Joint state–parameter estimation for nonlinear stochastic energy balance models

Fei Lu [1] Nils Weitzel[2] Adam Monahan[3]

[1]Department of Mathematics, Johns Hopkins

[2]Meteorological Institute, University of Bonn, Germany

[3]School of Earth and Ocean Science, University of Victoria, Canada

SIAM Minisymposium on Data Assimilation: Theory and Practice
JMM 2019, January 17

Outline

1. Motivation
   - Stochastic energy balance model
   - State space model representation
2. Bayesian inference
   - Particle MCMC
   - Regularized posterior
3. Numerical study
   - Diagnosis of Markov Chain
   - Parameter estimation
   - State estimation

Motivation

Paleoclimate: reconstruct past climate temperature from proxy data

- Spatio-temporal evolution
  - spatial correlations
  - physically laws: energy balance $\to$ SPDEs
- Sparse and noisy data
  - Proxy data: historical data, tree rings, ice cores, fossil pollen, ocean sediments, coral etc.

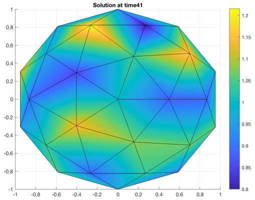
Plan: inference of SPDEs from sparse noisy data

- joint state-parameter estimation

Idealized atmospheric energy balance (Fanning&Weaver1996)

$$
\begin{aligned}
\partial_t u &= \underbrace{Q_T}_{\text{transport}} + \underbrace{Q_{SW}}_{\substack{\text{absorbed}\\ \text{shortwave}}} + \underbrace{Q_{SH}}_{\substack{\text{sensible}\\ \text{heat}}} + \underbrace{Q_{LH}}_{\substack{\text{latent}\\ \text{heat}}} + \underbrace{Q_{LW}}_{\substack{\text{longwave}\\ \text{surf.}\to\text{atmos.}}} - \underbrace{Q_{LPW}}_{\substack{\text{longwave}\\ \text{into space}}} \\
&= \nabla \cdot (\nu \nabla u) + \underbrace{\theta_0 + \theta_1 u + \theta_4 u^4} + W(t,x)
\end{aligned}
$$

- $u(t,x)$ normalized temperature (≈ 1)
- $\theta = (\theta_k)$: unknown parameters:
  - prior: a range of physical values
- $W(t,x)$: Gaussian noise,
  - white-in-time Matern-in-space



Data: observation model

Observation at sparse locations/regions:

$$y_{t_i} = \int_{A_i} u(t_i, x)\, dx + V_i,$$

- $\{A_i\}$ are regions/locations of observations
- Gaussian noise $\{V_i\}$, iid, variance known
- Linear operator in state u

State space model formulation

$$\text{SEBM:} \quad \partial_t u = \nabla \cdot (\nu \nabla u) + \sum_{k=0,1,4} \theta_k u^k + W(t,x)$$

$$\text{Observation data:} \quad y_{t_i} = H(u(t_i, x)) + V_i$$

Discretization (simplification):

- finite elements in space
- semi-backward Euler in time

State space model

$$\text{SEBM:} \quad U_n = g(\theta, U_{n-1}) + W_n$$

$$\text{Observation data:} \quad Y_n = HU_n + V_n$$

Goal: Given $y_{1:N}$, we would like to jointly estimate $(\theta, U_{1:N})$

- Bayesian approach to quantify uncertainty

Joint state-parameter estimation

Bayesian approach:

$$p(\theta, u_{1:N} | y_{1:N}) \propto p(\theta) p(u_{1:N}|\theta) p(y_{1:N}|u_{1:N})$$

- Posterior: quantifies the uncertainties

Approximate the posterior by sampling

- high dimensional ($> 10^3$),
- non-Gaussian, mixed types of variables $\theta, u_{1:N}$
- Gibbs Monte Carlo: $U_{1:N}|\theta$ and $\theta|U$ iteration
  - $U_{1:N}|\theta$ needs highD proposal density $\rightarrow$ Sequential MC
  - combine SMC with Gibbs (MCMC) $\rightarrow$

  Particle MCMC methods based on conditional SMC

Sampling: particle MCMC

Particle MCMC (Andrieu&Doucet&Holenstein10)

- Combines Sequential MC with MCMC:
  - SMC: seq. importance sampling $\rightarrow$ highD proposal density
  - conditional SMC: keep a reference trajectory in SMC
  - MCMC transition by conditional SMC
    $\rightarrow$ target distr invariant even w/ a few particles
- Particle Gibbs with Ancestor Sampling (Lindsten&Jordan&Schon14)
  - Update the ancestor of the reference trajectory
  - Improving mixing of the chain

However, standard Bayesian approach does not work:

for a Gaussian prior $p(\theta)$,
unphysical samples of posterior: systems blowing up

However, standard Bayesian approach does not work:

for a Gaussian prior $p(\theta)$,
unphysical samples of posterior: systems blowing up

Parameter estimation is ill-posed:

Singular Fisher infomation matrix for full perfect observation

$\rightarrow$ large oscillation in sample θ from Gibbs $\theta | \widehat{U}_{1:N}$

Regularized posterior

Recall the regularization in variational approach

Variational: $$(\widehat{\theta}, \widehat{u}_{1:N}) = \arg\min_{(\theta, u_{1:n})} C_{\lambda, y_{1:N}}(\theta, u_{1:N})$$

Bayesian : $$p_\lambda(\theta, u_{1:N} | y_{1:N}) \propto p(\theta)^\lambda p(y_{1:N} | u_{1:N}) p(u_{1:N} | \theta)$$

$$C_{\lambda, y_{1:N}}(\theta, u_{1:N}) = \underbrace{\lambda \log p(\theta)}_{\text{regularization}} + \underbrace{\log[p(y_{1:N} | u_{1:N}) p(u_{1:N} | \theta)]}_{\text{likelihood}}$$

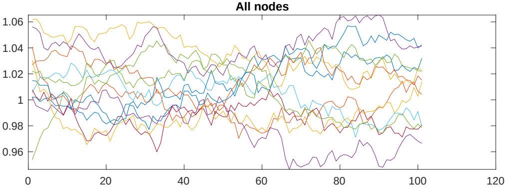
- $\lambda = 1$: Standard posterior $\xrightarrow{N \to \infty} \sim$ likelihood
- $\lambda = N$: regularized posterior

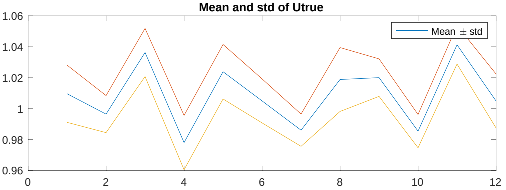
Physical parameter set up:

- Gaussian prior

	θ_0	θ_1	θ_4
mean	30.11	-24.08	-5.40
std	0.82	0.46	0.20

- temperature near an equilibrium point (normalized, ≈ 1)





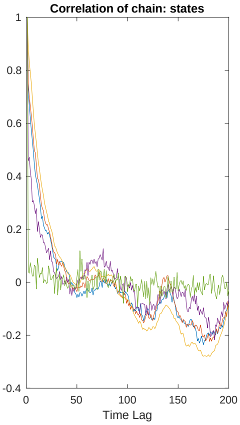
Dimension of the states: 1200

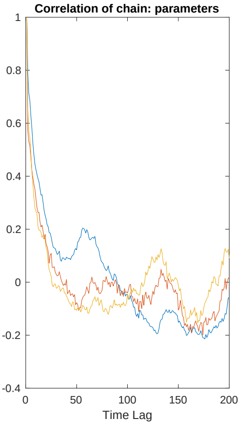
- 12 spatial nodes
- 100 time steps
- observe 6 nodes each time;

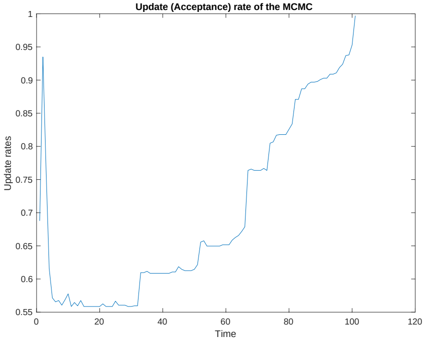
Randomly generate a true value for parameter from prior

Diagnosis of Markov Chain

Chain length: 1000 (with 30% burnin)



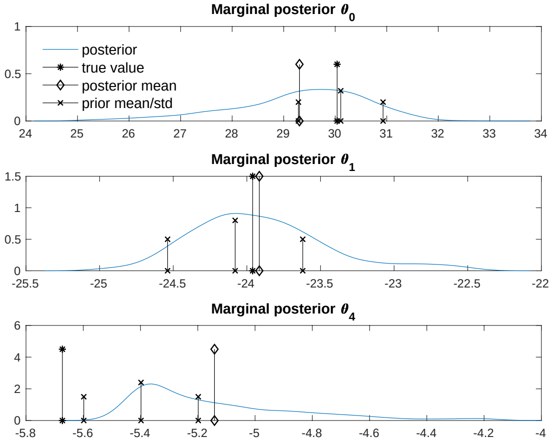




State update rate: > 0.55

Correlation length: 10-30 steps

Parameter estimation



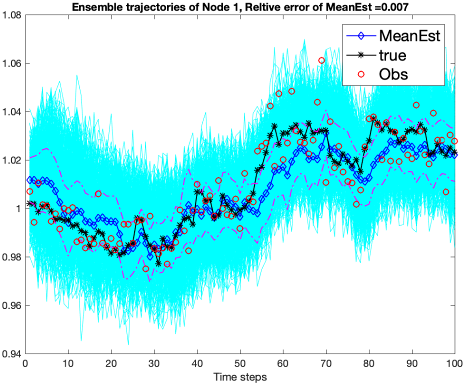
Marginals of the posterior

- θ_0, θ_1 OK, +bias θ_4
- posterior close to prior
- Errors in 100 simulations

	θ_0	θ_1	θ_4
Mean	-0.74	0.11	0.22
Std	0.73	0.46	0.20

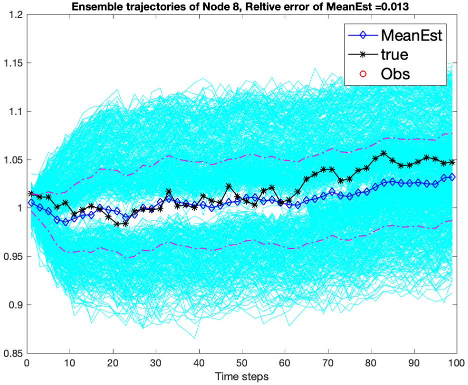
  - -bias θ_0, +bias in θ_4

Ensemble of sample trajectories:



Observed node:

filtered out noise



Unobserved node:

large spread, mean close to truth

Observe more or less nodes

When more modes are observed:

- State estimation gets more accurate
- Parameter estimation does not improve much:
  the posterior keeps close to prior.

Summary and discussion

Bayesian approach to jointly estimate parameter-state

- a stochastic energy balance model
- sparse and noisy data
- Estimate both parameters and states
  - regularized posterior due to singular Fisher matrix
  - Gibbs sampling via PGAS

Summary and discussion

Bayesian approach to jointly estimate parameter-state

- a stochastic energy balance model
- sparse and noisy data
- Estimate both parameters and states
  - regularized posterior due to singular Fisher matrix
  - Gibbs sampling via PGAS

Results:

- State estimation:
  - filtered noise on observed nodes;
  - large uncertainty in unobserved modes
- Parameter estimation:
  - slightly biased estimators
  - posterior close to prior

Open questions

1. Re-parametrization: avoid singular Fisher information matrix?

2. How many nodes need to be observed (for large mesh)?
(theory of determining modes)

Open questions

1. Re-parametrization: avoid singular Fisher information matrix?

2. How many nodes need to be observed (for large mesh)?
(theory of determining modes)

Thank you!