Part 2: Modelling a building using CTSM-R

2a: 2-state model of a single room

In this exercise, I will estimate the following model where the impact of the measured solar 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. The model is

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

Let's estimate the model

```
#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])
```

```r
  #lines(bs[ 14:27,3])
  #lines(bs[ 14:27,4])
  #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, i.e. with and without the spline-fit for the relative radiation impact.

```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 = -9.704192"
```

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

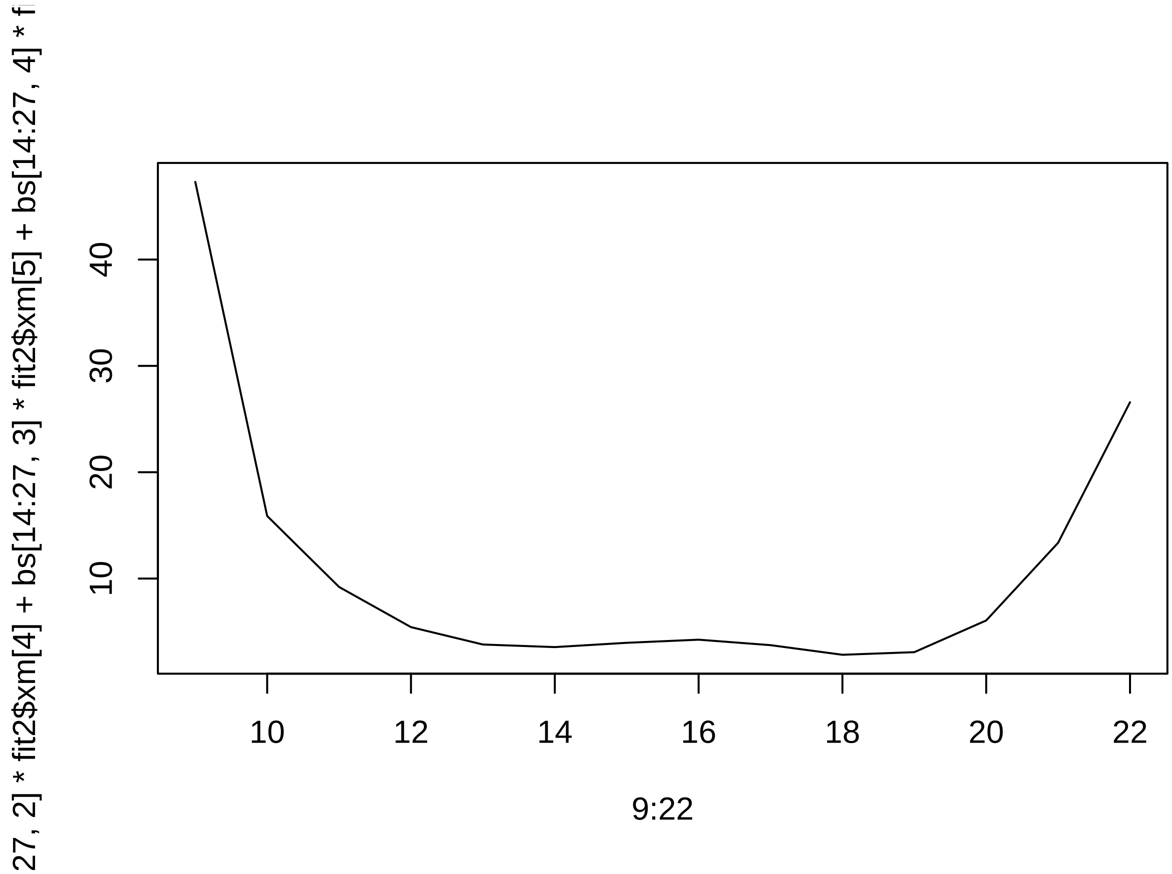
```r
summary(fit2, extended=T)
```

```
## Coefficients:
##         Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0   2.3599e+01  1.8687e-01  1.2629e+02  0.0000e+00 -4.4067e-01     4e-04
## Tm0   2.0988e+01  7.6874e-01  2.7302e+01  0.0000e+00 -1.0769e-01     2e-04
## a1    1.5884e+01  1.6699e+00  9.5122e+00  0.0000e+00  1.5806e-02     0e+00
## a2   -2.2530e+00  1.5086e+00 -1.4935e+00  1.3542e-01 -2.9110e-03     0e+00
## a3    1.2806e+01  2.7271e+00  4.6960e+00  2.7724e-06 -9.5178e-02     0e+00
## a4   -6.8625e+00  3.5952e+00 -1.9088e+00  5.6380e-02  4.7843e-02     0e+00
## a5    4.7306e+01  1.8826e+01  2.5128e+00  1.2027e-02 -9.7711e-03     0e+00
## Ci    8.5968e+00  2.0324e-01  4.2300e+01  0.0000e+00 -6.4099e-02     0e+00
## Cm    4.9063e+04  1.2524e+05  3.9176e-01  6.9526e-01  3.1704e-02     2e-04
## e11  -1.5331e+01  3.0499e-01 -5.0267e+01  0.0000e+00 -2.8737e-02     0e+00
## p11  -1.6166e+00  2.1809e-02 -7.4127e+01  0.0000e+00 -6.0683e-01     0e+00
## p22  -9.2340e-01  5.6930e-02 -1.6220e+01  0.0000e+00 -6.6264e-02     0e+00
## Ria   1.4344e+01  6.0685e+00  2.3637e+00  1.8154e-02  4.5962e-02     0e+00
## Rim   4.6107e-01  1.7451e-02  2.6420e+01  0.0000e+00 -3.0208e-01     0e+00
##
## Correlation of coefficients:
##     Ti0   Tm0   a1    a2    a3    a4    a5    Ci    Cm    e11   p11   p22
```

```
## Tm0  0.12
## a1   0.02 -0.08
## a2  -0.02 -0.04 -0.53
## a3   0.03 -0.03  0.57 -0.84
## a4  -0.06 -0.03 -0.36  0.79 -0.87
## a5   0.04 -0.04  0.31 -0.40  0.63 -0.62
## Ci   0.00  0.01  0.32 -0.05  0.11 -0.05  0.09
## Cm   0.00  0.02  0.00  0.00  0.00  0.00  0.00  0.01
## e11  0.02  0.07 -0.30  0.03 -0.08 -0.16 -0.04  0.03 -0.30
## p11  0.04 -0.09  0.21 -0.07  0.04 -0.04  0.01  0.01  0.00  0.38
## p22  0.01  0.11 -0.27 -0.05 -0.07  0.00 -0.10  0.00 -0.02 -0.09 -0.51
## Ria  0.02 -0.27  0.15  0.22  0.09  0.15  0.17  0.12  0.00 -0.36  0.04 -0.26
## Rim  0.04 -0.12 -0.16 -0.12 -0.09 -0.09 -0.14 -0.27  0.00  0.27  0.31  0.08
##     Ria
## Tm0
## a1
## a2
## a3
## a4
## a5
## Ci
## Cm
## e11
## p11
## p22
## Ria
## Rim -0.11
```

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* heat input from radiation, but it is a *weighting* of the measured 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. First, 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.

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 instead 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 (model 3) is then

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

Let's estimate model 3

```
set.seed(101)
source(paste0(path,"2b-sdeT1T2TmAv.R"))
fit3 <- sdeT1T2TmAv_2b(AllDat,AllDat$yTi1,AllDat$Ph1)
```

Let's have a look at the model.

```
sprintf('Model 3: logL = %f', fit3$loglik)
```

```
## [1] "Model 3: logL = -9.698581"
```

```
sprintf('Likelihood ratio test: p-value = %f', 1-pchisq(abs(2*(fit2$loglik - fit3$loglik)),1))
```

```
## [1] "Likelihood ratio test: p-value = 0.915634"
```

```
summary(fit3, extended=T)
```

```
## Coefficients:
##          Estimate  Std. Error     t value    Pr(>|t|)      dF/dPar dPen/dPar
## T10   2.3600e+01  1.9686e-01  1.1988e+02  0.0000e+00  3.2866e-02    0.0004
## Tm0   2.0996e+01  7.5174e-01  2.7929e+01  0.0000e+00  4.0611e-03    0.0002
## a1    1.5881e+01  1.5832e+00  1.0030e+01  0.0000e+00 -5.8404e-03    0.0000
## a2   -2.2628e+00  1.4575e+00 -1.5525e+00  1.2064e-01 -3.2334e-03    0.0000
## a3    1.2803e+01  2.5945e+00  4.9347e+00  8.4627e-07  1.8938e-02    0.0000
## a4   -6.8610e+00  3.4517e+00 -1.9877e+00  4.6928e-02 -4.0893e-03    0.0000
## a5    4.7249e+01  1.8330e+01  2.5777e+00  9.9907e-03  2.3702e-04    0.0000
## Ci    8.5967e+00  2.1182e-01  4.0585e+01  0.0000e+00  6.8290e-02    0.0000
## Cm    8.3689e+04  1.4914e+05  5.6116e-01  5.7473e-01  4.3731e-03    0.0031
## e11  -1.5093e+01  4.0597e-01 -3.7177e+01  0.0000e+00 -4.9749e-02    0.0000
## p11  -1.6167e+00  2.1253e-02 -7.6069e+01  0.0000e+00  2.4712e-02    0.0000
## p22  -9.2330e-01  5.4037e-02 -1.7086e+01  0.0000e+00  3.0584e-02    0.0000
## R_21  1.3443e+01  9.5225e+00  1.4117e+00  1.5814e-01  5.4287e-04    0.0000
## Ria   1.4206e+01  6.3891e+00  2.2235e+00  2.6253e-02 -3.5968e-03    0.0000
## Rim   4.6120e-01  1.7871e-02  2.5807e+01  0.0000e+00  2.6249e-02    0.0000
##
## Correlation of coefficients:
##       T10   Tm0   a1    a2    a3    a4    a5    Ci    Cm    e11   p11   p22
## Tm0   0.11
## a1    0.00 -0.06
## a2    0.02 -0.08 -0.52
## a3    0.00  0.00  0.57 -0.81
## a4    0.00 -0.05 -0.34  0.75 -0.85
## a5   -0.01 -0.02  0.30 -0.37  0.62 -0.62
## Ci   -0.01  0.08  0.38 -0.08  0.14 -0.05  0.07
## Cm    0.00 -0.01  0.00  0.00  0.00  0.00  0.00  0.00
## e11  -0.01  0.10 -0.25 -0.03 -0.05 -0.03 -0.03 -0.04 -0.76
## p11  -0.02 -0.11  0.15 -0.06  0.04 -0.05  0.05 -0.06  0.00  0.13
## p22   0.02  0.12 -0.19 -0.09 -0.03 -0.02 -0.05  0.11  0.00  0.08 -0.48
```

```
## R_21  0.00  0.01  0.00  0.00  0.00  0.00  0.00  0.00 -0.99  0.75  0.00  0.00
## Ria   0.02 -0.32  0.15  0.27  0.09  0.16  0.15  0.13  0.01 -0.27  0.05 -0.25
## Rim  -0.01 -0.09 -0.16 -0.16 -0.08 -0.11 -0.11 -0.32 -0.01  0.06  0.27 -0.01
##      R_21  Ria
## Tm0
## a1
## a2
## a3
## a4
## a5
## Ci
## Cm
## e11
## p11
## p22
## R_21
## Ria  -0.02
## Rim   0.00 -0.20
```

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 $\hat{\rho} = 0.91$). For model extenstions; Use $\Phi = (\Phi_1 + \Phi_2)/2$ and additionally change the Φ-term in dT_1 to the term $(1-c)\Phi - c\Phi(T_1 - T_a)$, where $c \in [0,1]$. This makes a proportion c of the heating load escape directly through the window relative to the temperature difference[1] between inside and outside[2]. 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

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

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

```r
#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$Ph2)
source(paste0(path,"2b-2-sdeT1T2TmAv.R"))
#fit4 <- sdeT1T2TmAvfit4(AllDat,AllDat$yTi1)
```

However, I can't get his model to converge properly, even with reduced convergence criteria and increased function evaluation allowance.

[1] I would have liked it to be a ratio instead of difference, but that makes it much harder for CTSM to solve.

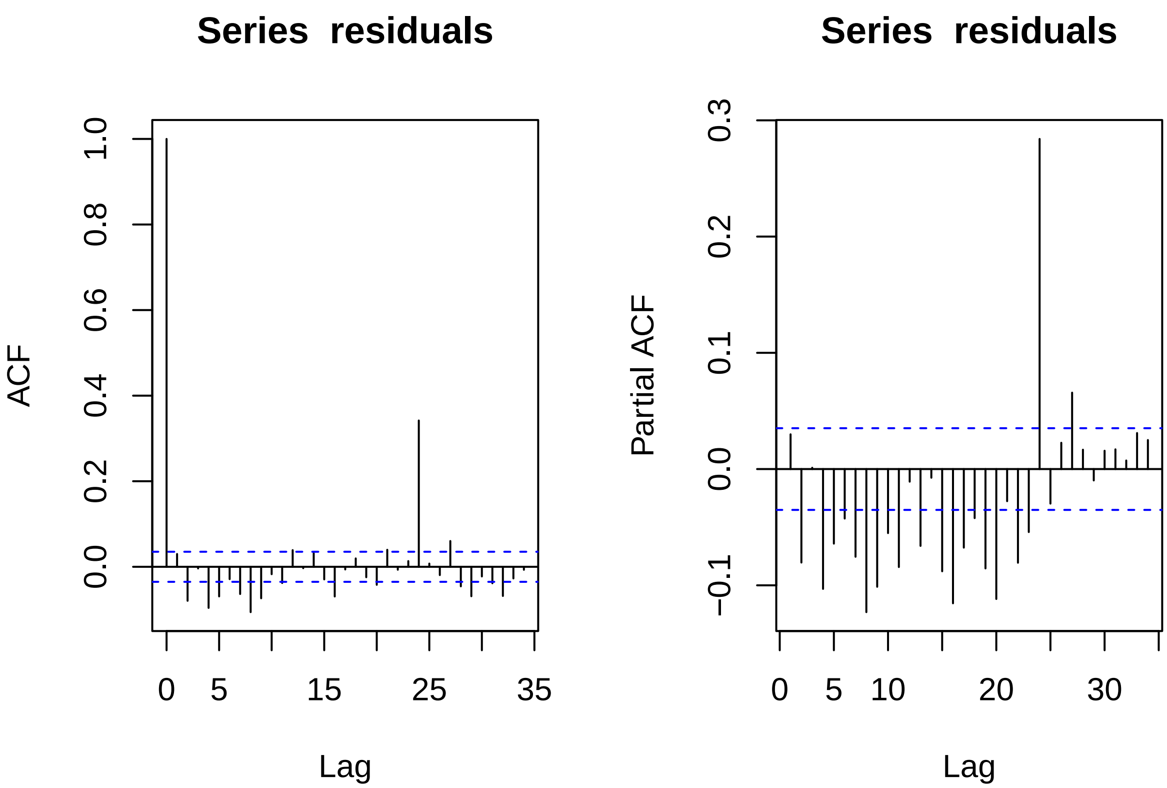
[2] The meaning of T_a is not explicit from the assignment. From plots I've concluded to assume it to be the outside temperature.

```
## Evaluate model
#fit4$loglik
#summary(fit4, extended=T)
#sprintf('Likelihood ratio test from previous model: p-value = %f', 1-pchisq(-2*(fit3$loglik - fit4$log
```

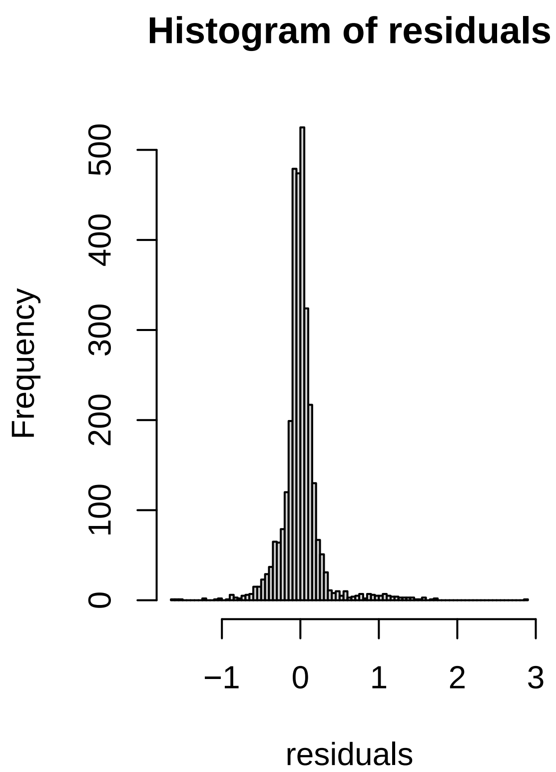
For this reason, I'm satisfied with model 3. Of course there are yet many possible extensions to try out, but for this single-room model, model 3 will suffice.

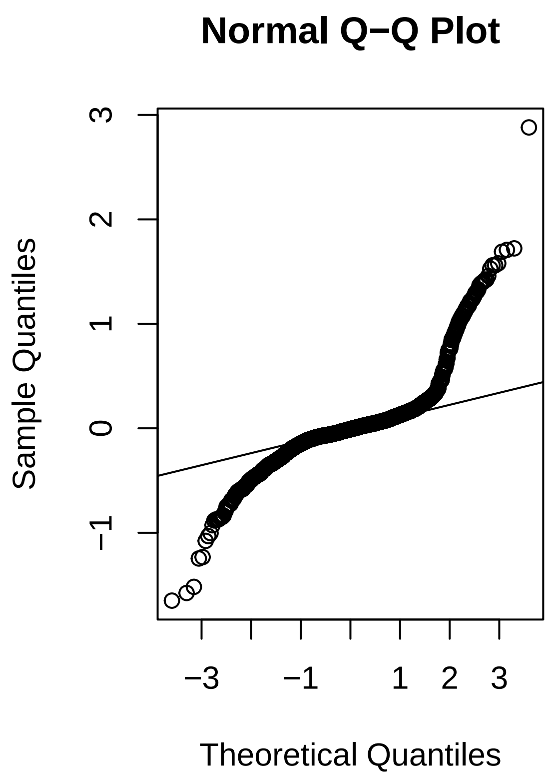
I will end this section with model validation of model 3.

```
# To assess the model fit, calculate the one-step predictions and residuals.
# Do residual analysis for model validation.

D = AllDat; D$Ph <- AllDat$Ph1; D$T2 <- AllDat$yTi2; D$yT1 <- AllDat$yTi1;
preds <- predict(fit3, newdata=D)
yTi1Hat <- preds$output$pred$yT1
residuals <- AllDat$yTi1 - yTi1Hat

par(mfrow=c(1,2))
acf(residuals)
pacf(residuals)
```

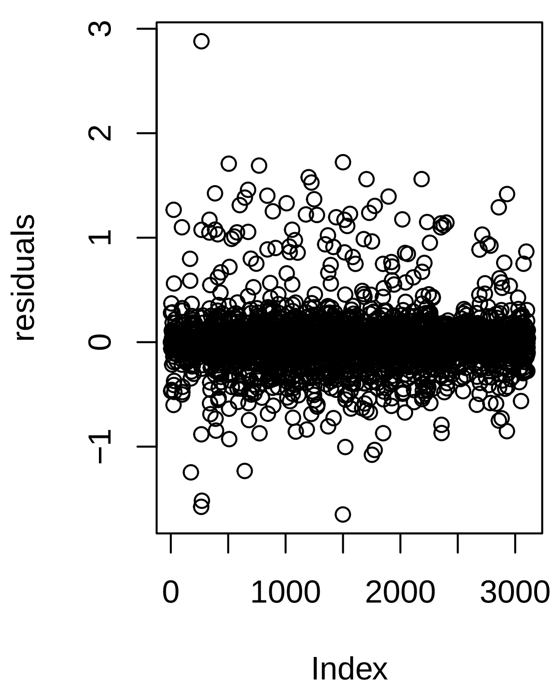


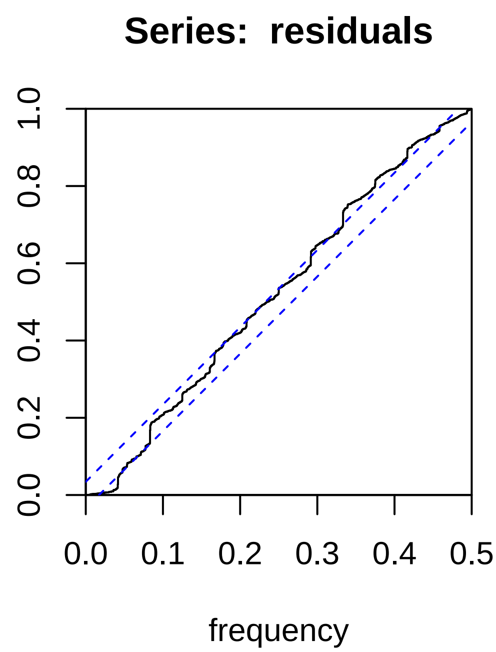

```
hist(residuals, breaks=100)
qqnorm(residuals)
qqline(residuals)
```






```
plot(residuals)
cpgram(residuals)
```





There is clearly further work to be done - the residuals are far from white noise. Inspecting the autocorrelation and partial autocorrelation, it would clearly be beneficial to adopt this autoregression further, particularly in lag 24. There is a strong over-dispersion in the residuals.

3 Making a multi-room model

In this exercise, I will fit a continuous-discrete state space model to the 4 rooms. I will extend model 3 to the multi-room case with the following considerations and simplifications (for computational feasibility and identifiability):

1. The thermal diffusion from one room to another is only present for neighboring rooms - and enters *through* the thermal mass.
2. Only room 1 and 4 are neighbours to the outside (T_a)
3. Use $\Phi = (\Phi_1 + \Phi_2)/2$
4. Use R_{iw} for the resistance of interior walls as a generalization of $R_{2,1}$.
5. Assume shared parameters where relevant, e.g. assume similar walls R_{iw} (simplification)
6. Assume equal variances across rooms $\{\sigma_1^2, \sigma_2^2\}$ (simplification)
7. Use A_w instead of splines (simplification)

The model is as follows

$$
\begin{aligned}
dT_1 &= \frac{1}{C_i}\left(\frac{1}{R_{ia}}(T_a - T_1) + \frac{1}{R_{im}}(T_{1m} - T_1) + \Phi + \left(\sum_{k=1}^{N} a_k bs_k(t)\right) G_v\right) dt + \sigma_1 dw_{1,1} \\
dT_{1m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_1 - T_{1m}) + \frac{1}{R_{iw}}(T_2 - T_1)\right) dt + \sigma_2 dw_{1,2}. \\
dT_2 &= \frac{1}{C_i}\left(\frac{1}{R_{im}}(T_{2m} - T_2) + \Phi\right) dt + \sigma_1 dw_{2,1} \\
dT_{2m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_2 - T_{2m}) + \frac{1}{R_{iw}}(T_1 - T_2) + \frac{1}{R_{iw}}(T_3 - T_2)\right) dt + \sigma_2 dw_{2,2}. \\
dT_3 &= \frac{1}{C_i}\left(\frac{1}{R_{im}}(T_{3m} - T_3) + \Phi\right) dt + \sigma_1 dw_{3,1} \\
dT_{3m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_3 - T_{3m}) + \frac{1}{R_{iw}}(T_2 - T_3) + \frac{1}{R_{iw}}(T_4 - T_3)\right) dt + \sigma_2 dw_{3,2}. \\
dT_4 &= \frac{1}{C_i}\left(\frac{1}{R_{ia}}(T_a - T_4) + \frac{1}{R_{im}}(T_{4m} - T_4) + \Phi + \left(\sum_{k=1}^{N} a_k bs_k(t)\right) G_v\right) dt + \sigma_1 dw_{4,1} \\
dT_{4m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_4 - T_{4m}) + \frac{1}{R_{iw}}(T_3 - T_4)\right) dt + \sigma_2 dw_{4,2}. \\
\forall_{i \in \{1,2,3,4\}} yT_i &= T_i + e_i
\end{aligned}
$$

Let's fit the model

```r
D = AllDat[1:2000,]# Subset the data for computional feasibility. 3111 for all.
source(paste0(path,"3-multi-room.R"))
fitmultim = sde_multi_room(D)
```

The model diverges when trained on the whole dataset, but returns a proper fit when trained on 2/3 of the data. Let's evaluate the fit:

```r
fitmultim$loglik
```

```
## [1] 33.8177
```

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

```
## Coefficients:
##          Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## T10    1.5026e+01  6.4721e-01  2.3216e+01  0.0000e+00 -3.8409e+02     1e-04
## T1m0   1.4989e+01  6.4655e-01  2.3182e+01  0.0000e+00  3.1702e-01     1e-04
## T20    1.5021e+01  6.4712e-01  2.3212e+01  0.0000e+00 -3.4665e+02     1e-04
## T2m0   1.4988e+01  6.4655e-01  2.3182e+01  0.0000e+00  3.0118e-01     1e-04
## T30    1.5022e+01  6.4714e-01  2.3213e+01  0.0000e+00 -3.6768e+02     1e-04
## T3m0   1.4988e+01  6.4655e-01  2.3182e+01  0.0000e+00  3.2069e-01     1e-04
## T40    1.5021e+01  6.4712e-01  2.3212e+01  0.0000e+00 -3.4001e+02     1e-04
## T4m0   1.4988e+01  6.4655e-01  2.3182e+01  0.0000e+00  2.9620e-01     1e-04
## Aw     5.5879e+00  2.6127e-01  2.1387e+01  0.0000e+00  1.7650e+00     1e-04
## Ci     2.7024e+01  4.4580e+00  6.0618e+00  1.4124e-09 -6.9022e+03     0e+00
## Cm     1.0020e+03  6.8433e+01  1.4642e+01  0.0000e+00  5.5301e+01     0e+00
## e11   -9.7489e+00  1.1349e+00 -8.5902e+00  0.0000e+00  2.6939e+02     0e+00
## e21   -9.6753e+00  1.1340e+00 -8.5319e+00  0.0000e+00  4.7490e+01     0e+00
## e31   -9.7101e+00  1.1357e+00 -8.5496e+00  0.0000e+00  4.6547e+01     0e+00
## e41   -9.8000e+00  1.1405e+00 -8.5924e+00  0.0000e+00  2.8754e+01     0e+00
## p11   -3.4090e+00  2.2857e-01 -1.4915e+01  0.0000e+00  1.4164e+03     0e+00
## p22    4.7350e+00  6.3773e-02  7.4248e+01  0.0000e+00  9.0049e+02     2e-04
## Ria    1.8743e+01  1.2970e+00  1.4452e+01  0.0000e+00  1.5643e+04     0e+00
## Rim    1.7277e+01  1.1643e+00  1.4838e+01  0.0000e+00  4.3039e+03     0e+00
## Riw    2.0006e+01  1.3799e+00  1.4499e+01  0.0000e+00 -8.3891e+01     0e+00
##
## Correlation of coefficients:
##      T10   T1m0  T20   T2m0  T30   T3m0  T40   T4m0  Aw    Ci    Cm    e11
## T1m0  0.00
## T20   0.00  0.00
## T2m0  0.00  0.00  0.00
## T30   0.00  0.00  0.00  0.00
## T3m0  0.00  0.00  0.00  0.00  0.00
## T40   0.00  0.00  0.00  0.00  0.00  0.00
## T4m0  0.00  0.00  0.00  0.00  0.00  0.00  0.00
## Aw    0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
## Ci    0.01  0.00  0.01  0.00  0.01  0.00  0.01  0.00 -0.06
## Cm    0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
## e11  -0.01  0.00 -0.01  0.00 -0.01  0.00 -0.01  0.00  0.04 -0.54  0.00
## e21  -0.01  0.00 -0.01  0.00 -0.01  0.00 -0.01  0.00  0.04 -0.54  0.00  0.35
## e31  -0.01  0.00 -0.01  0.00 -0.01  0.00 -0.01  0.00  0.04 -0.54  0.00  0.35
## e41  -0.01  0.00 -0.01  0.00 -0.01  0.00 -0.01  0.00  0.04 -0.54  0.00  0.35
## p11  -0.01  0.00 -0.01  0.00 -0.01  0.00 -0.01  0.00  0.08 -0.99  0.00  0.61
## p22   0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00 -0.04  0.97  0.00 -0.53
## Ria   0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.01 -0.09  0.00  0.05
## Rim   0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00 -0.07  0.98  0.00 -0.59
## Riw   0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
##      e21   e31   e41   p11   p22   Ria   Rim
## T1m0
## T20
## T2m0
## T30
## T3m0
## T40
## T4m0
## Aw
## Ci
```

```
## Cm
## e11
## e21
## e31   0.35
## e41   0.35  0.35
## p11   0.61  0.61  0.61
## p22  -0.54 -0.54 -0.54 -0.96
## Ria   0.05  0.05  0.06  0.09 -0.04
## Rim  -0.58 -0.58 -0.59 -0.98  0.93 -0.05
## Riw   0.00  0.00  0.00  0.00  0.00  0.00  0.00
```

Inspecting the fit is evident that

1. All parameters are significant
2. None of the parameters are too strongly correlated
3. Some partial derivatives are still fairly large.

As a final model I will try to lift two simplifications:

1. Let each room temperature (T_i) and room thermal mass (T_{im}) have unique variances $\{\sigma^2_{i,1}, \sigma^2_{i,2}\}$
2. Expand the A_w to the spline-fit which proved very beneficial in the first model. Still assume equal weighting of splines for the northen and southern room.

In other words, the final model is:

$$
\begin{aligned}
dT_1 &= \frac{1}{C_i}\left(\frac{1}{R_{ia}}(T_a - T_1) + \frac{1}{R_{im}}(T_{1m} - T_1) + \Phi + A_w G_v\right) dt + \sigma_{1,1} dw_{1,1} \\
dT_{1m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_1 - T_{1m}) + \frac{1}{R_{iw}}(T_2 - T_1)\right) dt + \sigma_{1,2} dw_{1,2}. \\
dT_2 &= \frac{1}{C_i}\left(\frac{1}{R_{im}}(T_{2m} - T_2) + \Phi\right) dt + \sigma_{2,1} dw_{2,1} \\
dT_{2m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_2 - T_{2m}) + \frac{1}{R_{iw}}(T_1 - T_2) + \frac{1}{R_{iw}}(T_3 - T_2)\right) dt + \sigma_{2,2} dw_{2,2}. \\
dT_3 &= \frac{1}{C_i}\left(\frac{1}{R_{im}}(T_{3m} - T_3) + \Phi\right) dt + \sigma_{3,1} dw_{3,1} \\
dT_{3m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_3 - T_{3m}) + \frac{1}{R_{iw}}(T_2 - T_3) + \frac{1}{R_{iw}}(T_4 - T_3)\right) dt + \sigma_{3,2} dw_{3,2}. \\
dT_4 &= \frac{1}{C_i}\left(\frac{1}{R_{ia}}(T_a - T_4) + \frac{1}{R_{im}}(T_{4m} - T_4) + \Phi + A_w G_v\right) dt + \sigma_{4,1} dw_{4,1} \\
dT_{4m} &= \frac{1}{C_m}\left(\frac{1}{R_{im}}(T_4 - T_{4m}) + \frac{1}{R_{iw}}(T_3 - T_4)\right) dt + \sigma_{4,2} dw_{4,2}. \\
\forall_{i\in\{1,2,3,4\}} yT_i &= T_i + e_i
\end{aligned}
$$

Let's estimate it

```
D = AllDat[1:500,]# Subset the data for computional feasibility. 3111 for all.
source(paste0(path,"3b-multi-room.R"))
fitmulti2 = sde2_multi_room(D)

# Also fit the previous model on the same subset of data to allow for AIC-comparison of the models
fitmulti1 = sde_multi_room(D)
```

Let's evaluate the fit

```
fitmulti2$loglik
```

```
## [1] 1258.297
```

```
sprintf("AIC for comparison of non-nested models:")
```

```
## [1] "AIC for comparison of non-nested models:"
```

```
sprintf("AIC of final model: %f", -2*fitmulti2$loglik + 2*length(fitmulti2$xm))
```

```
## [1] "AIC of final model: -2456.594737"
```

```
sprintf("AIC of previous model: %f", -2*fitmultim$loglik + 2*length(fitmulti1$xm))
```

```
## [1] "AIC of previous model: -27.635394"
```

```
summary(fitmulti2, extended=T)
```

```
## Coefficients:
```

```
##          Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## T10   2.3600e+01  1.4943e-01  1.5794e+02  0.0000e+00  9.3213e-05    0.0004
## T1m0  2.5704e+01  1.1966e+00  2.1481e+01  0.0000e+00  2.5545e-05    0.0005
## T20   2.2300e+01  8.2600e-02  2.6998e+02  0.0000e+00 -3.5473e-04    0.0003
## T2m0  2.0760e+01  6.0168e-01  3.4504e+01  0.0000e+00  3.6248e-04    0.0002
## T30   2.2698e+01  7.6120e-02  2.9819e+02  0.0000e+00  2.8411e-04    0.0003
## T3m0  2.1986e+01  5.2753e-01  4.1678e+01  0.0000e+00 -1.0073e-04    0.0003
## T40   2.2100e+01  7.2344e-02  3.0549e+02  0.0000e+00  3.1575e-04    0.0003
## T4m0  2.1855e+01  4.0182e-01  5.4388e+01  0.0000e+00 -1.8100e-04    0.0003
## a1    6.7194e+00  5.7414e+00  1.1703e+00  2.4200e-01 -3.6149e-06    0.0000
## a2    1.9618e+01  6.0733e+00  3.2301e+00  1.2578e-03 -2.2470e-06    0.0000
## a3   -5.3347e+00  1.3684e+01 -3.8986e-01  6.9668e-01 -8.0002e-07    0.0000
## a4    6.4815e+01  2.7413e+01  2.3644e+00  1.8155e-02 -4.5573e-06    0.0000
## a5   -4.9587e+02  2.0538e+01 -2.4144e+01  0.0000e+00 -3.4210e-05    1.4566
## Ci    1.7609e+01  6.4780e-01  2.7183e+01  0.0000e+00  2.7203e-04    0.0000
## Cm    3.6776e+01  4.9994e+00  7.3561e+00  2.7889e-13 -1.2298e-05    0.0000
## e11  -2.0922e+01  2.6397e+01 -7.9261e-01  4.2810e-01  7.5146e-05    0.0001
## e21  -1.9884e+01  5.6451e+00 -3.5223e+00  4.3764e-04  5.3862e-05    0.0001
## e31  -8.4028e+00  2.1708e+00 -3.8708e+00  1.1207e-04  1.0814e-04    0.0000
## e41  -2.3283e+01  2.1797e+01 -1.0682e+00  2.8556e-01  9.0916e-05    0.0001
## p11a -2.0089e+00  6.3809e-02 -3.1483e+01  0.0000e+00 -1.4383e-04    0.0000
## p11b -2.6485e+00  5.0114e-02 -5.2850e+01  0.0000e+00 -1.0616e-04    0.0000
## p11c -2.6849e+00  1.8262e-01 -1.4702e+01  0.0000e+00 -6.5983e-05    0.0000
## p11d -2.5470e+00  4.8264e-02 -5.2773e+01  0.0000e+00 -3.2216e-04    0.0000
## p22a -1.9729e-01  9.2786e-02 -2.1263e+00  3.3606e-02 -1.2669e-06    0.0000
## p22b -7.6345e-01  8.4183e-02 -9.0689e+00  0.0000e+00 -6.6005e-06    0.0000
## p22c -9.7224e-01  9.3287e-02 -1.0422e+01  0.0000e+00 -1.5807e-05    0.0000
## p22d -1.6128e+00  1.3176e-01 -1.2240e+01  0.0000e+00 -7.7866e-06    0.0000
## Ria   2.6558e+00  1.8356e-01  1.4468e+01  0.0000e+00 -5.7077e-05    0.0000
## Rim   2.7835e-01  2.0865e-02  1.3340e+01  0.0000e+00 -9.3373e-05    0.0000
## Riw   6.0544e-01  1.3764e-01  4.3986e+00  1.1488e-05  2.9782e-07    0.0000
##
## Correlation of coefficients:
##      T10   T1m0  T20   T2m0  T30   T3m0  T40   T4m0  a1    a2    a3    a4
## T1m0  0.30
## T20  -0.03  0.00
## T2m0 -0.03  0.03  0.35
## T30  -0.02 -0.02  0.02 -0.02
## T3m0  0.00 -0.03  0.02  0.04  0.25
## T40   0.00  0.02 -0.01 -0.02 -0.01 -0.01
## T4m0  0.00  0.07 -0.02  0.05  0.01  0.01  0.12
## a1    0.00 -0.16  0.02 -0.01  0.00  0.03 -0.01 -0.22
## a2   -0.03  0.08  0.01  0.03  0.00  0.00 -0.03  0.08 -0.71
## a3    0.02 -0.06  0.00 -0.01  0.00  0.00  0.00 -0.10  0.65 -0.93
## a4   -0.02  0.04  0.00  0.03 -0.01  0.01  0.01  0.07 -0.50  0.81 -0.92
## a5    0.02  0.07 -0.02 -0.12 -0.04 -0.09 -0.01 -0.01 -0.05 -0.01 -0.01 -0.01
## Ci   -0.01  0.08  0.00  0.09  0.01  0.12 -0.01  0.10  0.08  0.04  0.03 -0.04
## Cm    0.03  0.03 -0.01  0.11  0.03  0.11  0.06  0.21 -0.13  0.05 -0.05  0.04
## e11  -0.02 -0.05 -0.02 -0.10  0.00  0.04 -0.01  0.00 -0.02  0.04 -0.04  0.05
## e21   0.03  0.05 -0.01  0.04  0.02 -0.06  0.02 -0.04  0.04 -0.03  0.06 -0.06
## e31  -0.05 -0.06  0.00 -0.06 -0.05  0.02 -0.05  0.04  0.01  0.02 -0.02  0.04
## e41   0.03  0.06  0.02  0.12  0.00 -0.03  0.02  0.02  0.00 -0.04  0.04 -0.06
## p11a  0.02 -0.07 -0.03 -0.08 -0.01  0.01 -0.01 -0.08  0.16 -0.15  0.15 -0.14
## p11b  0.01  0.03  0.03  0.03 -0.03  0.03 -0.03  0.00  0.04 -0.08  0.08 -0.09
```

```
## p11c  0.04  0.05 -0.01  0.05  0.05 -0.01  0.04 -0.03 -0.01 -0.04  0.03 -0.05
## p11d  0.02  0.06 -0.01 -0.01  0.01 -0.01  0.01  0.01  0.04  0.06 -0.04  0.02
## p22a  0.02  0.15  0.00 -0.08  0.05 -0.03  0.00 -0.07  0.03 -0.06  0.08 -0.10
## p22b  0.04  0.05 -0.01 -0.14  0.02 -0.06  0.01 -0.14  0.08 -0.10  0.10 -0.14
## p22c -0.03 -0.02  0.02 -0.12 -0.02 -0.05 -0.04 -0.10  0.06 -0.05  0.07 -0.09
## p22d -0.01  0.02  0.02  0.04  0.03  0.04  0.00  0.09 -0.13  0.02 -0.03  0.01
## Ria   0.01 -0.08 -0.04 -0.09 -0.01 -0.08 -0.04 -0.31 -0.06  0.02  0.00  0.02
## Rim   0.02  0.03 -0.02 -0.20  0.02 -0.13  0.02 -0.21  0.10 -0.20  0.18 -0.27
## Riw  -0.01 -0.03  0.00 -0.03 -0.02 -0.08 -0.03 -0.16  0.11  0.00  0.03  0.02
##      a5    Ci    Cm    e11   e21   e31   e41   p11a  p11b  p11c  p11d  p22a
## T1m0
## T20
## T2m0
## T30
## T3m0
## T40
## T4m0
## a1
## a2
## a3
## a4
## a5
## Ci   -0.07
## Cm   -0.02  0.25
## e11   0.07 -0.10 -0.05
## e21  -0.05  0.21  0.10 -0.86
## e31   0.04 -0.12 -0.17  0.16 -0.50
## e41  -0.06  0.13  0.08 -0.99  0.85 -0.24
## p11a  0.00  0.17  0.14 -0.05  0.12 -0.10  0.06
## p11b -0.01  0.12  0.15 -0.01  0.08 -0.08  0.02  0.11
## p11c -0.03  0.13  0.20 -0.15  0.49 -0.95  0.22  0.12  0.11
## p11d -0.05  0.29  0.01 -0.08  0.17 -0.13  0.10  0.16  0.04  0.16
## p22a  0.02  0.27 -0.26 -0.01  0.07 -0.07  0.02 -0.20  0.11  0.08  0.17
## p22b  0.04  0.08 -0.38  0.00  0.04 -0.04  0.00  0.11 -0.18  0.05  0.17  0.37
## p22c  0.03  0.01 -0.43  0.07 -0.19  0.44 -0.10  0.00  0.02 -0.48  0.06  0.28
## p22d  0.03  0.04  0.02  0.01 -0.04  0.03  0.00 -0.08  0.01 -0.05 -0.47  0.00
## Ria   0.04 -0.15 -0.26  0.08 -0.05  0.01 -0.10  0.03  0.03 -0.01 -0.01  0.11
## Rim   0.06  0.08 -0.21  0.00  0.08 -0.13  0.00  0.26  0.22  0.17  0.24  0.46
## Riw  -0.01 -0.12 -0.32  0.02  0.01 -0.03 -0.04 -0.04  0.01  0.03  0.04  0.11
##      p22b  p22c  p22d  Ria   Rim
## T1m0
## T20
## T2m0
## T30
## T3m0
## T40
## T4m0
## a1
## a2
## a3
## a4
## a5
## Ci
## Cm
```

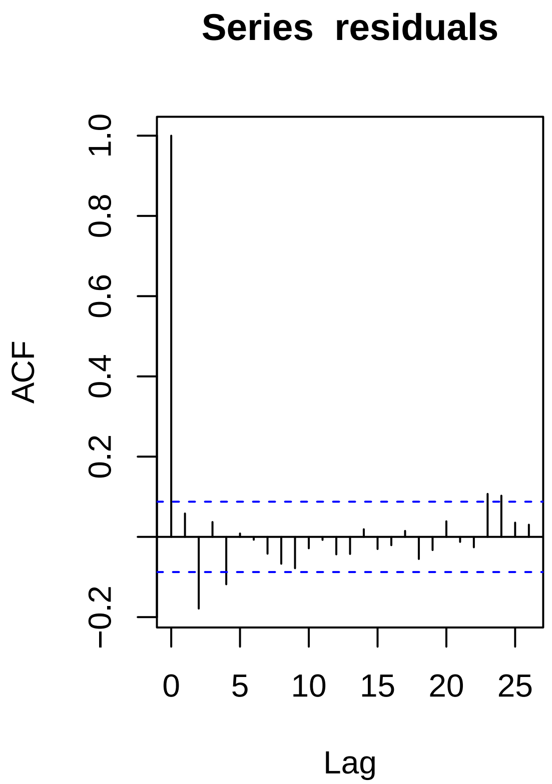
```
## e11
## e21
## e31
## e41
## p11a
## p11b
## p11c
## p11d
## p22a
## p22b
## p22c  0.32
## p22d -0.05  0.00
## Ria   0.17  0.17 -0.04
## Rim   0.54  0.36 -0.15  0.30
## Riw   0.04  0.03 -0.18  0.34  0.09
```

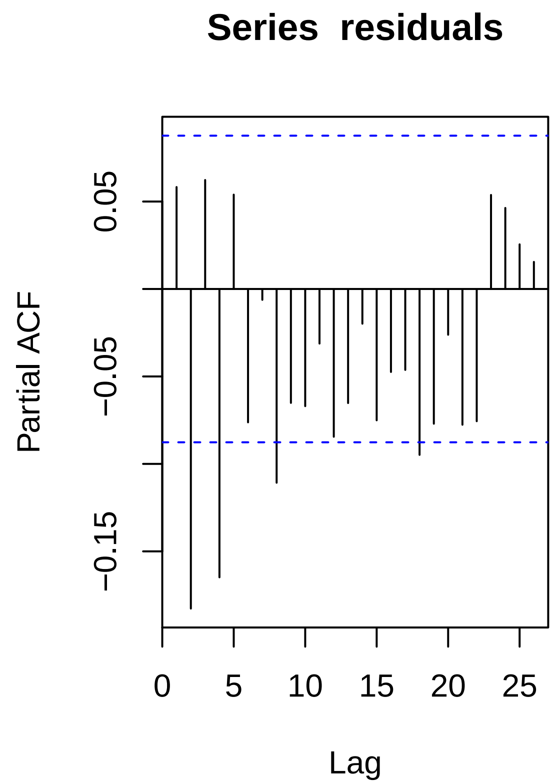
Inspecting the fit is evident that

1. All parameters are significant, except for a_2, a_4 and e_4 which is not of great importance.
2. None of the parameters are too strongly correlated.
3. Convergence has reasonably been met.
4. A very large improvement in AIC from the previous model.

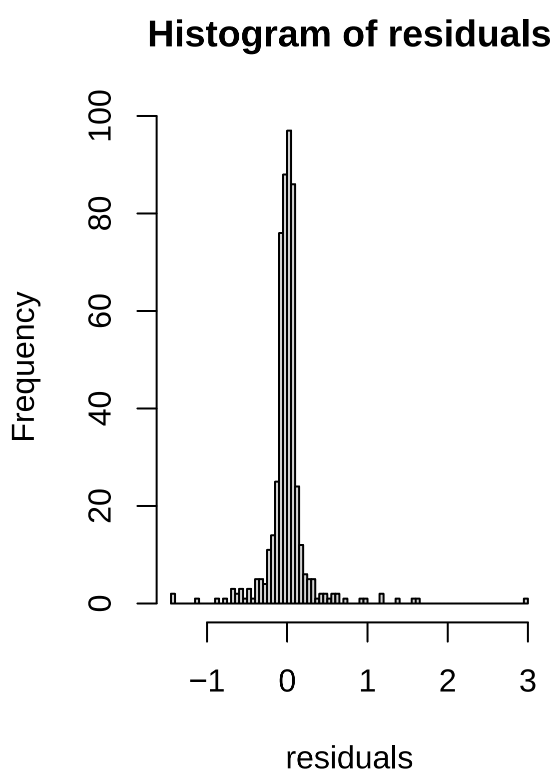
I will end this assignment with model validation of the final model. In order not to overload the reader, I will only show residuals for room 1 (which is then comparable to the previous single-room exercise).

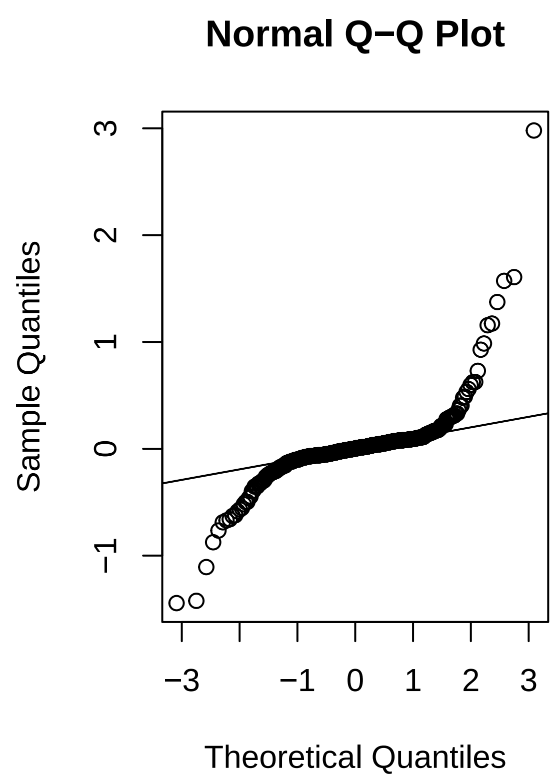
```
# To asses the model fit, calculate the one-step predictions and residuals.
# Do residual analysis for model validation.
D$Ph = (D$Ph1 + D$Ph2)/2; D$yT2 <- D$yTi2; D$yT1 <- D$yTi1; D$yT3 <- D$yTi3; D$yT4 <- D$yTi4;
preds <- predict(fitmulti2, newdata=D)

yTi1Hat <- preds$output$pred$yT1
residuals <- D$yTi1 - yTi1Hat

par(mfrow=c(1,2))
acf(residuals)
pacf(residuals)
```



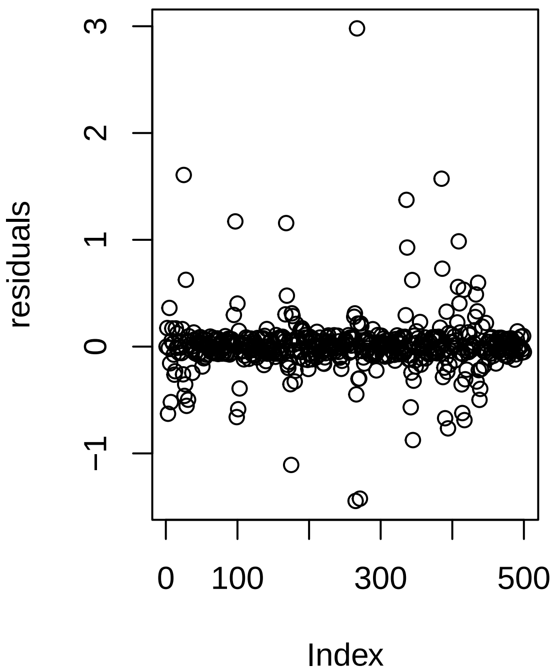


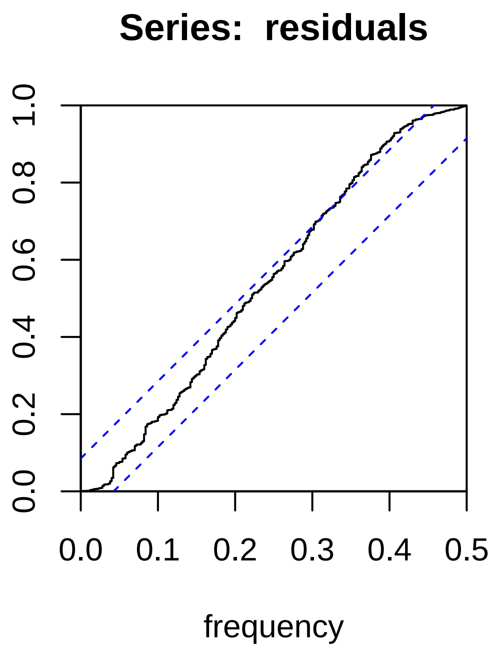

```
hist(residuals, breaks=100)
qqnorm(residuals)
qqline(residuals)
```





```
plot(residuals)
cpgram(residuals)
```





We see a clear improvement in the residuals from the single-room model, however there is still room for improvement. Next step would be to let the variance be non-constant, letting the diffusion-term be a function of the states.