Part 2: Modelling a building using CTSM-R

2a: 2-state model of a single room

We will estimate the following model where the impact of the measured radiation G_v is scaled relative the sun's angle through the window. Since we do not have access to this information, a non-parametric fit will be done using B-splines.

$$
\begin{aligned}
dT_i &= \frac{1}{C_i}\left(\frac{1}{R_{ia}}(T_a - T_i) + \frac{1}{R_{im}}(T_m - T_i) + \Phi + \left(\sum_{k=1}^{N} a_k bs_k(t)\right) G_v\right) dt + \sigma_1 dw_1 \\
dT_m &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_i - T_m)\right) dt + \sigma_2 dw_2. \\
yT_i &= T_i + e_1,
\end{aligned}
$$

```
#install.packages("ctsmr", repo = "http://ctsm.info/repo/dev")
#install.packages("pkgbuild")

# For git pushing
## git push https://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@github.com/davidripsen/3-Assignment.git

library(ctsmr)
library(splines)
source("CompEx3_E18/sdeTiTm.R")
# Load data
if (Sys.info()[7] == "davidipsen")
  {path <- "~/Documents/DTU/3. Semester (MSc)/Advanced Time Series/Assignments/3-Assignment/CompEx3_E18/
} else {path <- "CompEx3_E18/"}
load(paste0(path,"Exercise3.RData"))
#AllDat

########## Initial model ############
  fit1 <- sdeTiTm(AllDat,AllDat$yTi1,AllDat$Ph1) # Original model

  summary(fit1,extended=TRUE)
  fit1$loglik

  Hour <- as.numeric(strftime(AllDat$date, format="%H"))

  Pred <- predict(fit1)
  #plot(Pred[[1]]$state$pred$Ti - AllDat$yTi1 ~ Hour)

  # Fit only splines for radiation hours
  #plot(AllDat$Gv ~ Hour) #


  idx <- (Hour>8 & Hour < 23) # It is impossible to fit a window area for the hours without any sun, so
  bs = bs(Hour[idx],df=5,intercept=TRUE) # Dvs. 4 knots / 5 basis splines

  # What does the splines look like?
  #plot(bs[14:27,1],type='l')
  #lines(bs[ 14:27,2])
  #lines(bs[ 14:27,3])
  #lines(bs[ 14:27,4])
```

```r
  #lines(bs[ 14:27,5])

  bs1 <- bs2 <- bs3 <- bs4 <- bs5 <- bs6 <- numeric(dim(AllDat)[1])

  bs1[idx] = bs[,1]
  bs2[idx] = bs[,2]
  bs3[idx] = bs[,3]
  bs4[idx] = bs[,4]
  bs5[idx] = bs[,5]

  AllDat$bs1 = bs1
  AllDat$bs2 = bs2
  AllDat$bs3 = bs3
  AllDat$bs4 = bs4
  AllDat$bs5 = bs5


### IMPLEMENT THE MENTIONED MODEL ###
source(paste0(path,"sdeTiTmAv.R"))
fit2 <- sdeTiTmAv(AllDat,AllDat$yTi1,AllDat$Ph1)
```

Let's compare the two models

```r
sprintf('Model 1: logL = %f', fit1$loglik)
```

```
## [1] "Model 1: logL = -1099.990110"
```

```r
sprintf('Model 2: logL = %f', fit2$loglik)
```

```
## [1] "Model 2: logL = -9.696171"
```

I.e. we see a very large improvement in likelihood (for only 4 extra parameters).

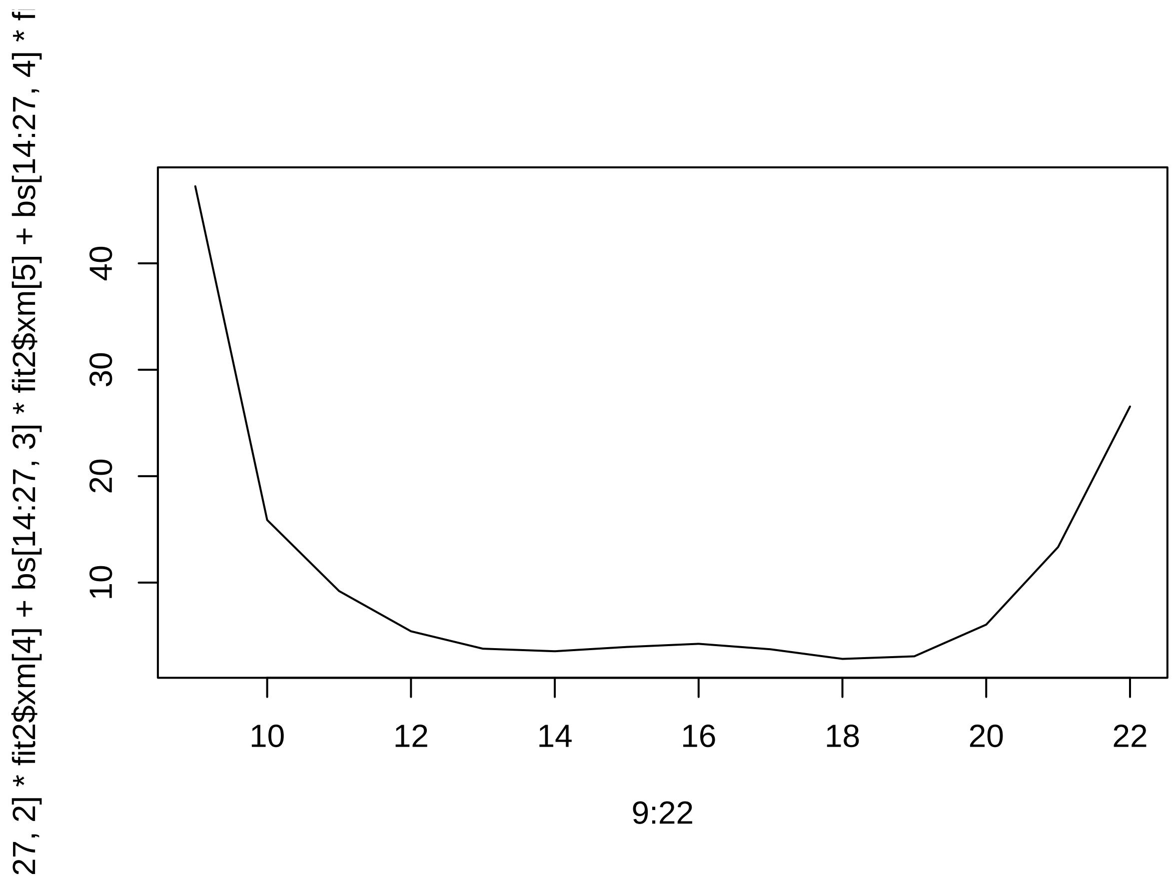
```r
summary(fit2, extended=T)
```

```
## Coefficients:
##        Estimate  Std. Error    t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.3600e+01  1.8990e-01  1.2427e+02  0.0000e+00 -8.3189e-05    0.0004
## Tm0  2.0995e+01  3.1471e-01  6.6712e+01  0.0000e+00 -2.3551e-04    0.0002
## a1   1.5880e+01  1.8993e+00  8.3607e+00  0.0000e+00 -4.0460e-05    0.0000
## a2  -2.2628e+00  1.4018e+00 -1.6142e+00  1.0658e-01  1.1018e-05    0.0000
## a3   1.2802e+01  2.6667e+00  4.8008e+00  1.6578e-06 -3.3692e-05    0.0000
## a4  -6.8605e+00  3.4944e+00 -1.9633e+00  4.9705e-02  3.3437e-06    0.0000
## a5   4.7242e+01  1.8673e+01  2.5300e+00  1.1454e-02  1.9608e-07    0.0000
## Ci   8.5965e+00  2.3986e-01  3.5840e+01  0.0000e+00 -1.8386e-04    0.0000
## Cm   8.8387e+04  1.3726e+04  6.4393e+00  1.3934e-10 -3.6652e-04    0.0066
## e11 -3.0435e+01  3.8155e+00 -7.9767e+00  2.2204e-15  4.0628e-04    0.0004
## p11 -1.6167e+00  2.5923e-02 -6.2366e+01  0.0000e+00  1.4017e-04    0.0000
## p22 -9.2324e-01  6.2581e-02 -1.4753e+01  0.0000e+00 -4.9533e-05    0.0000
## Ria  1.4211e+01  6.8962e+00  2.0607e+00  3.9420e-02 -2.0718e-05    0.0000
## Rim  4.6121e-01  1.7869e-02  2.5810e+01  0.0000e+00 -9.7439e-05    0.0000
##
## Correlation of coefficients:
##     Ti0   Tm0   a1    a2    a3    a4    a5    Ci    Cm    e11   p11   p22
## Tm0 -0.05
## a1   0.03  0.16
```

```
## a2   0.03 -0.08 -0.39
## a3   0.02  0.12  0.59 -0.77
## a4  -0.01 -0.10 -0.34  0.76 -0.84
## a5   0.06  0.09  0.37 -0.34  0.62 -0.59
## Ci   0.01  0.16  0.51 -0.04  0.23 -0.08  0.17
## Cm  -0.06 -0.03 -0.39  0.05 -0.37  0.23 -0.53 -0.26
## e11  0.05  0.06  0.38 -0.12  0.38 -0.28  0.55  0.24 -0.98
## p11  0.04  0.13  0.43 -0.04  0.19 -0.11  0.17  0.22 -0.28  0.27
## p22 -0.05 -0.07 -0.47 -0.05 -0.22  0.07 -0.20 -0.19  0.37 -0.34 -0.57
## Ria  0.07 -0.10  0.25  0.28  0.13  0.15  0.19  0.26 -0.48  0.31  0.17 -0.32
## Rim  0.00 -0.09 -0.21 -0.11 -0.14 -0.06 -0.16 -0.26  0.12 -0.10  0.17  0.14
##     Ria
## Tm0
## a1
## a2
## a3
## a4
## a5
## Ci
## Cm
## e11
## p11
## p22
## Ria
## Rim -0.21
```

We see that all parameters are significant, except for a2. For completeness, a2 is kept in the model. Let's visualize the spline-fit

```
plot(9:22, bs[14:27,1]*fit2$xm[3]+bs[14:27,2]*fit2$xm[4]+bs[14:27,3]*fit2$xm[5]+bs[14:27,4]*fit2$xm[6]+
```



Remember, that the above fit is not the actual input radiation to the room, but it is a *weighting* of the actual radiation. Apparently, the model would like to have more emphasis on the radiation when the sun is low.

2b: Improving the single-room model

In this part, I will try to expand the model. 1. Note how the effect of the temperature in room 3 and 4 affects only room 1 through the medium of room 2. In other words: T_1 is *conditionally independent* of T_3 and T_4 when T_2 is known.

2. Initially, I added the temperature of the neighboring room (T_2) as an additional state and then had it impact dT_1. However, I realized that the temperature of room 2 only affects room 1 through *the wall* as a medium, i.e. the thermal mass of room 1. And since we have direct measurements of $T2$ and only focus on modelling the single-room ($T1$), I add it as input to the model with the term:

$$\frac{1}{R_{2,1}}(T_2 - T_1)$$

thereby letting the change in thermal mass of room 1 (dT_m) being proportional to the difference in temperature between the two rooms.

The model is then

$$
\begin{aligned}
dT_1 &= \frac{1}{C_1}\left(\frac{1}{R_{ia}}\left(T_a - T_1\right) + \frac{1}{R_{im}}\left(T_m - T_1\right) + \Phi + \left(\sum_{k=1}^{N} a_k bs_k(t)\right) G_v\right) dt + \sigma_1 dw_1 \\
dT_m &= \frac{1}{C_m}\left(\frac{1}{R_{im}}\left(T_1 - T_m\right) + \frac{1}{R_{2,1}}(T_2 - T_1)\right) dt + \sigma_2 dw_2. \\
yT_1 &= T_1 + e_1,
\end{aligned}
$$

Let's estimate it

```
#source(paste0(path,"2b-sdeT1T2TmAv.R"))
#fit3 <- sdeT1T2TmAv(AllDat,AllDat$yTi1,AllDat$Ph1)
```

Let's have a look at the model.

```
#sprintf('Model 3: logL = %f', fit3$loglik)
#sprintf('Likelihood ratio test: p-value = %f', 1-pchisq(abs(2*(fit2$loglik - fit3$loglik)),1))
#summary(fit3, extended=T)
```

Again, a large improvement in likelihood is seen (at the cost of 1 extra parameter). Since the two models are nested, at likehood ratio test is performed and shows a very strong signficant difference. Again, all parameters are significant, except for the observation error which might in fact have 0 as the true value, however we do see some strong correlations in the coefficients (!).

Now expanding further on the model, I see that the radiator is placed close to the window in room 1, i.e. some of the heating might go straight out of the window. Additionally, I note that the North Heating Circuit and South Heating Circuit are piped together before entering the building, so the actual heating in room 1 is a mix of the two (however, the two are higly correlated with $\hat{\rho} = 0.91$). I.e. denote $= \Phi = (\Phi_1 + \Phi_2)/2$ and change then the Φ term in dT_1 to the term $(1-c)\Phi - c\Phi\frac{T_1 - T_a}{T_1}$, where $c \in [0, 1]$. This makes a proportion c of the heating load escape directly through the window relative to the temperature ratio between inside and outside[1]. Note that using this formulation, the hypothesis H0: $c = 0$ is of interest (meaning no loss of heating) and when the outdoor and indoor temperature is the same, the second term cancels (meaning no loss of heating).

The model then becomes

$$
\begin{aligned}
dT_1 &= \frac{1}{C_1}\left(\frac{1}{R_{ia}}\left(T_a - T_1\right) + \frac{1}{R_{im}}\left(T_m - T_1\right) + ((1-c)\Phi - c\Phi\frac{T_1 - T_a}{T_1}) + \left(\sum_{k=1}^{N} a_k bs_k(t)\right) G_v\right) dt + \sigma_1 dw_1 \\
dT_m &= \frac{1}{C_m}\left(\frac{1}{R_{im}}\left(T_1 - T_m\right) + \frac{1}{R_{2,1}}(T_2 - T_1)\right) dt + \sigma_2 dw_2. \\
yT_1 &= T_1 + e_1,
\end{aligned}
$$

Where $\Phi = (\Phi_1 + \Phi_2)/2$.

```
#plot(AllDat$date, AllDat$Ta, type='l')
#plot(AllDat$date, (AllDat$Ph1 + AllDat$Ph2)/2, type='l')
#plot(AllDat$Ph1, AllDat$Ph2); cor(AllDat$Ph1, AllDat$Ph2)
#source(paste0(path,"2b-2-sdeT1T2TmAv.R"))
#fit4 <- sdeT1T2TmAv(AllDat,AllDat$yTi1)
```

[1] The meaning of T_a is not explicit from the assignment. From plots I've concluded to assume it to be the outside temperature.