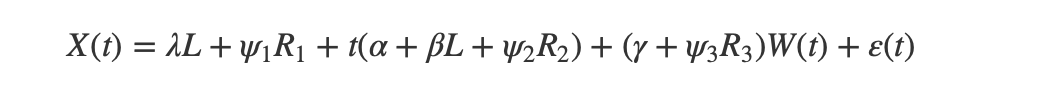
Comments on simulation model 5Aug21

Original model



Maybe set up as

W(t) = r1 W(t-1) + d1(t)

W2(t) = r2 W2(t-1) + d2(t)

U(t) = hW(t) + W2(t)

X(t) = l0 + l1L + fR + (b0 + b1L)t + gW(t) + sUU(t) + see(t)

where W(t) is measured and U(t) isn't. Therefore, U(t) serves as the source of spatial and temporal autocorrelation in the random variation. Note that U(t) can contain W(t) so that there is correlation between the errors and W(t) to make estimation of g hard.

e ~ Normal iid

d1, d2 ~ Normal or Cauchy (see below) scaled so the variance is 1

S = f(D,theta) [exclude the nugget] where f() is the tapered covariance function

R is the 2D sine function + a random term in which Tx = Ty can be manipulated to change the wavelength relative to the granularity.

Granularity fixed at 8

For time-series simulations, have a burn-in time of 10 years.

Cauchy distribution (same as a t-distribution with 1 df) - truncated at abs() < 10

Note: the correlation for V has to be high to get correlations in E

library(mvtnorm)

nrep <- 10^4

V1 <- matrix(c(1,.99,.99,1), 2, 2)

# V2 <- matrix(c(1,0,0,1), 2, 2)

X <- matrix(NA,nrow=nrep, ncol=2)

for(i in 1:nrep){

e1 <- t(rmvnorm(n=1, sigma=V1))

e2 <- t(rmvnorm(n=1, sigma=V1))

X[i,] <- e1/abs(e2)

}

XX <- X[abs(X[,1]) < 10 & abs(X[,2]) < 10,]

par(mfrow=c(2,1))

hist(XX[,1],1,breaks=20)

plot(XX)

cor(XX)

cov(XX)

Provisional list of experiments; these really are provisional, since you know more about the simulations.

**Sim1a:** Test the impact of map size on PARTs ability to estimate only spatial effects.

a. Map widths of 104, 144, 200, and 280 pixels

b. Beta0 = beta1 = gamma = 0

c. Random term d2 ~ Normal

d. Set g = h = f = 0. I'm not sure the base values for the other parameters, but use the ones you have been using if they give reasonable simulations. Well, maybe increase q to 0.1 to give more spatial autocorrelation.

e. Set l1 = 0 and l1 > 0 in two sets of simulations; for the l1 > 0, it would be nice to pick l1 so that the rejection rate for a significance level of 0.05 was about 30% for the smallest maps, so this simulation would show the effect of map size on power.

f. Plot simulation example of U

g. Plot est l1 - sim l1 and rejection rates if you can do 500 simulations)

Sim1b: Test the impact of non-gaussian errors on estimating only spatial effects.

a. Do a study like sim1a with 104 pixel maps, but with random term d2 ~ T3

b. Plot est l1 - sim l1; only do rejection rates if you can do 500 simulations (i.e., these aren't as important)

c. plot comparison of l1 among sim 1a (104 px) and sim 1b (i.e., compare gaussian vs T)

Sim1c: Test the impact of introducing random spatial variation to the spatial parameter.

\* (this replaces old sim2 that looks a granularity, doing so without time; it is also similar to old sim 3 in looking at grain)

a. Do a study like sim1a with 104 pixel maps, but f > 0 (R included), with Tx = Ty picked to give 1, 22, 32, ..., 82 peaks. At 82 peaks, estimates of l1 should be bad, since this is the same scale as the granularity.

b. Plot est l1 - sim l1 and rejection rates (only if you can do 500 simulations), including comparison from when f = 0 from sim 1a

Sim2a: Test the effect of temporal autocorrelation on estimates of **intercept** time trends

a. 30 time points

b. g = h = 0

c. lambda = gamma = 0

d. b0 = 0 and b0 = 1/30; b1 = 0

e. To make things hard, use r2 = 0 and r2 = 0.6 (i.e., temporally autocorrelated U)

f. Plot est b0 - sim b0 but not rejection rates

Sim2b: Test the effect of temporal autocorrelation on estimates of time trends **interaction effects**.

a. 30 time points

b. g = h = 0

c. lambda = gamma = 0 (like 2a)

d. b0 = 0; b1 = 1/30

f. To make things hard, use r2 = 0 and r2 = 0.6 (i.e., temporally autocorrelated U)

g. Plot est b1 - sim b1 but not rejection rates

Sim3: Test the effects of a nugget on estimates of climate variable effects (gamma)

a. Simulate cases g = 0 and g > 0 (maybe g = 0.1 like you used before)

b. r1 = r2 = 0.6, b0 = b1 = 0

c. lambda = 0??

c. Simulate two cases h = 0 and h > 1

d. Plot est g - sim g but not rejection rates

Total time estimates (not accounting the burn-in period):

* If we simulate 500 datasets for each case, it will take a total of 31 days on 4 cores:
  + Sim 1: 498 hrs = 20.75 days
    - 1a: 310 hrs
    - 1b: 38 hrs
    - 1c: 150 hrs
  + Sim 2: 113 hours = 4.7 days
    - 2a: 75 hrs
    - 2b: 38 hours
  + Sim 3: 150 hours = 6.25 days
* If we reduce this to 200 per case, it might take roughly 31\*0.4 = 12.4 days and calculations of rejection rates would be questionable.