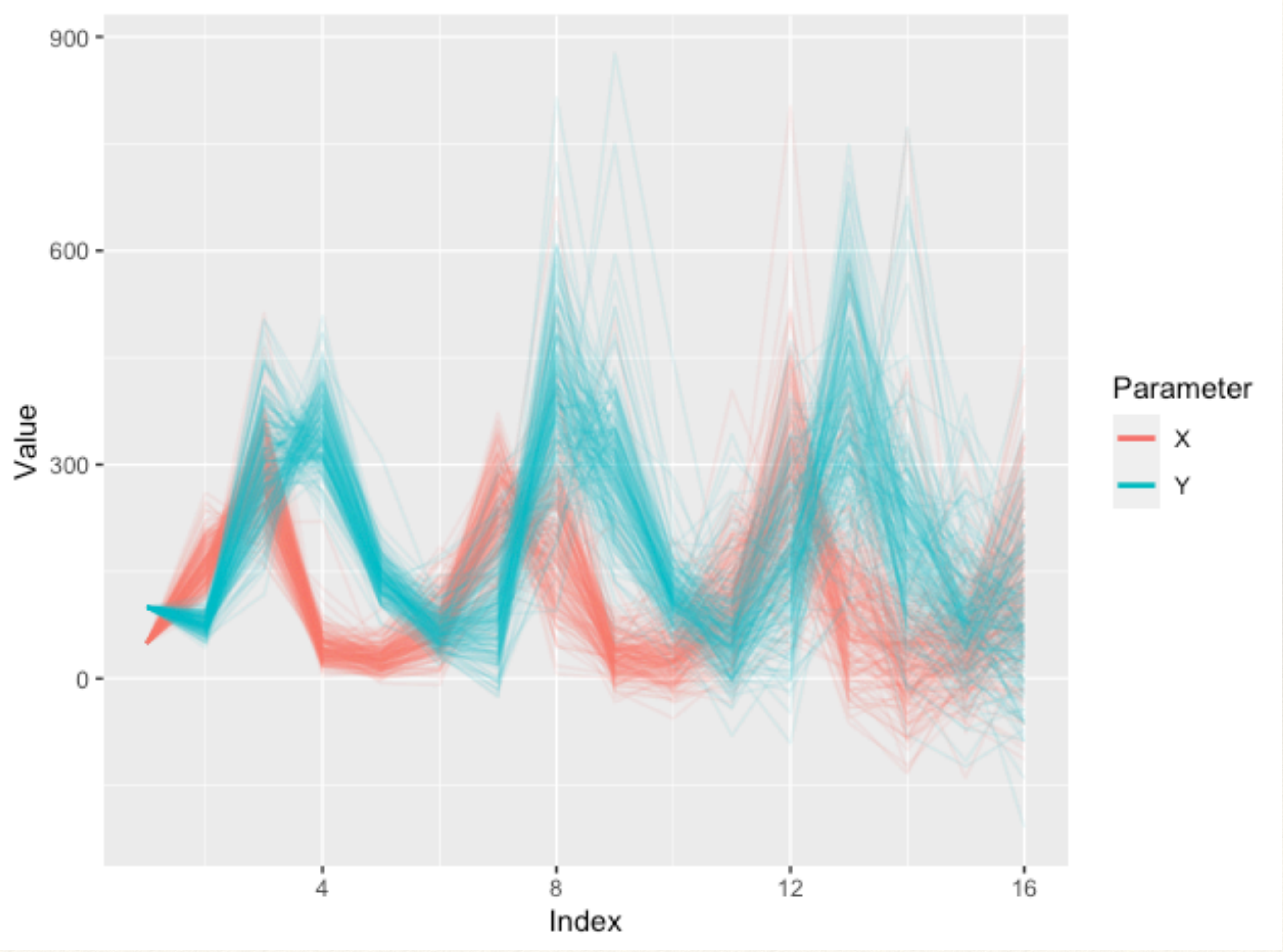
IEnKI-ABC

- For each θ…
- Simulate points from the model and take the summary for each. For $j = 1 : M$
$$y^{(j)} \sim f\left(\cdot \mid \theta\right) \qquad s^{(j)} = S\left(y^{(j)}\right).$$
- Perform an ensemble Kalman update of these points for $t = 0 : T - 1$.
- At iteration $t + 1$, for $j = 1 : M$
$$s_{t+1}^{(j)} = s_t^{(j)} + \hat{K}_{t+1}\left(s_{\text{obs}} - \tilde{s}_{t+1}^{(j)}\right)$$
where
$$\tilde{s}_{t+1}^{(j)} \sim \mathcal{N}\left(s_t^{(j)}, \left(\alpha_{t+1} - \alpha_t\right)^{-1} \epsilon \mathbf{I}_{d_S}\right)$$
and
$$\hat{K}_{t+1} = \hat{C}_t^{s_t s_t}\left(\hat{C}_t^{s_t s_t} + \left(\alpha_{t+1} - \alpha_t\right)^{-1} \epsilon \mathbf{I}_{d_S}\right)^{-1}$$
$$\hat{C}_t^{s_t s_t} \frac{1}{M-1} \sum_{j=1}^{M}\left(s_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} s_t^{(j)}\right)\left(s_t^{(j)} - \frac{1}{M}\sum_{k=1}^{M} s_t^{(j)}\right)^T$$

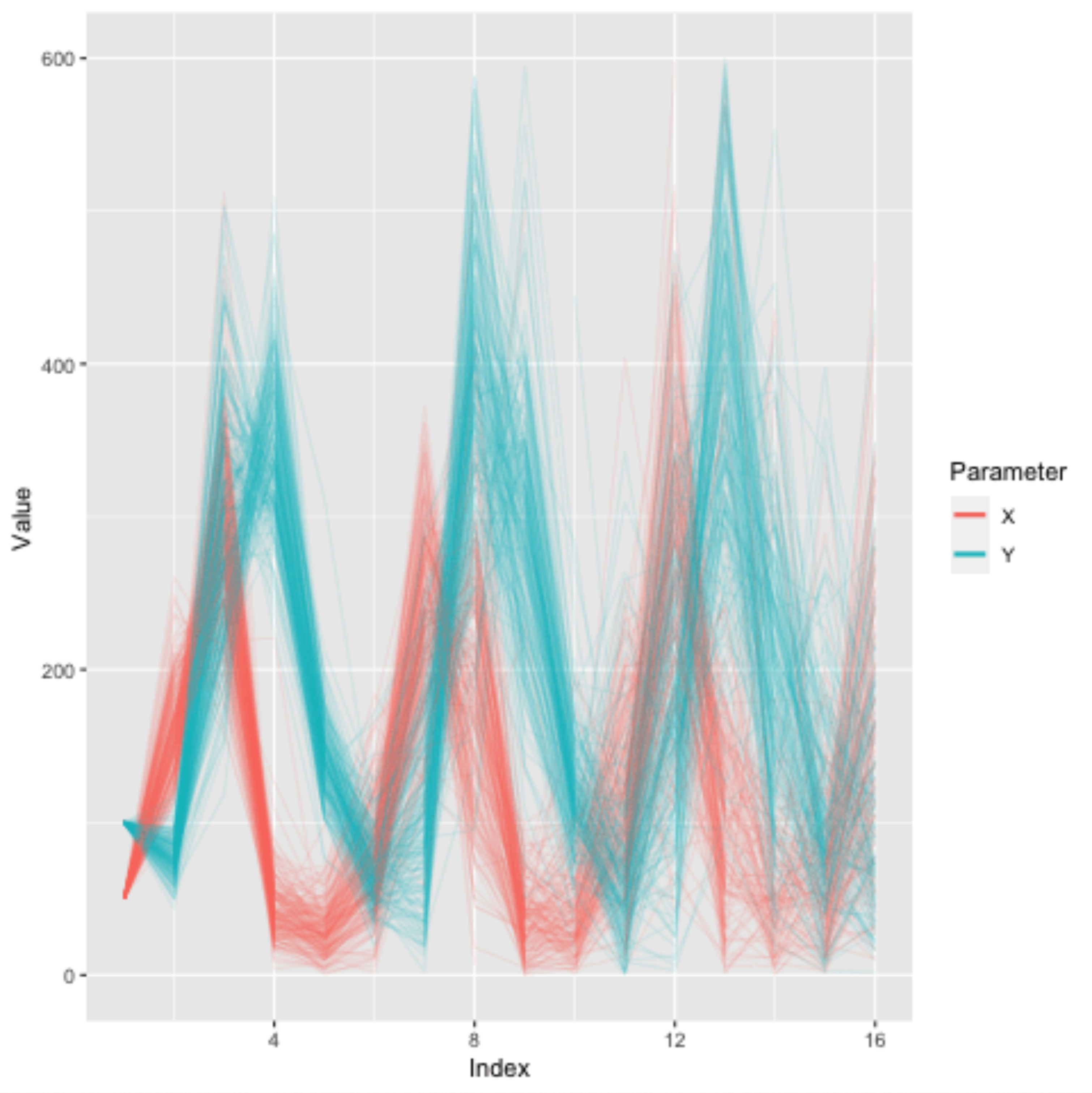
Modelling animal populations



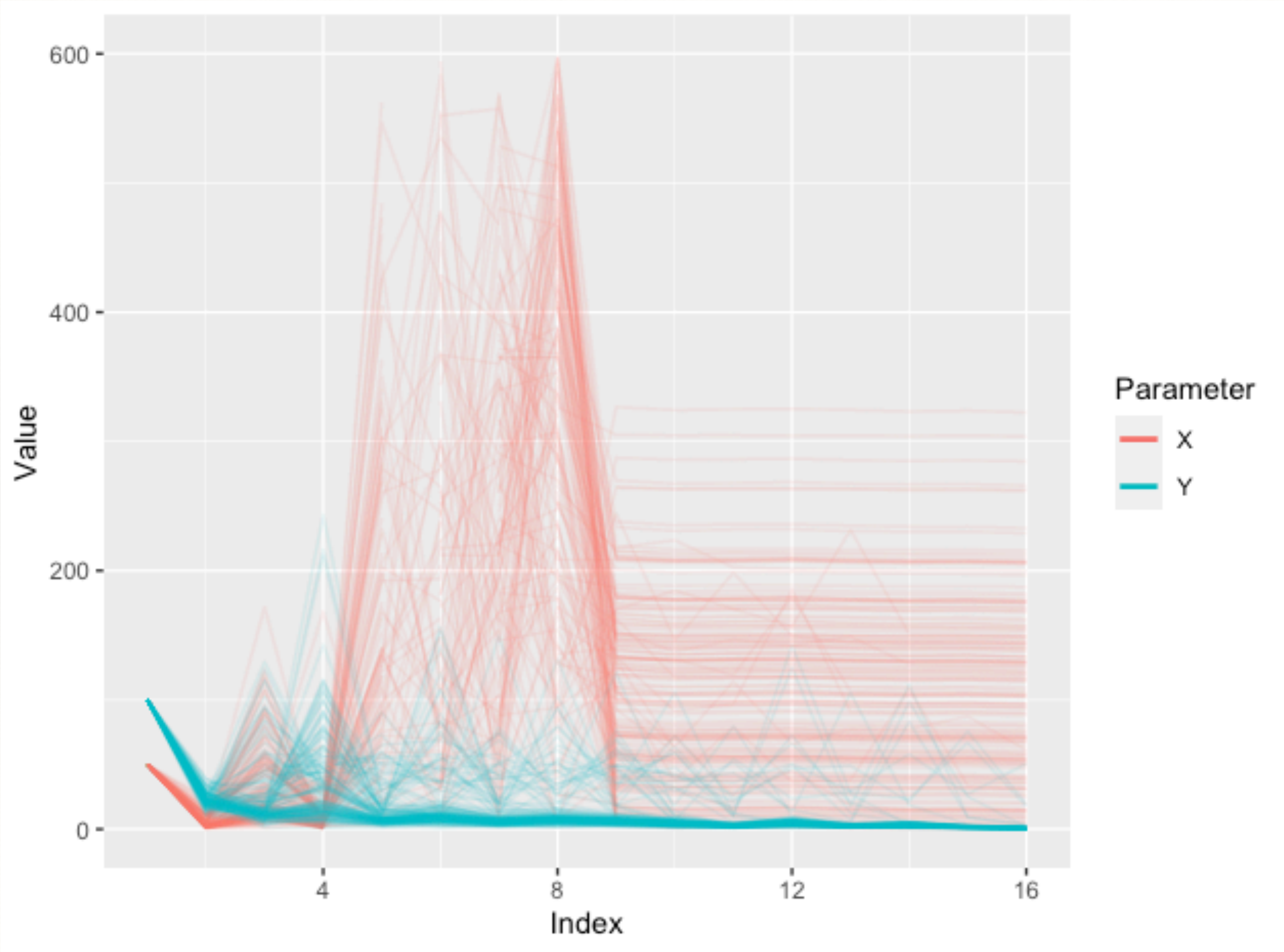
ABC simulation



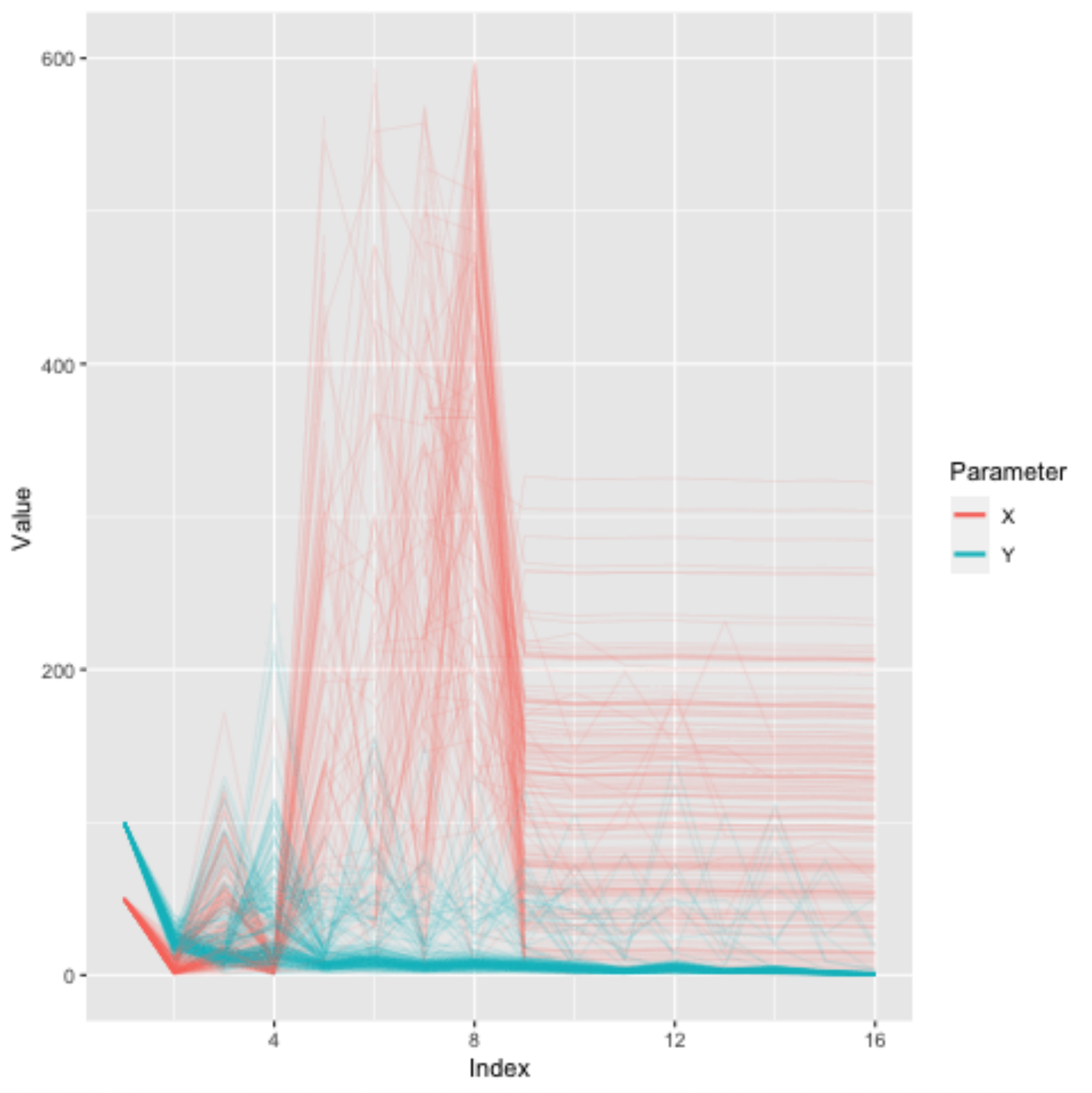
IEnKI-ABC simulation



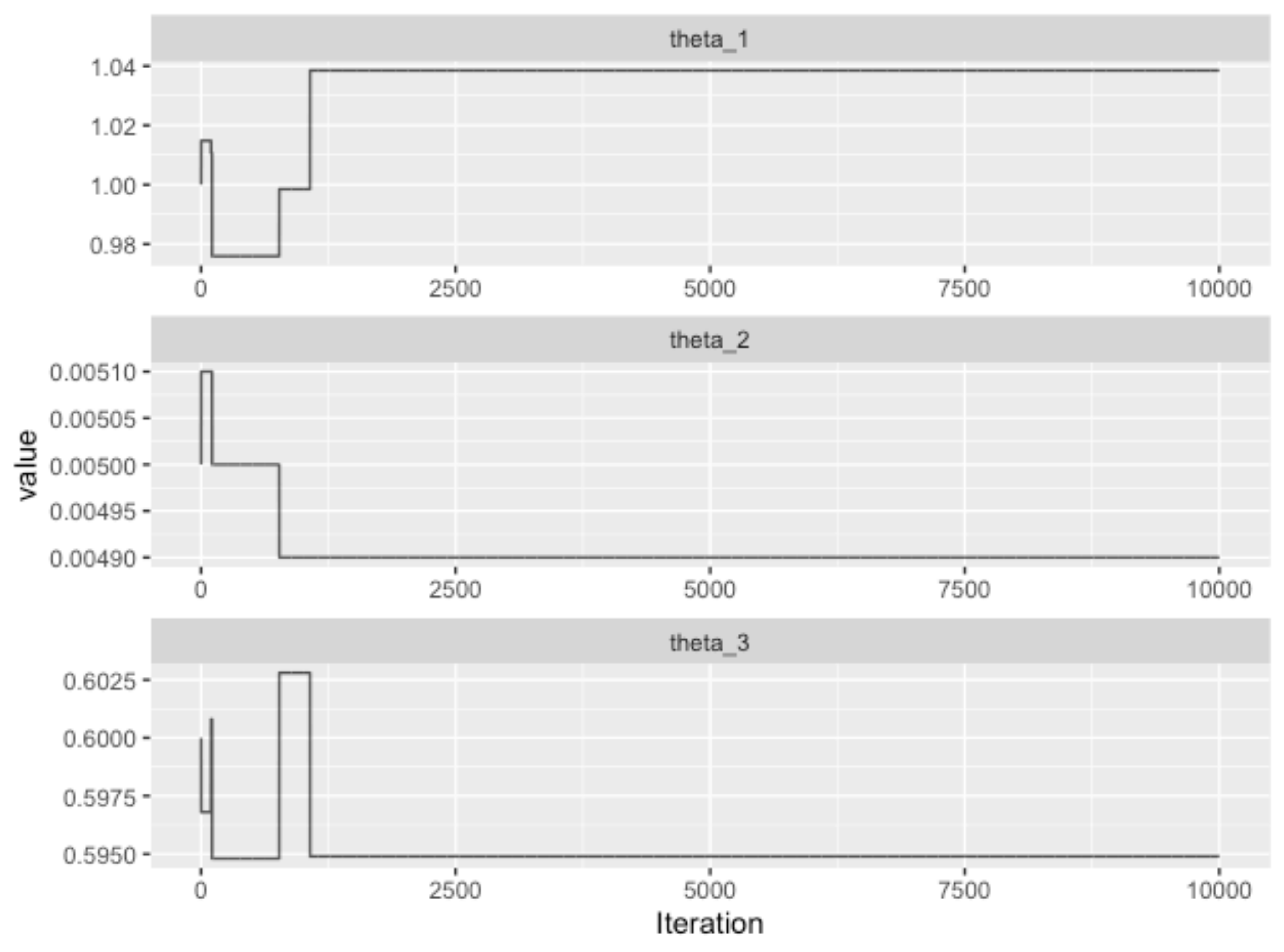
ABC simulation



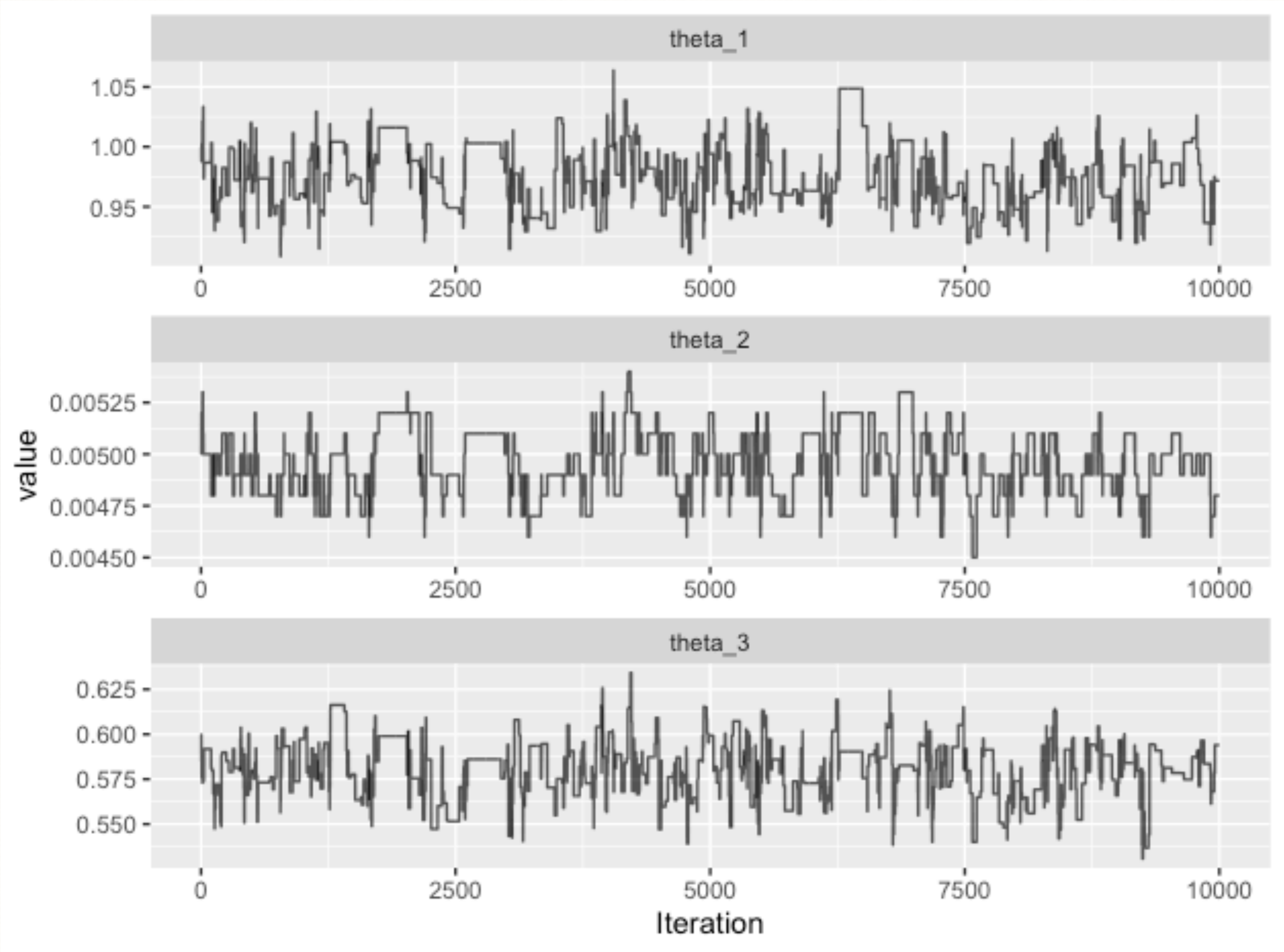
IEnKI-ABC simulation



ABC-MCMC results



IEnKI-ABC-MCMC results



Conclusions

- An improved pseudo-marginal-style ABC-MCMC
  - scales to higher dimensions than ABC;
  - uses an IEnKI “correction” on simulations from the model - IEnKI runs on the observation/summary statistic space (contrast to particle filter);
  - IEnKI can choose the sequence of targets adaptively - only one additional parameter to ABC.
- **Software:** ilike, ggsmc.