Bayesian integrated population modeling using JAGS

State-space models

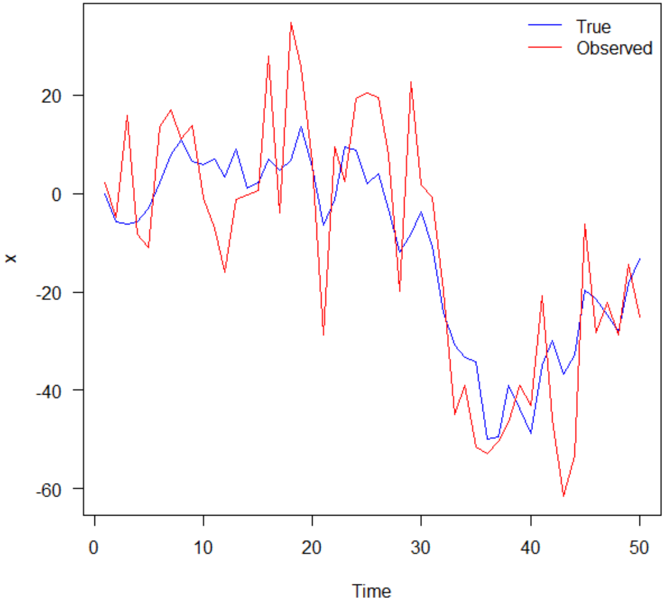
Introduction

- State-space models (SSM) provide a general framework for analyzing deterministic and stochastic systems that are measured or observed through a stochastic process
- Applied in engineering, statistics, computer science and economics to solve a broad range of dynamical systems problems
- The term state space originated in 1960s in the area of control engineering (Kalman 1960)
- A well studied SSM is the Kalman filter, which defines an algorithm for linear Gaussian systems
- SSM are also known as hidden Markov models or latent process models
- SSM are composed of 2 sets of equations:
  - *State process equations*: describe the development of the true state of a system over time
  - *Observation equations*: relate the true state to the observation

Example 1: random walk with noise

$$x_{t+1} \mid x_t \sim Normal\left(x_t, \sigma^2_{proc}\right)$$

$$y_t \mid x_t \sim Normal\left(x_t, \sigma^2_{obs}\right)$$



Data simulation with:

- $x_1 = 0$
- $\sigma_{proc} = 7$
- $\sigma_{obs} = 10$

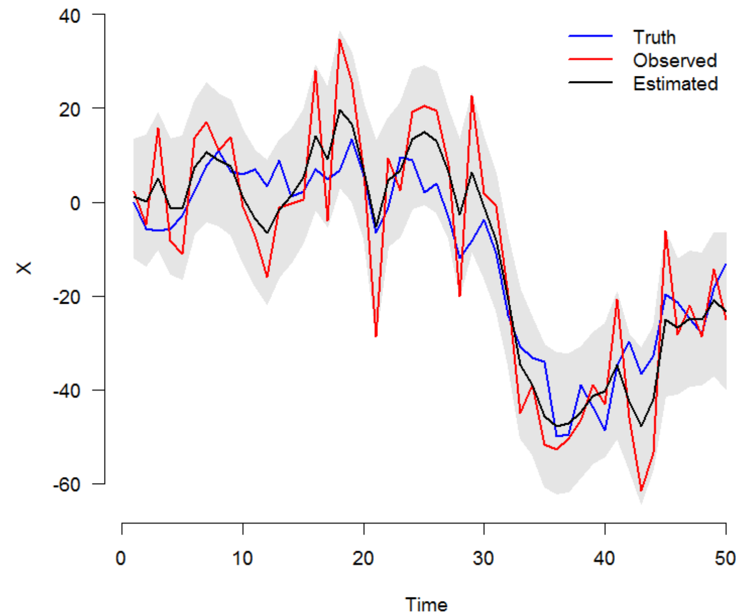
Example 1: random walk with noise

Analysis in JAGS

$$x_1 \sim Normal(0,100)$$

$$x_{t+1} \sim Normal(x_t, \sigma^2_{proc})$$

$$y_t \sim Normal(x_t, \sigma^2_{obs})$$



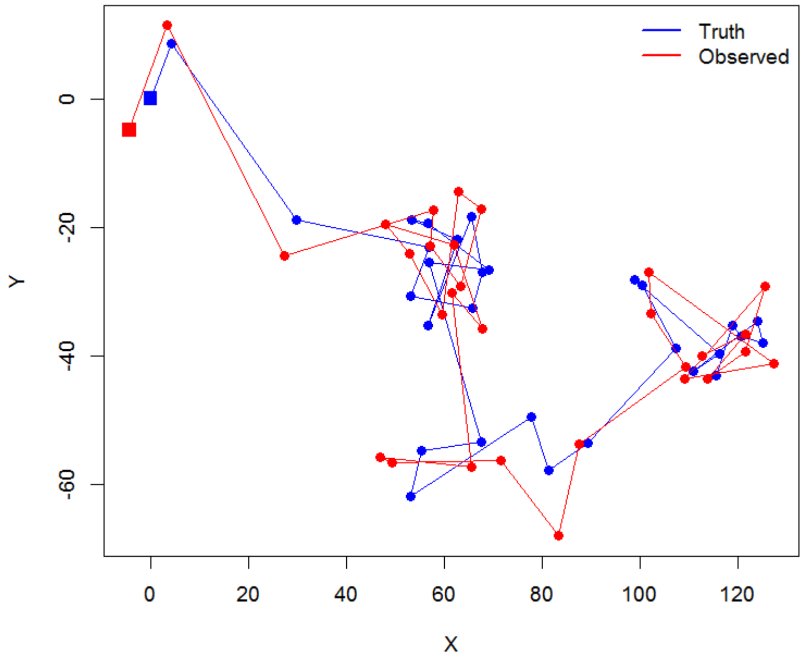
Estimates:

	mean	sd	2.5%	50%	97.5%	overlap0	f	Rhat	n.eff
x[1]	-1.457	6.310	-13.943	-1.494	10.397	TRUE	0.598	1.000	4002
...									
sigma1	7.521	2.449	3.787	7.110	13.267	FALSE	1.000	1.032	79
sigma2	11.539	1.893	7.672	11.549	15.266	FALSE	1.000	1.010	247

Example 2: movement pattern with noise

$$x_{t+1} \mid x_t \sim Normal\left(x_t, \sigma_x^2\right)$$
$$y_{t+1} \mid y_t \sim Normal\left(y_t, \sigma_y^2\right)$$

$$x.obs_t \mid x_t \sim Normal\left(x_t, \sigma_{obs}^2\right)$$
$$y.obs_t \mid y_t \sim Normal\left(y_t, \sigma_{obs}^2\right)$$

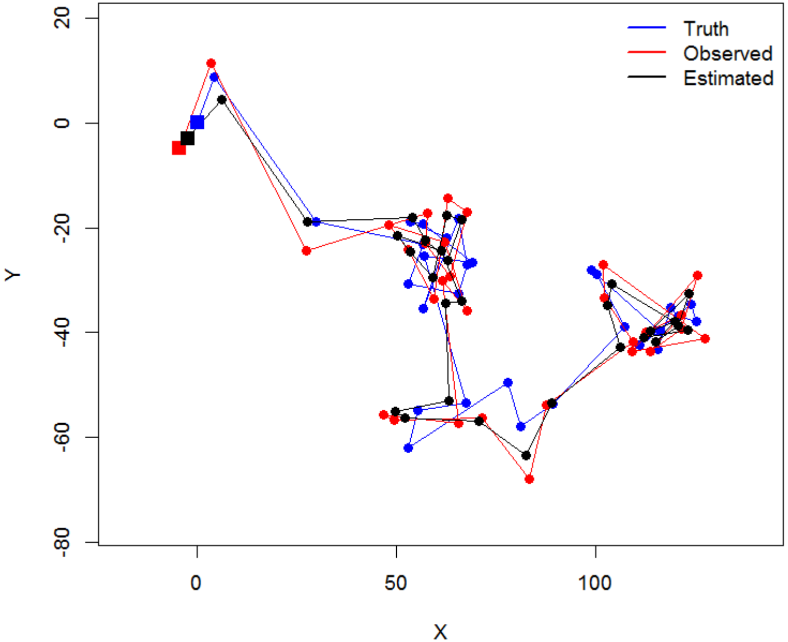


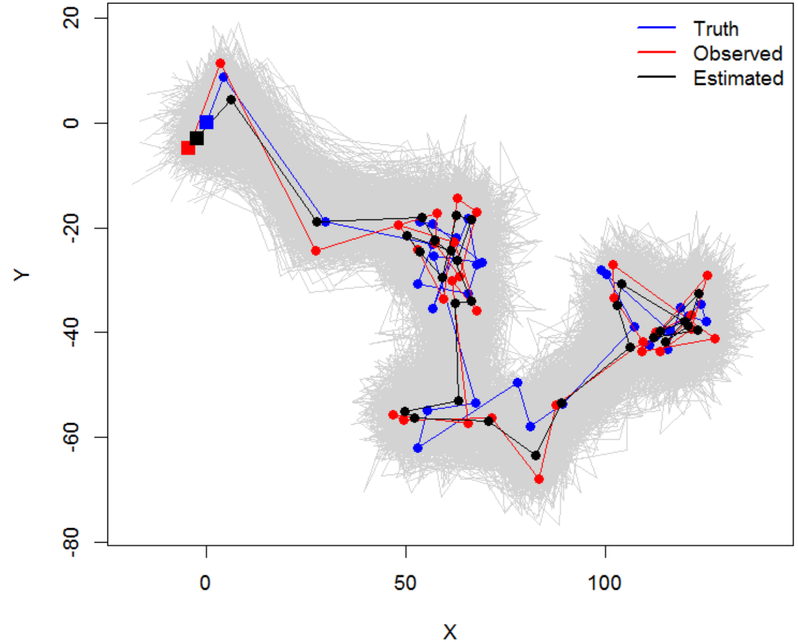
Data simulation with:

- $x_1 = 0,\ y_1 = 0$
- $\sigma_x = 10,\ \sigma_y = 12$
- $\sigma_{obs} = 5$

Example 2: movement pattern with noise

Analysis in JAGS





	mean	sd	2.5%	50%	97.5%	overlap0	f	Rhat	n.eff
sigmaX	12.692	2.148	9.017	12.572	17.528	FALSE	1.000	1.003	558
sigmaY	10.057	2.563	5.647	9.875	15.177	FALSE	1.000	1.000	1251
sigma.obs	4.699	2.114	0.440	4.893	8.558	FALSE	1.000	1.010	323

Application to count data

- Data: annual counts (a time-series) from a population
- Aims:
  - Understand dynamics of the population (population growth rate, density-dependence, impact of environment)
  - Predict the behaviour in the future
- Classical time-series analyses such as described in Dennis et al. (1991) or Lande et al. (2003)
- Challenge: count data are usually not free from observation errors
- Idea: use state-space models (SSM) to analyse such time-series of counts (e.g. Lindley 2003; Buckland et al. 2004; Dennis et al. 2006; Flesch 2014)

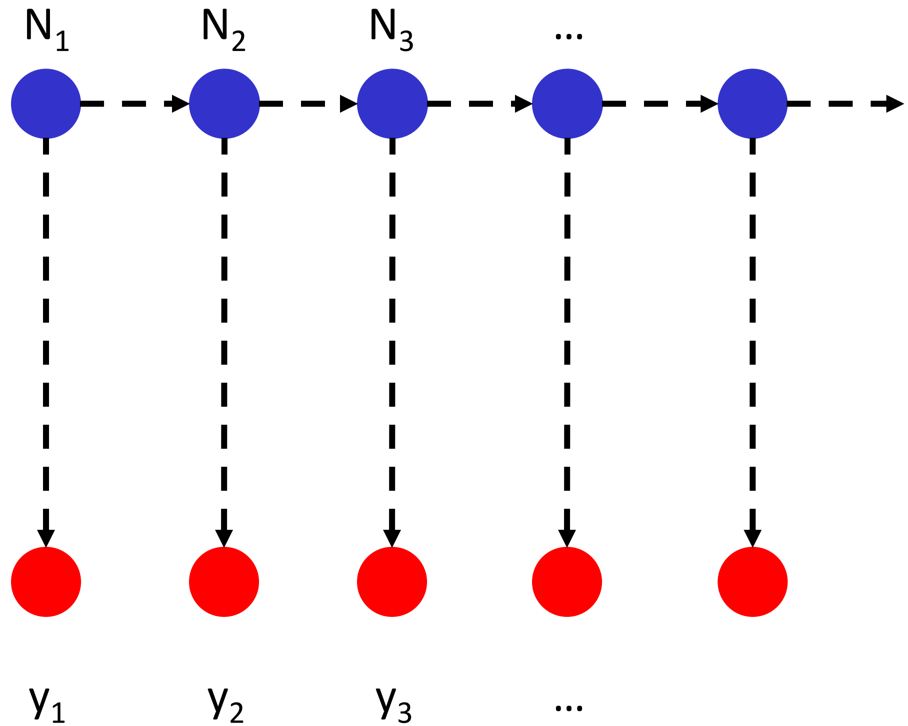
Time-series of counts

State process

True population size

Observation process

Observed count



Stochastic process

Deterministic process

State-space likelihood to analyse count data

1. State process

$$N_1 \sim Uniform(a,b)$$

$$N_{t+1} | N_t = \lambda_t N_t$$

$$\lambda_t \sim Normal(\bar{\lambda}, \sigma_\lambda^2)$$

where,
N_t : true population size at time t
λ_t : population growth rate between time t and t+1
$\bar{\lambda}$: mean population growth rate («population trend»)
σ_λ^2 : temporal variation of the population growth rate (environmental stochasticity)

2. Observation process

$$y_t | N_t \sim Normal(N_t, \sigma_y^2)$$

where,
y_t : population count at time t (our data)
σ_y^2 : residual variation («observation error»)

State-space likelihood to analyse count data (on log scale)

1. State process

$$\log(N_1) \sim \mathit{Normal}(5.6, 0.01)$$

$$\log(N_{t+1}) \,|\, \log(N_t) = \log(N_t) + r_t$$

$$r_t \sim \mathit{Normal}(\bar{r}, \sigma_r^2)$$

where,
N_t : true population size at time t
r_t : population growth rate (on log scale) between time t and t+1
$\bar{r}$: mean population growth rate («population trend»)
σ_r^2 : temporal variation of the population growth rate (environmental stochasticity)

2. Observation process

$$\log(y_t) \,|\, \log(N_t) \sim \mathit{Normal}(\log(N_t), \sigma_y^2)$$

where,
y_t : population count at time t (our data)
σ_y^2 : residual variation («observation error»)

Summary comments

- SSM are powerful
- Yet, estimation problems occur frequently, in particular
  - when the observation error is much larger than the process error (Auger-Méthé et al. 2016)
  - when autocorrelation of the state-process is weak (Auger-Méthé et al. 2016)
  - when models with density-dependence are fitted (Knape 2008)
- Potential solutions
  - use longer time-series …
  - add information about observation error
  - add information about the state-process → as in IPM