A state-space model simulation

Anthony Davidson

2020-01-08

Table of Contents

[Overview 2](#_Toc40797579)

[To do 2](#_Toc40797580)

[Introduction 2](#_Toc40797581)

[Methods 3](#_Toc40797582)

[Simulation scope 3](#_Toc40797583)

[Three 3](#_Toc40797584)

[Step Two 3](#_Toc40797585)

[Test datasets 4](#_Toc40797586)

[Data input 5](#_Toc40797587)

[Applications 10](#_Toc40797588)

[Fitting a single model 10](#_Toc40797589)

[Example one (beech/mice) 10](#_Toc40797590)

[Example two (weather/mice) 10](#_Toc40797591)

[Final Words 11](#_Toc40797592)

[References 11](#_Toc40797593)

[My notes 11](#_Toc40797594)

[Modelling differences 11](#_Toc40797595)

[Another simulation? 11](#_Toc40797596)

[Literature 12](#_Toc40797597)

[SSM with multiple sources 12](#_Toc40797598)

[The connection 12](#_Toc40797599)

# Overview

Here I am attempting to demostrate the affect of predictable differences in the size of variance in the observation part of a state-space model. The differences between the true underlying variance structure of these datasets is unknown but this is alway the case with true ecological experiments. In the ecological literature CR datasets can generate systamatic fitting issues (better or worse fitting models related to the variance alone).

## To do

1. Contruct a generic state-space modelling framework that allows for the partioning of observation and process error in the data
2. How does the differences between accuracy of abundance estimates and seed estimates?
3. This project can also be used as a simulation template for PFNZ2050.

myPaths <- .libPaths("C:/Program Files/R/R-3.6.2/library")  
myPaths <- c(myPaths)  
.libPaths(myPaths) # add new path  
.libPaths()

## [1] "C:/Program Files/R/R-3.6.2/library"

# Introduction

**Estimating abundance**

* Ecological questions focused on linking the observed abundance of species with the true abundance and the underlying interactions with other species, density and environmental factors
* By informing the dynamics as the visualisations of the process model we can understand the quality of observation (abundance estimates) are needed to capture the true underlying population dynamics.
* But estimating variance can be hard and a new thing to some degree (look at literature review agian too….)
* An this has been acknowledged in both the statistics and ecology literature by using case studies and simulations. ….?

**Intergrated population modelling**

* Using an intergrated population model we can use information from multiple different sources and intergrate these sources of information into our SSM.

**Variance estimation**

* Often the target of field work is to reduce this
* However maybe general processes are better explained at a different (broader scale??)

# Methods

Here is the simple set of steps to simulate a state-space model (SSM).

## Simulation scope

I am interested in the process model:

The first step was to add a collection of common but simple model structures of abundance and resource into a single SSM for comparing the differences between variance of abundance estimates.

We parameterise this model with the following structure:

#add process diagram here. With equation associated with it

### Two

Varying observational data from the following common observational models (CR abundance ) for the data () as:

### Three

Use the observed data to see what happens..

## Step Two

Defining the model in JAGS.

mod <- "model {  
  
 # Observation model  
 for(i in 1:nYears) {  
 for(j in 1:nGrids) {  
 Y[i, j] ~ dnorm(X[i], tau[1])  
 }  
 }  
  
 X[1] ~ dnorm(0, 1)  
 simX[1] <- X[1]  
  
 # Process model  
 for(i in 2:nYears) {  
  
 X[i] ~ dnorm(predX[i], tau[2])  
 predX[i] <- X[i-1] + b + beta.res \* seed[i-lag] - d \* X[i-1]  
  
 simX[i] ~ dnorm(predsimX[i], tau[2]) # fitted model  
 predsimX[i] <- simX[i-1] + b + beta.res \* seed[i-lag] - d \* simX[i-1]  
 }  
  
 b ~ dnorm(0, 0.001)T(0, )  
 beta.res ~ dnorm(0, 0.01)  
 d ~ dnorm(0, 0.01)T(0, )  
 mu.r ~ dnorm(0, 0.001)  
  
 for(i in 1:2) {  
 tau[i] <- 1 / (sigma[i] \* sigma[i])  
 sigma[i] ~ dunif(0, 10)  
 }  
  
 }"#model finish  
   
 write(mod, "./models/mod.txt")

## Test datasets

The data needed for this analysis is as follows:

Mouse abundance: A dataset with the estimated abundance (N), se of the estimate N (se.N) and the time of observation.

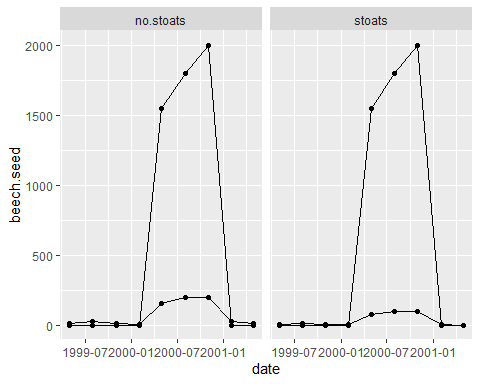
Seed abundance: A dataset with the estimated seedfall (on the same scale as the abundance estimation).

### Data input

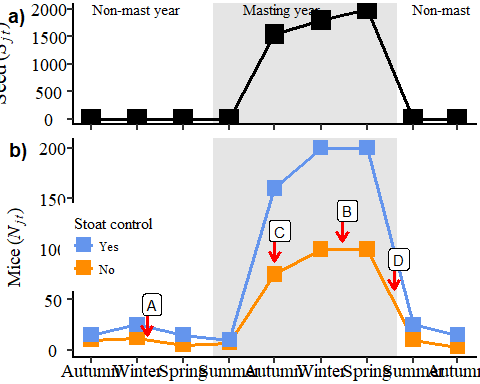
There are all assumed to be capture/recapture datasets for now.

#### Simulated dataset

# glimpse(dat1)  
ggplot(dat1, aes(x = date, y = beech.seed)) +  
 geom\_point() +   
 geom\_line() +  
 geom\_point(data = dat1, aes(x = date, y = value)) +  
 geom\_line(data = dat1, aes(x = date, y = value)) +  
 facet\_wrap(~control)



**PLOT**

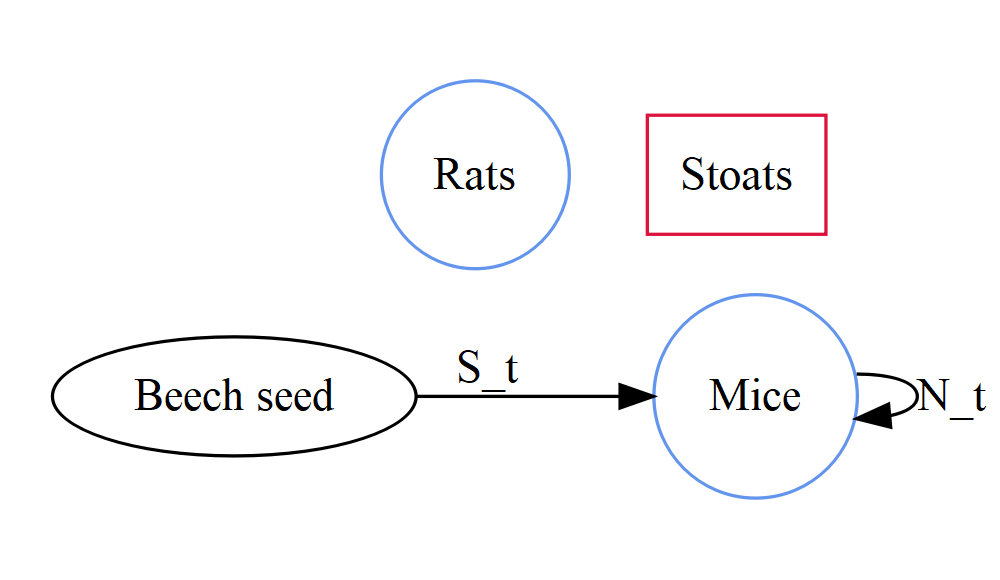


#### Beech dataset

**Dynamic relationships**

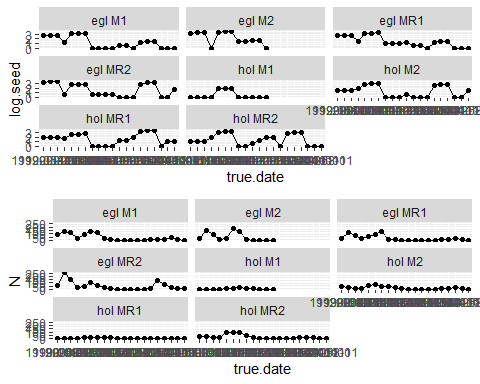
# long.dat <- gather(data = red.dat, value = "abund",key = "spp", c(N:log.seed))  
# long.dat

library("DiagrammeR")  
  
grViz("   
 digraph CFA { graph [layout = circo]  
 # Multiple level nodes  
   
 # # Terminal branch nodes  
 node [shape = ellipse, color = Drakgreen]  
 a [label = '@@1']  
   
 node [shape = circle, color=CornflowerBlue]  
 b [label = '@@2'];   
 c [label = '@@3'];  
 {rank = same; b; c}  
   
 # Terminal branch nodes  
 node [shape = box, color = Crimson]  
 d [label = '@@4'];  
   
 # Connect nodes with edges and labels  
 a -> c [label = 'S\_t']  
 c -> c [label = 'N\_t']  
 }  
  
[1]: 'Beech seed'   
[2]: 'Rats'  
[3]: 'Mice'  
[4]: 'Stoats'  
")

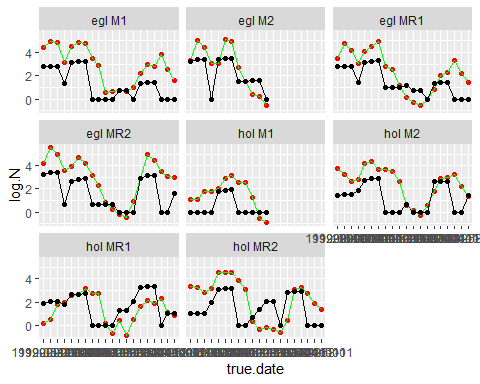


where:

??

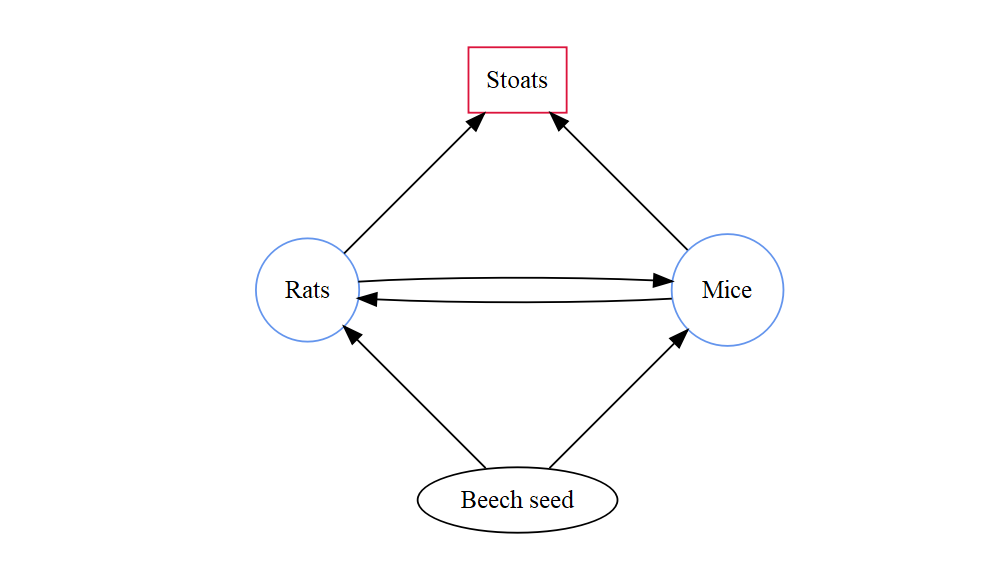


# p3 <-   
 ggplot(red.dat, aes(x = true.date, y = log.N, group = grid)) +  
 geom\_point(colour = "red") +  
 geom\_line(colour = "green") +  
 geom\_point(data = red.dat, aes(x = true.date, y = log.seed, group = grid)) +  
 geom\_line(data = red.dat, aes(x = true.date, y = log.seed, group = grid)) +  
 facet\_wrap(~grid)



#### MPD dataset

[coming]



#### Other Options

*[coming]*

# Applications

Some *significant* applications are demonstrated in this chapter.

## Fitting a single model

Here I have began with a dataset that has already been fitted to this dataset with a simplier model. **(Credit Em and others in team?!)**

### Example one (beech/mice)

# mod <- jags(  
 # model.file = "./models/mod.txt",  
 # data = list(  
 # Y = Y,  
 # J = J,  
 # z = z,  
 # N\_grid = N\_grid,  
 # N\_animals = N\_animals,  
 # samp = samp,  
 # ny = ny  
 # ),  
 # parameters.to.save = c("mu.p", "sigma.p", "N"),  
 # n.chains = 5,  
 # n.iter = 2000,  
 # n.burnin = 500,  
 # n.thin = 1,  
 # parallel = T  
 # )  
   
 # mod2  
 # mod2.sum <- mod2$summary

### Example two (weather/mice)

# mod <- jags(  
 # model.file = "./models/mod.txt",  
 # data = list(  
 # Y = Y,  
 # J = J,  
 # z = z,  
 # N\_grid = N\_grid,  
 # N\_animals = N\_animals,  
 # samp = samp,  
 # ny = ny  
 # ),  
 # parameters.to.save = c("mu.p", "sigma.p", "N"),  
 # n.chains = 5,  
 # n.iter = 2000,  
 # n.burnin = 500,  
 # n.thin = 1,  
 # parallel = T  
 # )  
   
 # mod2  
 # mod2.sum <- mod2$summary

# Final Words

We have finished a nice book.

# References

# My notes

## Modelling differences

**between discrete (observations in time) and dynamic models (continuous in time)**

Interested in the relationship between:

**and…**

### Another simulation?

THis will address repeatibility>?? See eDNA example of this?

We will start with very simple model of mice CR data as the base data and augmentation 5 ,10 ,100 ,500 ,1000 zeros to the longest line of captures.

1. And then we will augmentation the same numbers of zeros but even independent of the number of captures in the trip.
2. Modify p
3. Use observed CR data and augment across same range

# Literature

Here are the key papers that suggest that this is an important question to address.

## SSM with multiple sources

## The connection