Advanced Time Series Analysis: Computer Exercise 3

Anders Launer Bæk (s160159)

28 November 2017

Sparring partners:

- Anja Liljedahl Christensen (s162876)
- Marie Mørk (s112770)

Part 1: Simulation and discretization of diffusion processes

Equation 2a and 2b from the description have been discretized and they are showed in eqn. 1.

$$
\begin{aligned}
Y_{n+1}^1 &= Y_n^1 + \theta_3 \left(Y_n^1 + Y_n^2 - \frac{1}{3} \left(Y_n^1 \right)^3 + \theta_4 \right) \Delta + \sigma \Delta W_{n+1}^1 \\
Y_{n+1}^2 &= Y_n^2 - \frac{1}{\theta_3} \left(Y_n^1 + \theta_2 Y_n^2 - \theta_1 \right) \Delta
\end{aligned} \tag{1}
$$

The initial parameters for this diffusion process are given in eqn. 2.

$$
\begin{aligned}
Y_0^1 &= -1.9 \\
Y_0^2 &= 1.2 \\
\theta_{1,2,3,4} &= [0.7,\, 0.8,\, 3,\, -0.34] \\
\Delta &= 0.0019531 \\
\sigma &= 0 \\
T &= 100 \\
t &= 1 : \Delta : T \\
\Delta W_{n+1}^1 &\sim \mathcal{N}\left(0,\, \Delta\right)
\end{aligned} \tag{2}
$$

Question 1a

It is possible to change the process by changes the value of σ. An increase in σ will provide a bigger variation in the Wiener process. Below there have been plotted realizations of Y_k^1 and Y_k^2 wrt. time and a phase plot of Y_k^1 and Y_k^2. The following function `model_func()` has been used to plot

```
# function ----
model_func <- function(sigma, delta, t, Theta, init_values) {
    # initilize data.frame and initial values
    data <- data.frame(T = t, Y_1 = NA, Y_2 = NA)
    data$Y_1[1] <- init_values[1]
    data$Y_2[1] <- init_values[2]
    # simulate winer process
    set.seed(22)
    data$W <- rnorm(nrow(data), mean = 0, sd = delta)
    # run the simulation loop
    for (k in 1:(nrow(data) - 1)) {
        # Y_k^1
```

```
        data$Y_1[k + 1] <- data$Y_1[k] + Theta[3] * (data$Y_1[k] + data$Y_2[k] -
            1/3 * data$Y_1[k]^3 + Theta[4]) * delta + sigma * data$W[k + 1]
        # Y_k^2
        data$Y_2[k + 1] <- data$Y_2[k] - 1/Theta[3] * (data$Y_1[k] + Theta[2] *
            data$Y_2[k] - Theta[1]) * delta
    }
    # realizations
    re_plot <- ggplot2::ggplot(data) + ggplot2::geom_point(ggplot2::aes(x = T,
        y = Y_1, color = "Y_k^1"), alpha = 1/2) + ggplot2::geom_point(ggplot2::aes(x = T,
        y = Y_2, color = "Y_k^2"), alpha = 1/2) + ggplot2::labs(x = "t", y = "Y_k^*(t) ",
        color = "") + theme_TS()
    # phase
    ph_plot <- ggplot2::ggplot(data) + ggplot2::geom_point(ggplot2::aes(x = Y_1,
        y = Y_2, color = "Phase"), alpha = 1/2) + ggplot2::labs(x = "Y_k^1(t)",
        y = "Y_k^2(t)", color = "") + theme_TS()
    return(list(sim = data, re_plot = re_plot, ph_plot = ph_plot))
}
```

The realizations of Y_k^1 and Y_k^2 wrt. time and phase plots of Y_k^1 and Y_k^2 are constructed for different $\sigma = [0.0, 0.1, 0.2, 0.3, 0.4]$.

$\sigma = 0.00$

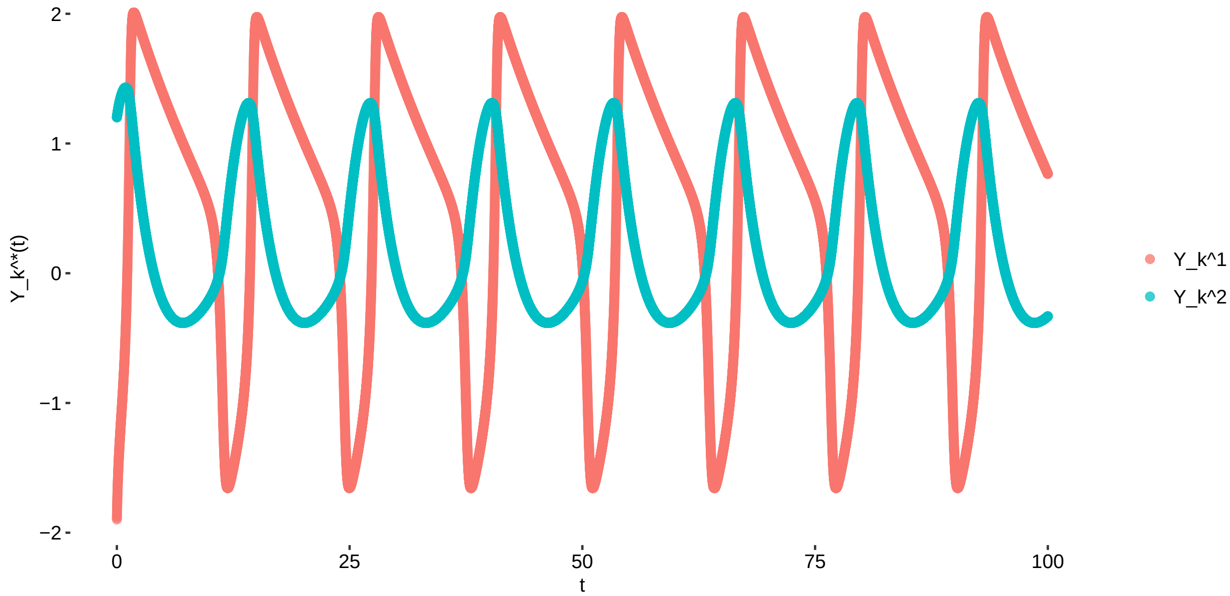

Figure 1: Plot of the simulation realizations with sigma = 0.0.

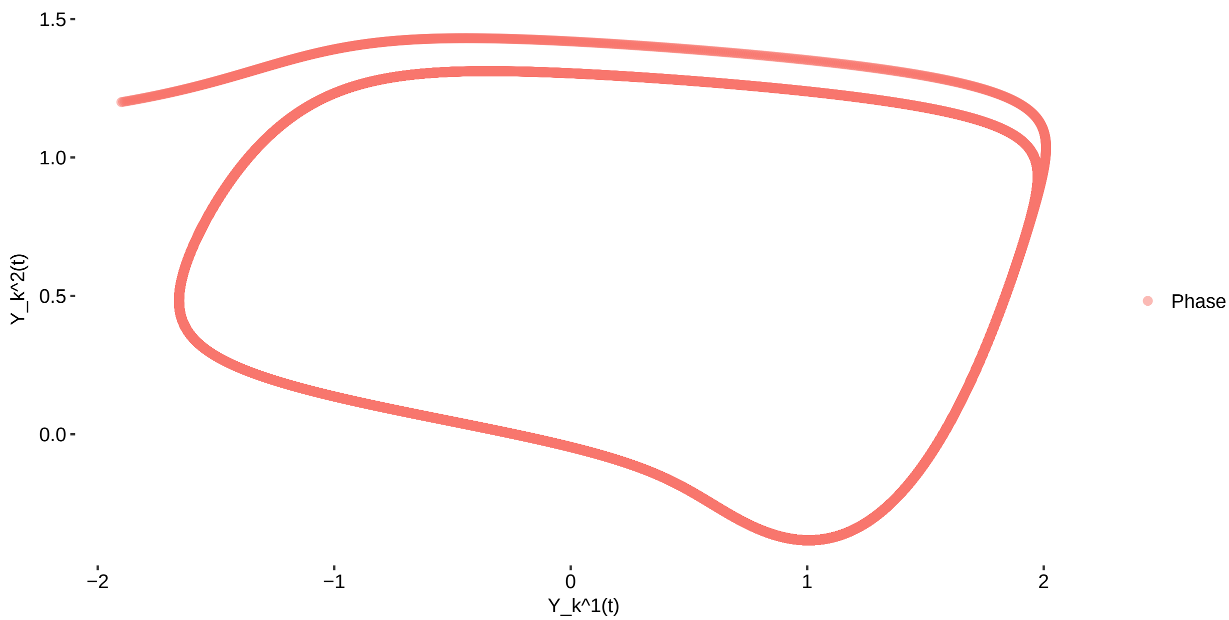

Figure 2: Phase plot of the simulation with sigma = 0.0.

$\sigma = 0.0$ returns a stable cyclic system. It takes ≈ 10 time steps for the system to converge to its cyclic pattern. The ≈ 10 time steps will be the same for all simulations. This is caused by the initial parameters of Y_0^1 and Y_0^2.

$\sigma = 0.10$

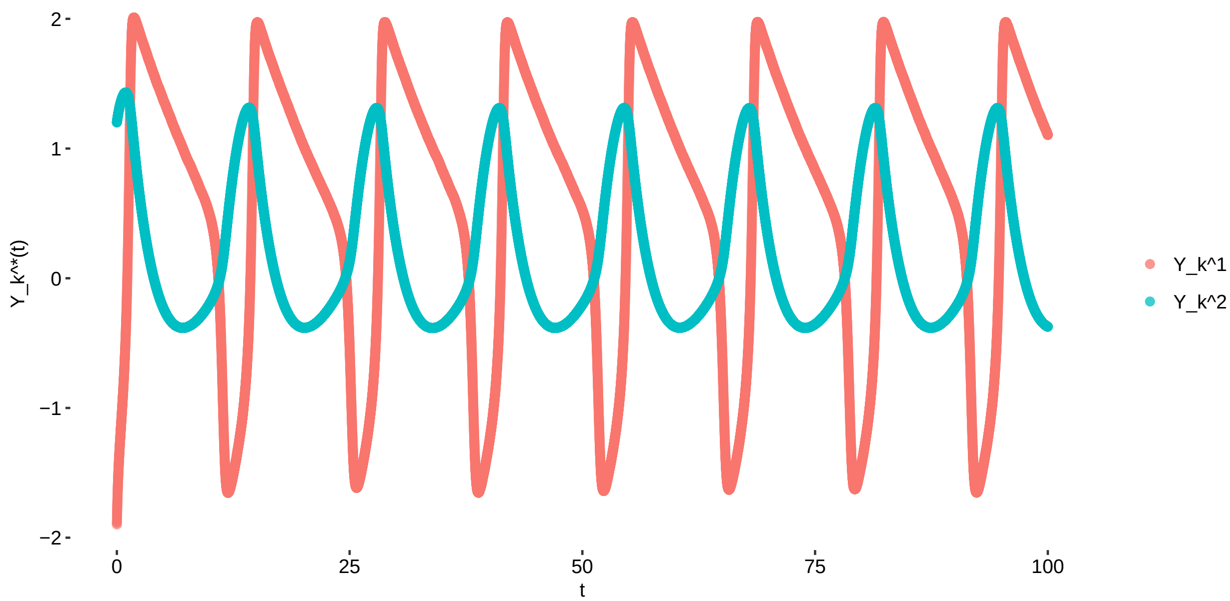

Figure 3: Plot of the simulation realizations with sigma = 0.10.

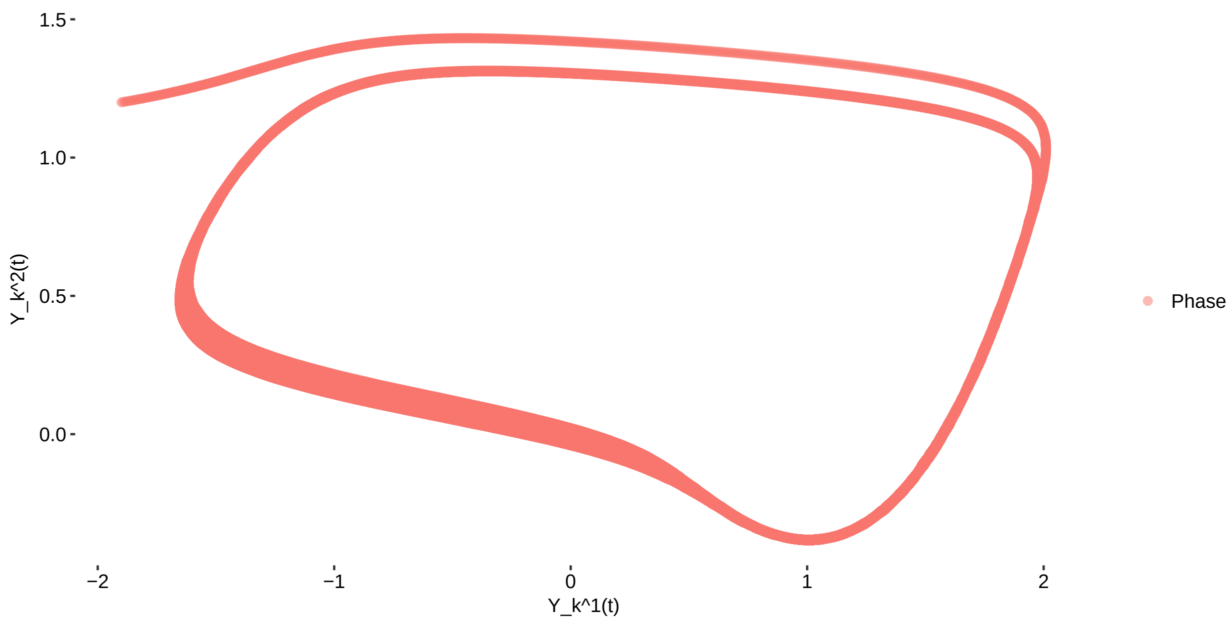

Figure 4: Phase plot of the simulation with sigma = 0.10.

Changing σ to 0.10 does not make a huge visual effect on the realizations. But it is possible to see the change in the phase plot. It is possible to see a "thicker line" in the lower left corner, which indicates that the change in σ impact Y_k^2 the most for Y_k^1 values in the range $[-1.75; 0.5]$.

$\sigma = 0.20$

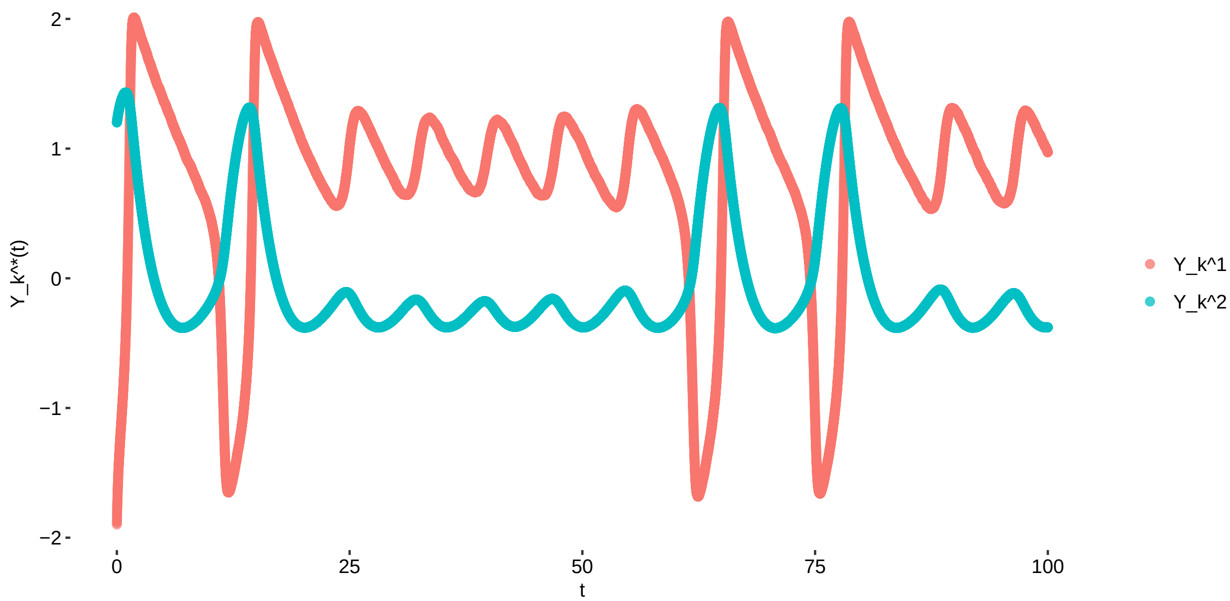

Figure 5: Plot of the simulation realizations with sigma = 0.20.

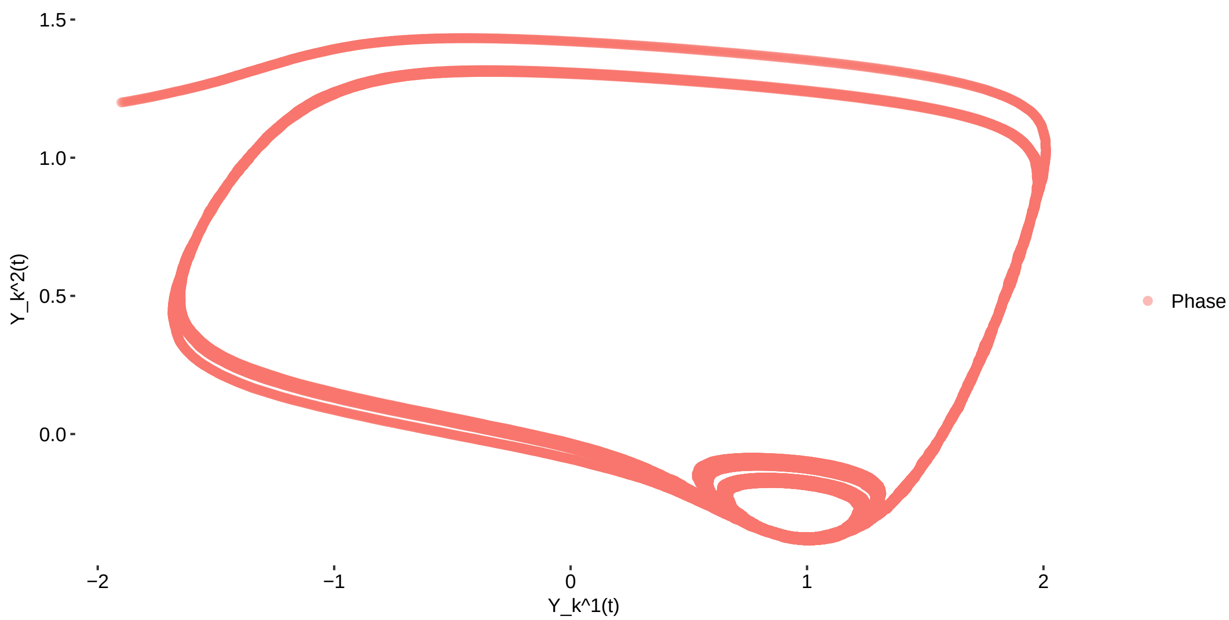

Figure 6: Phase plot of the simulation with sigma = 0.20.

Changing σ to 0.20 does make an effect in both the realizations and in the phase plot. The time sereis still need ≈ 10 time steps to converge to its normal cyclic pattern. The pattern changes dramatically after that for some cycles, back to normal and then back to the abnormal pattern again. The change in σ from 0.10 to 0.20 increase the effect of the Wiener process.

$\sigma = 0.30$

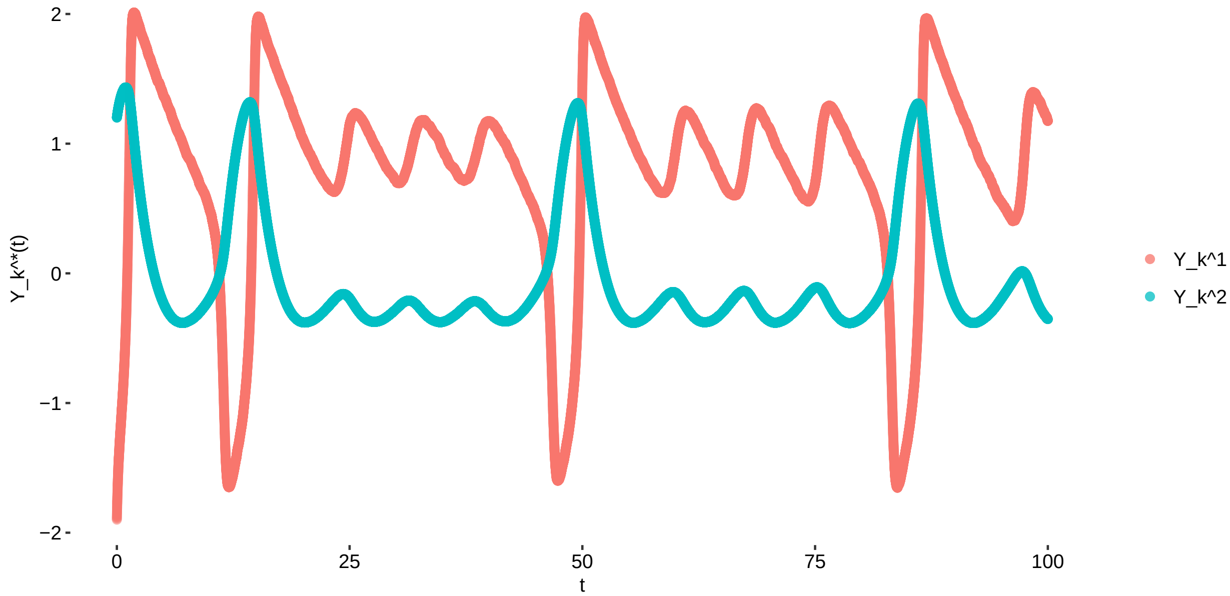

Figure 7: Plot of the simulation realizations with sigma = 0.30.

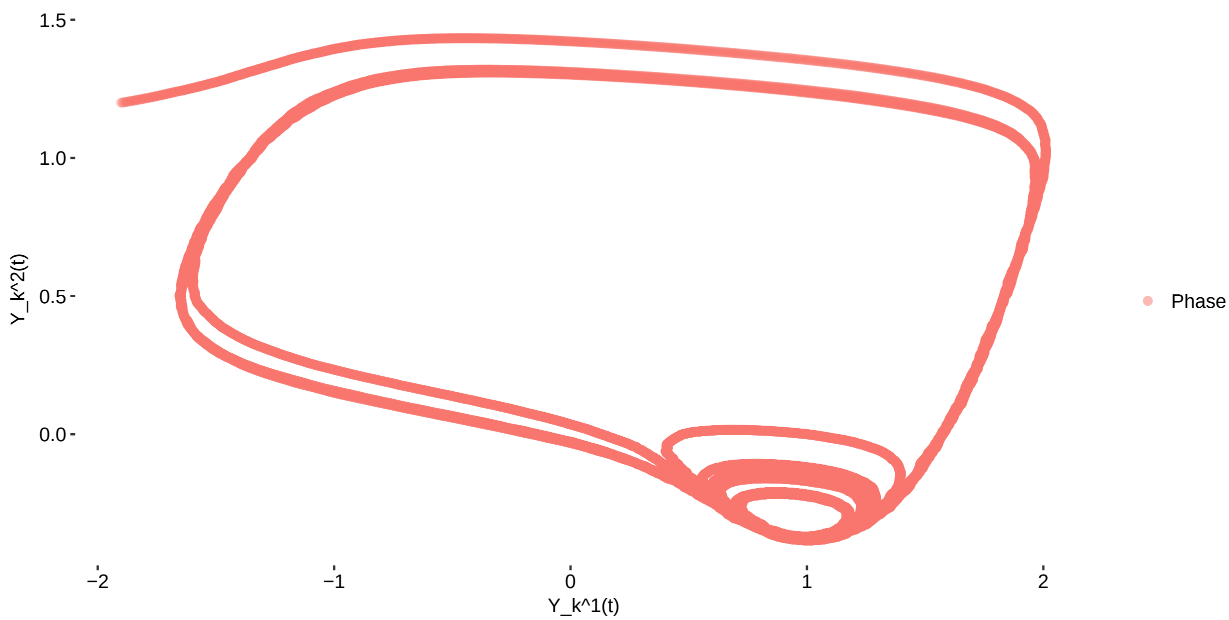

Figure 8: Phase plot of the simulation with sigma = 0.30.

Changing σ from 0.20 to 0.30 gives more less the same illustrations as before. It is noticeable to see a greater variance Y_k^2 when Y_k^1 is in the range $[-1.75; 1.5]$.

$\sigma = 0.40$

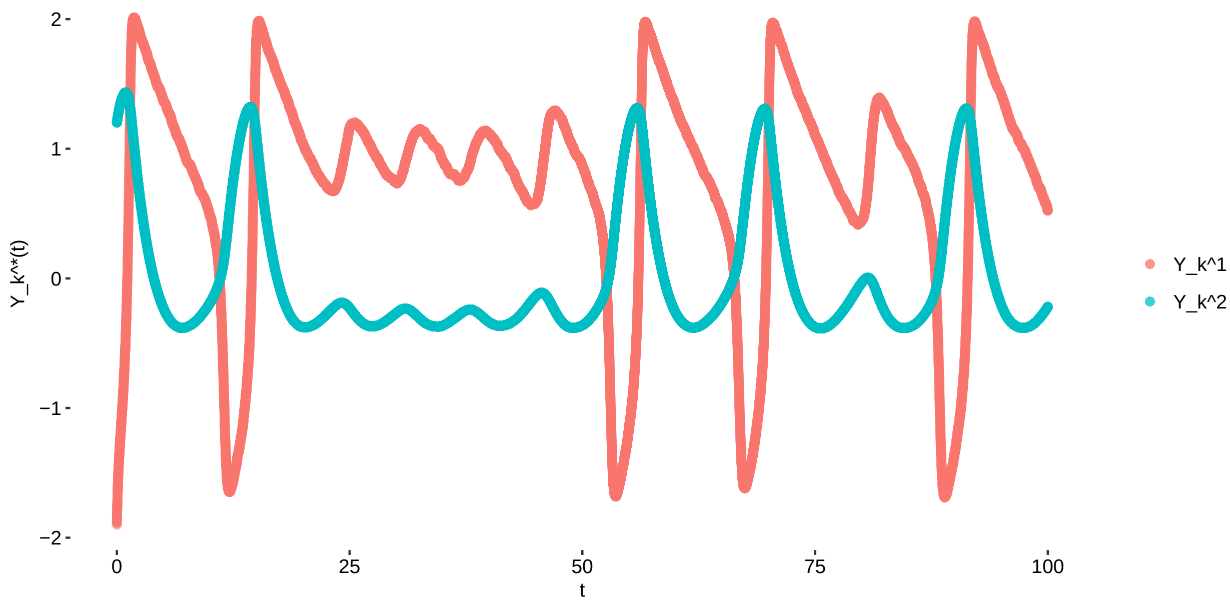

Figure 9: Plot of the simulation realizations with sigma = 0.40.

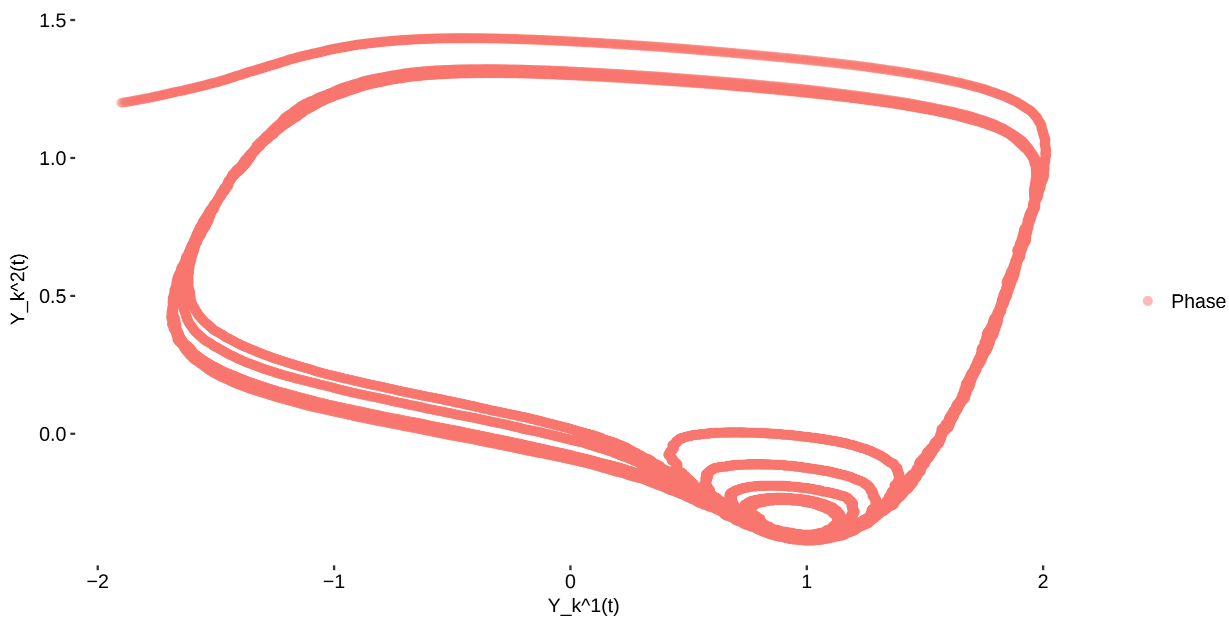

Figure 10: Phase plot of the simulation with sigma = 0.40.

For $\sigma = 0.40$ the changes between the normal cyclic pattern and the abnormal pattern is changing more frequently.

The length of the abnormal cycles has great influence of the amplitude of the time series. The amplitudes are more similar to the normal pattern when the abnormal pattern is short. For longer period of abnormal pattern entail smaller amplitudes. The smaller amplitudes creates more cycles in the lower right corner of the phase plot.

Comment on the effect of adding noise to the equations

- The first and most noticeable visual change by increasing the noise (σ) is the amplitude and the number of cycles in the abnormal pattern. Those effecting the "number of cycles" in the abnormal pattern where greater values of σ entail more frequent change between the normal pattern and the abnormal pattern.
- A greater value of σ has a small impact of the effected range of Y_k^1. The effect range of Y_k^1 is for all four simulation more less the given range $[-1.75; 1.5]$. Where the effect range of Y_k^2 is increasing in all four simulations.
- MORE? TODO

Question 1b

The same function (`model_func()`) has been applied to create the simulations for given values of σ. The `stat_bin2d(bins=100)` function has been used to create the 100x100-grid in the phase plot in order to count the number of trajectories in each cell.

The simulated phase plots of Y_k^1 and Y_k^2 are constructed for following values: $\sigma = [0.0, 0.1, 0.2, 0.3, 0.4]$.

$\sigma = 0.10$

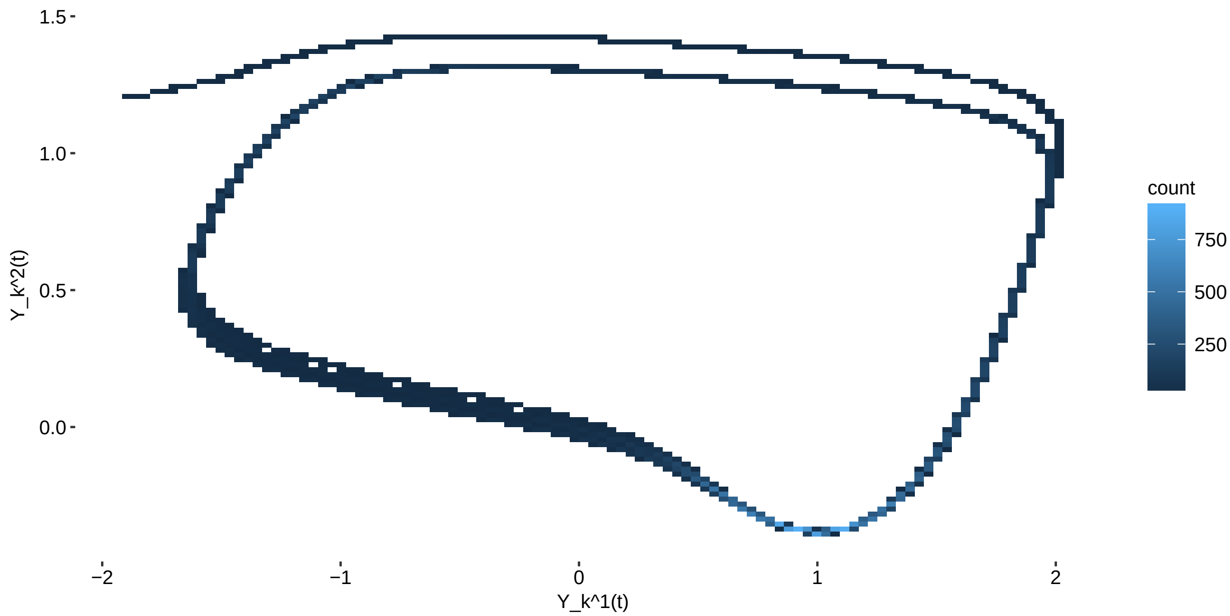

Figure 11: Phase plot of the simulation with sigma = 0.10.

$\sigma = 0.20$

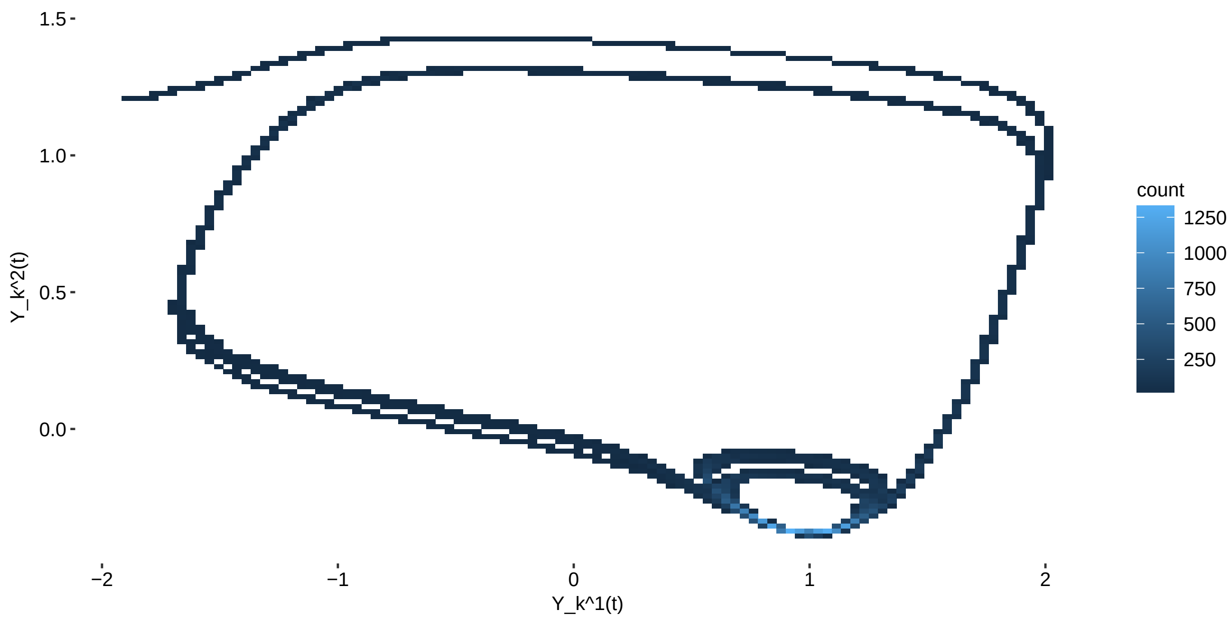

Figure 12: Phase plot of the simulation with sigma = 0.20.

$\sigma = 0.30$

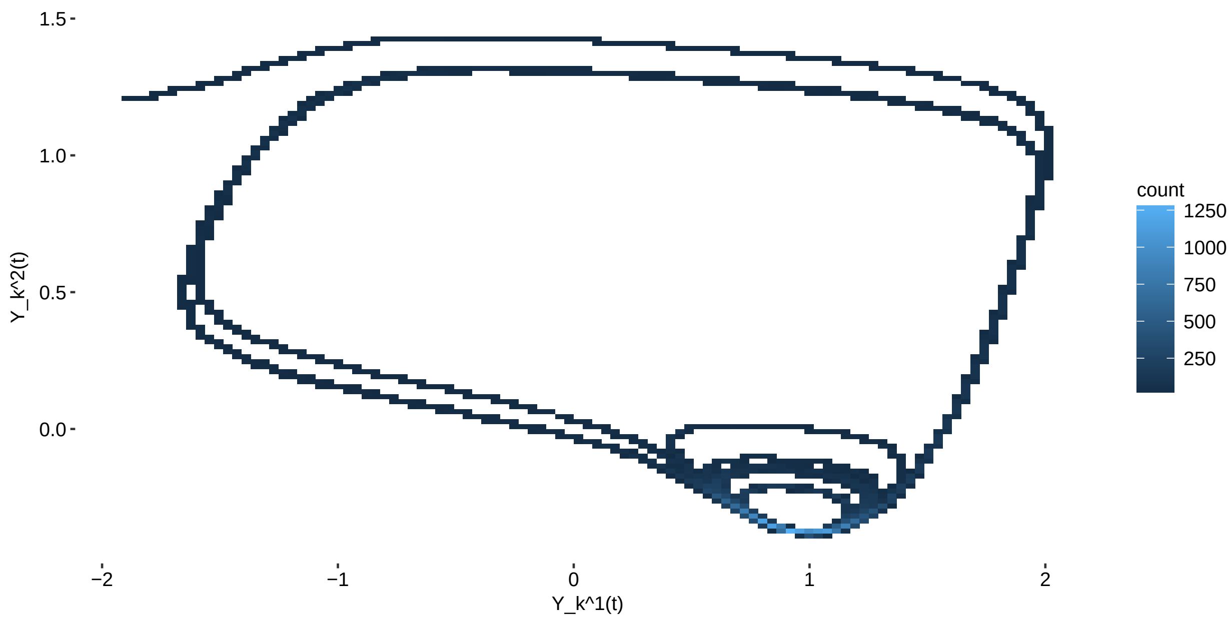

Figure 13: Phase plot of the simulation with sigma = 0.30.

$\sigma = 0.40$

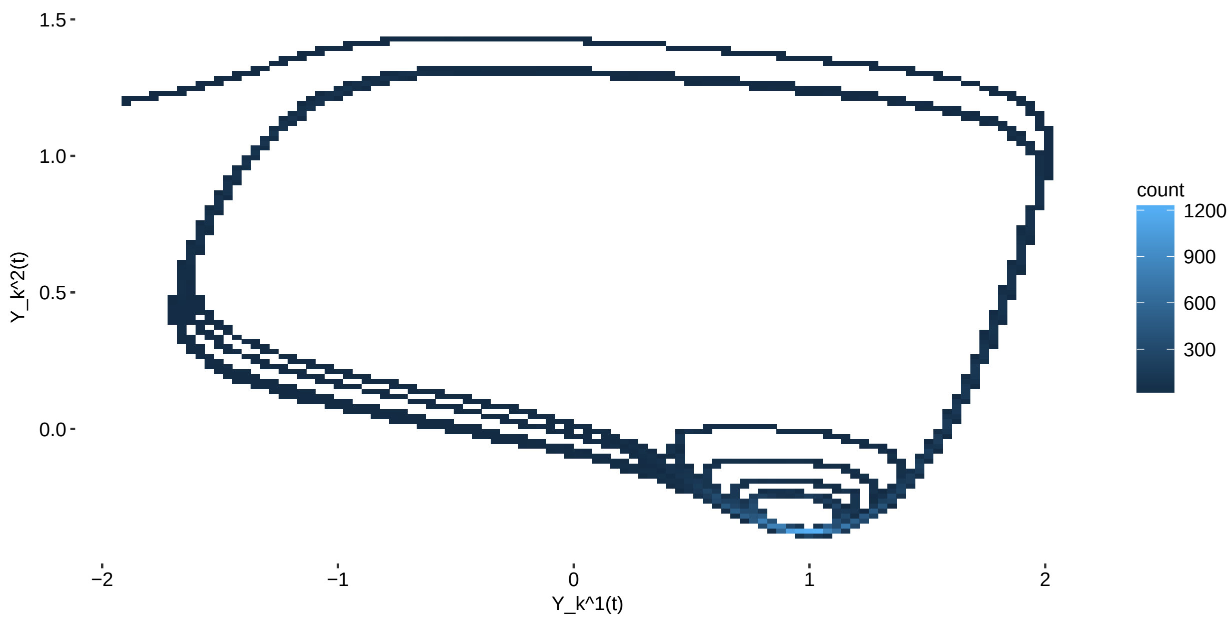

Figure 14: Phase plot of the simulation with sigma = 0.40.

Which extra information does the plot contain, compared to the standard phase-plot?

- Figure 10 and figure 14 illustrates the same phase plot with same value of σ. Figure 14 has added another dimension which adds additional value to the plot compared to figure 10.
- It is much easier to see where the system spent the most of its time when counting the number of trajectories in each cell.
- MORE? TODO

Part 2: Models for the heat dynamics of a high performance test building

Data

The provided data has following properties:

- *timedate* The time of the sample in UTC.
- TiE The indoor air temperature of the East room $[^{\circ}C]$.
- TiW The indoor air temperature of the West room $[^{\circ}C]$.
- Ta The ambient temperature $[^{\circ}C]$.
- Gv The solar radiation on a vertical surface facing south $\left[\frac{kW}{m^2}\right]$.
- PhE The power of the heater in the East room $[kW]$.
- PhW The power of the heater in the West room $[kW]$.

The sample period is $10\,[min]$.

NB: It is worth mention that there has not been been considered any kind of outlier detection prior to model fitting. But there is some "dramatic" behaviour around mid day between Oct. 12 and Oct. 13.

Question 2a

Question 2a focus at the East room and uses TiE as input variable yTi and TiW as target variable Tn.

The script `fitmodel.R` has been implemented step by step in the following sections.

1. Step

Figure 15 shows the interesting recorded time series in three sub-plots:

- A step sequence which tells when the heater is on Ph in the east room.
- The ambient temperature Tn, the indoor air temperature in the east room yTi and the indoor air temperature in the west room Ta.
- The solar radiation on a vertical surface facing south Gv.

It is possible to see how the Solar radiation, on surface facing south, increases the temperatures and the pattern for the heater in figure 15.

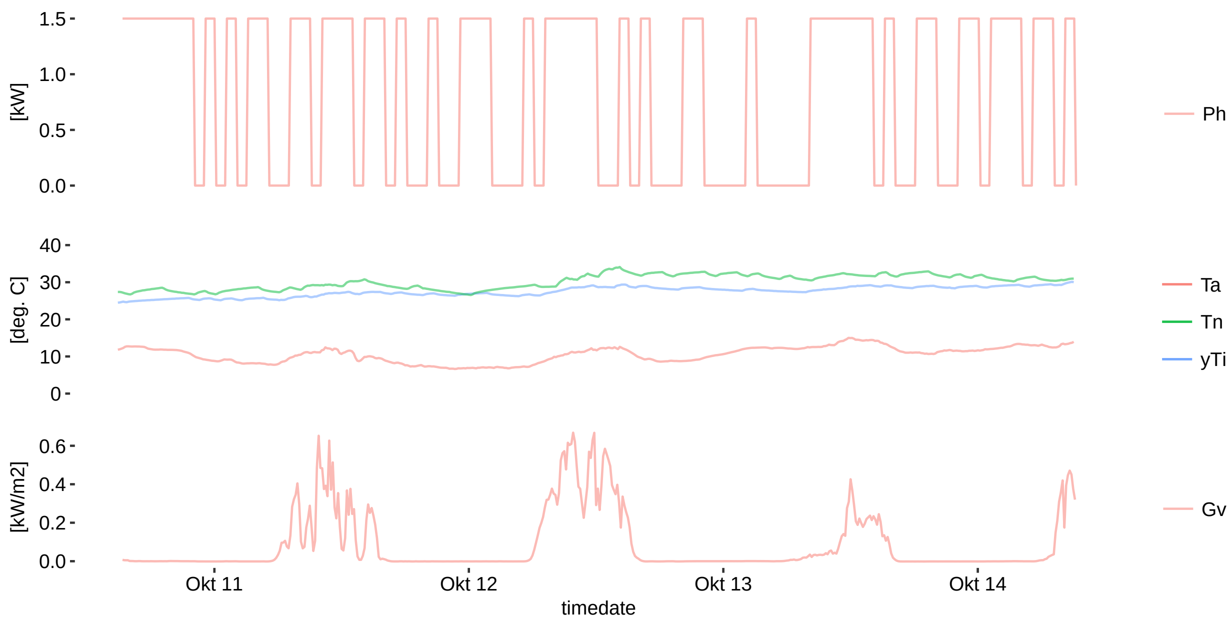

Figure 15: Plot of the captured time series.

2. Step

There has been implemented the most simple model of the system in step 2.

Equation 3 and eqn. 4 shows the system eqn. and measurement eqn. respectively.

$$dT_i = \left(\frac{1}{R_{ia}C_i} (T_a - T_i) + \frac{1}{C_i}\Phi_h \right) dt + \sigma_i d\omega_i \tag{3}$$

$$T_{t_k} = T_{i,\,t_k} + \epsilon_{t_k} \tag{4}$$

3. Step

Below is the summary of the fit and the estimated parameters.

```
## Coefficients:
##         Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.4502e+01  7.5142e-02  3.2607e+02  0.0000e+00 -3.4124e-03    0.2468
## Ci   1.0924e+01  2.0632e+00  5.2948e+00  1.7518e-07 -4.0554e-05    0.0003
## e11 -2.2092e+01  4.1092e+00 -5.3763e+00  1.1438e-07 -1.6859e-04    0.0001
## p11 -1.6420e+00  2.6167e-02 -6.2751e+01  0.0000e+00  5.6728e-04    0.0000
## Ria  9.9858e+00  3.7836e-02  2.6392e+02  0.0000e+00  6.0841e-03   49.5438
##
## Correlation of coefficients:
##     Ti0   Ci    e11   p11
## Ci  -0.11
## e11  0.24 -0.28
## p11 -0.12  0.05 -0.07
## Ria -0.22  0.03 -0.97  0.05
##
## [1] "loglikelihood = 599.62755058139"
```

The optimization procedure works out without any problems but we compare the value in $dF/dPar$ with the value in $dPen/dPar$. If the values are significantly different, the particular initial parameter value it is close to one of its limits. A solution to this is to loosen the particular initial parameter value. [1]

I fixed this issue by increasing the limits on both of the following two parameters Ria and $Ti0$. Then estimate the model again.

```
## Coefficients:
##        Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.4502e+01  7.3830e-02  3.3187e+02  0.0000e+00  0.0000e+00     8e-04
## Ci   3.4567e+00  2.4954e-01  1.3852e+01  0.0000e+00 -4.3892e-06     0e+00
## e11 -2.2491e+01  2.3055e+02 -9.7556e-02  9.2232e-01 -2.0623e-05     1e-04
## p11 -1.7459e+00  2.6627e-02 -6.5568e+01  0.0000e+00 -5.4510e-06     0e+00
## Ria  2.6260e+01  2.4301e+00  1.0806e+01  0.0000e+00 -1.4811e-05     0e+00
##
## Correlation of coefficients:
##     Ti0   Ci    e11   p11
## Ci   0.00
## e11  0.00  0.00
## p11 -0.01  0.04  0.00
## Ria  0.00  0.25  0.00  0.05
##
## [1] "loglikelihood = 653.325639755608"
```

It is possible to see the new estimated parameters in the output above. The estimated values of the two affected parameters are now $Ti0 = 24.502$, which was close to a initial binding value for $Ti0$, and $Ria = 26.260$ which is roughly about twice as munch as the initial upper limit value for Ria.

By fixing the initial limit issues also increased the log-likelihood significantly and decreased the correlations among the parameters which is desirable.

[1] P.15 http://ctsm.info/pdfs/ctsmr-reference.pdf

4. Step

The estimated model is analysed in step 4. We are interested in the residuals of the one step ahead prediction.

```
# Calculate the one-step predictions of the state
pred <- predict(fit.Ti)
# Extract the estimated value of yTi
data$yTiHat <- pred[[1]]$output$pred$yTi
# Calculate the residuals and add them to the data frame
data$residuals <- data$yTi - data$yTiHat
```

Time Series of the residuals

Figure 16 shows a time Series plot of residuals from the one step ahead predictions. It is clear to see that the residuals if not white noise which indicates that there is some systematic behaviour left.

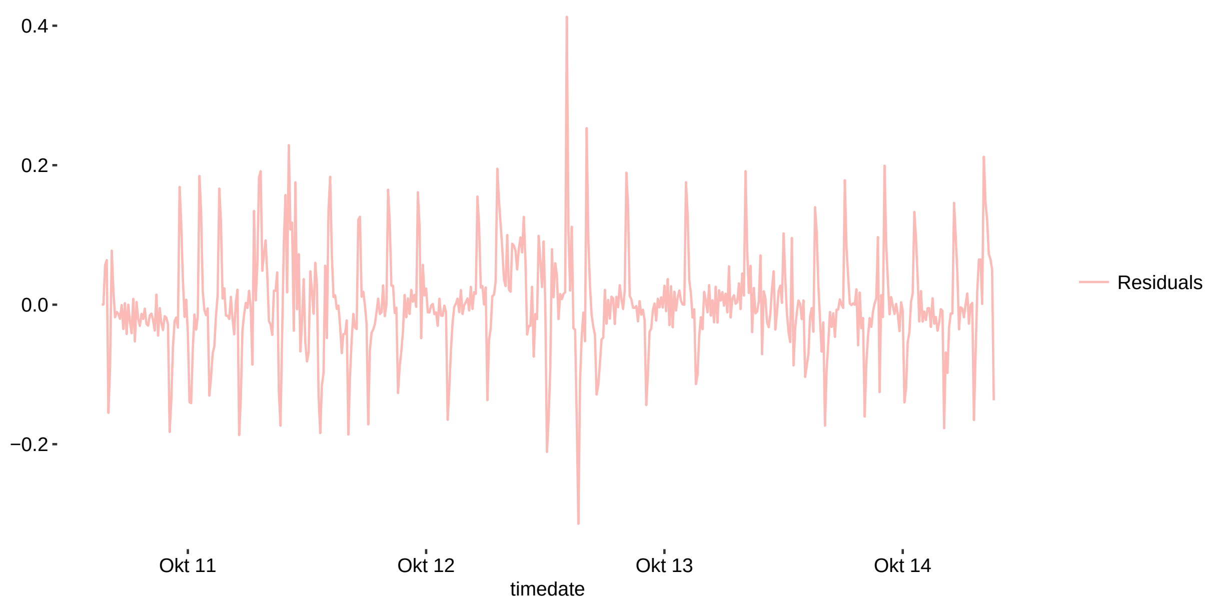

Figure 16: Time serie plot of the residuals.

Distribution of the residuals

Figure 17 shows a histogram of the residuals. This plot supports the statement from above that the residuals is not Gaussian distributed and therefore not white noise.

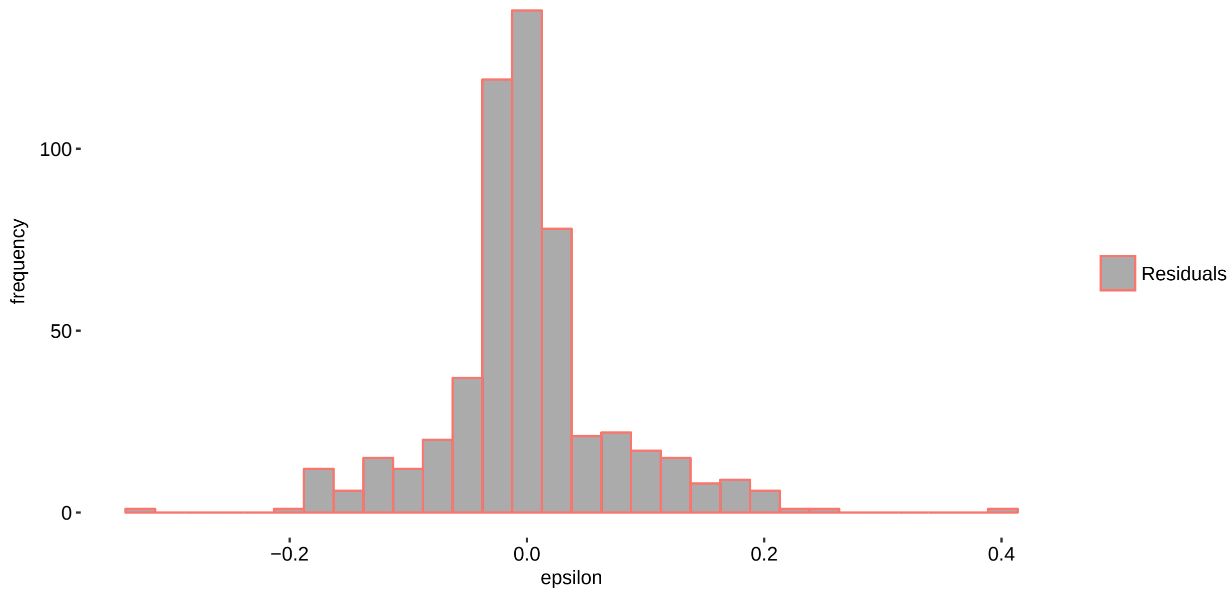

Figure 17: Histogram of the residuals.

ACF and PACF of the residuals

Figure 18 and figure 19 shows the ACF and the PACF of the residuals.

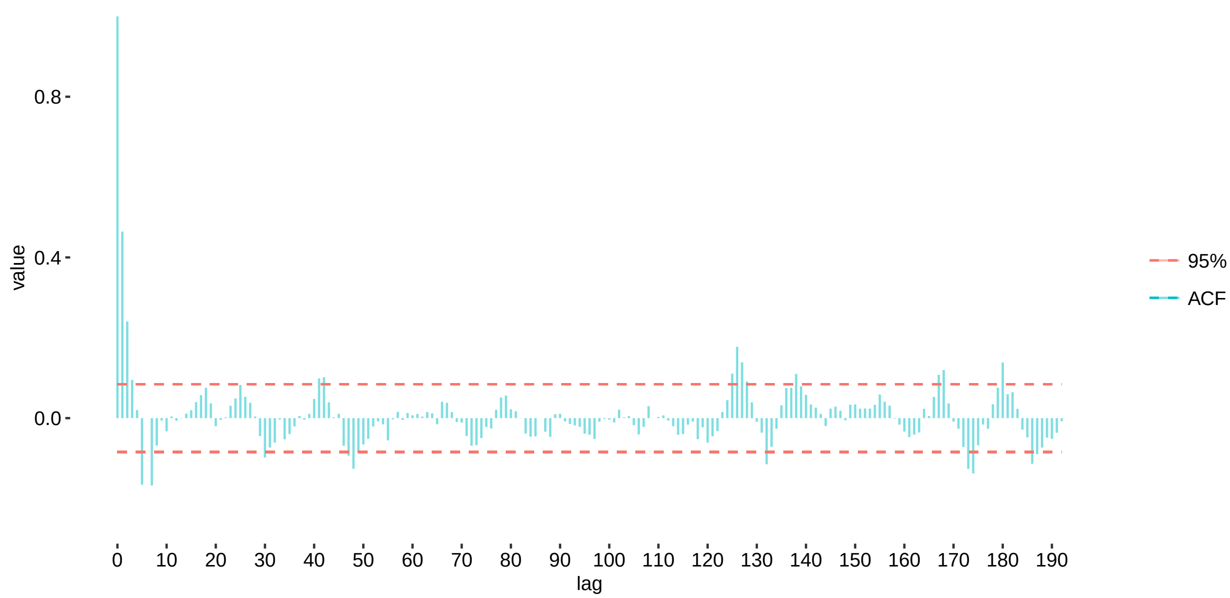

Figure 18: ACF of the residuals.

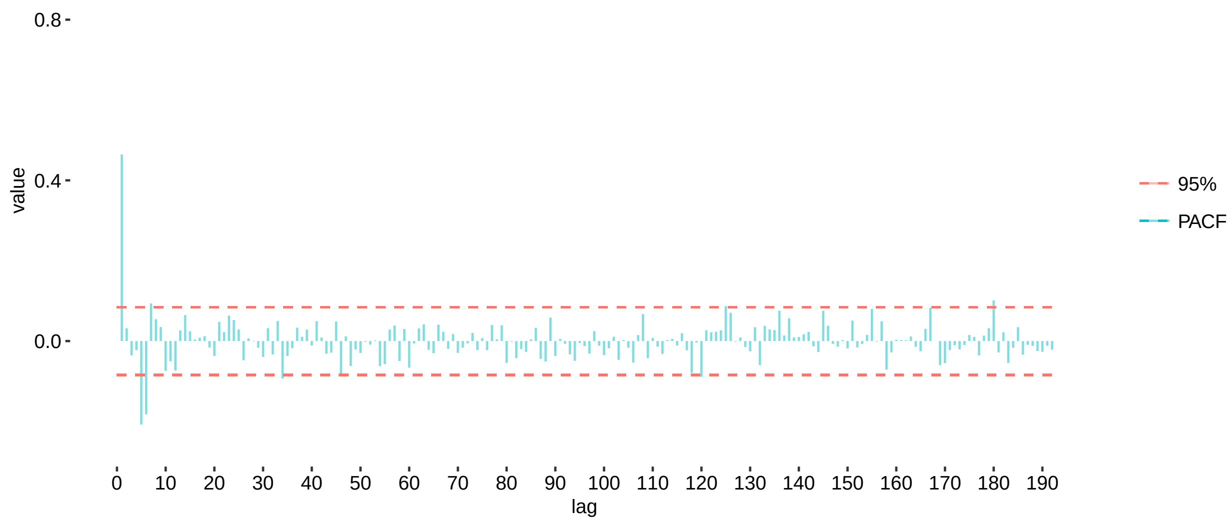

Figure 19: PACF of the residuals.

Both the ACF and the PACF of the residuals shows some correlation in the residuals, which indicates that the model can be improved.

Periodogram of the residuals

Figure 20 shows the periodogram of the residuals.

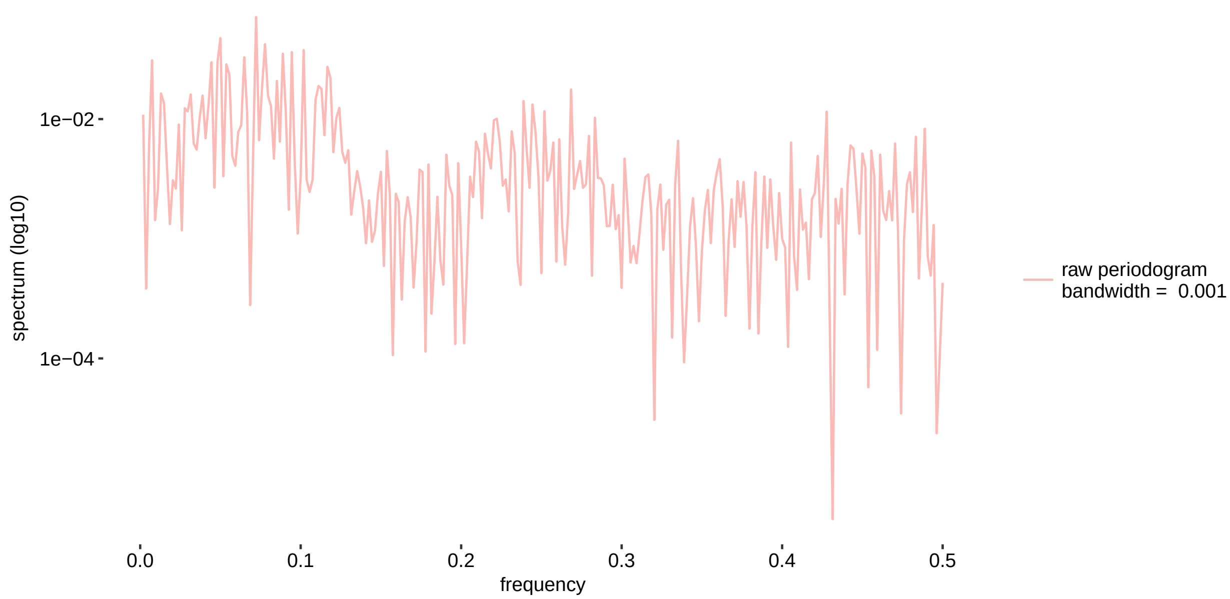

Figure 20: Periodogram of the residuals.

Cumulated periodogram of the residuals

Figure 21 shows the cumulated periodogram of the residuals. It is possible to see that the residuals are dependent of each other, which identicates that there is still unexplored systematic behaviour left in the residuals. The residuals will be on the diagonal in cumulated periodogram if they are white noise.

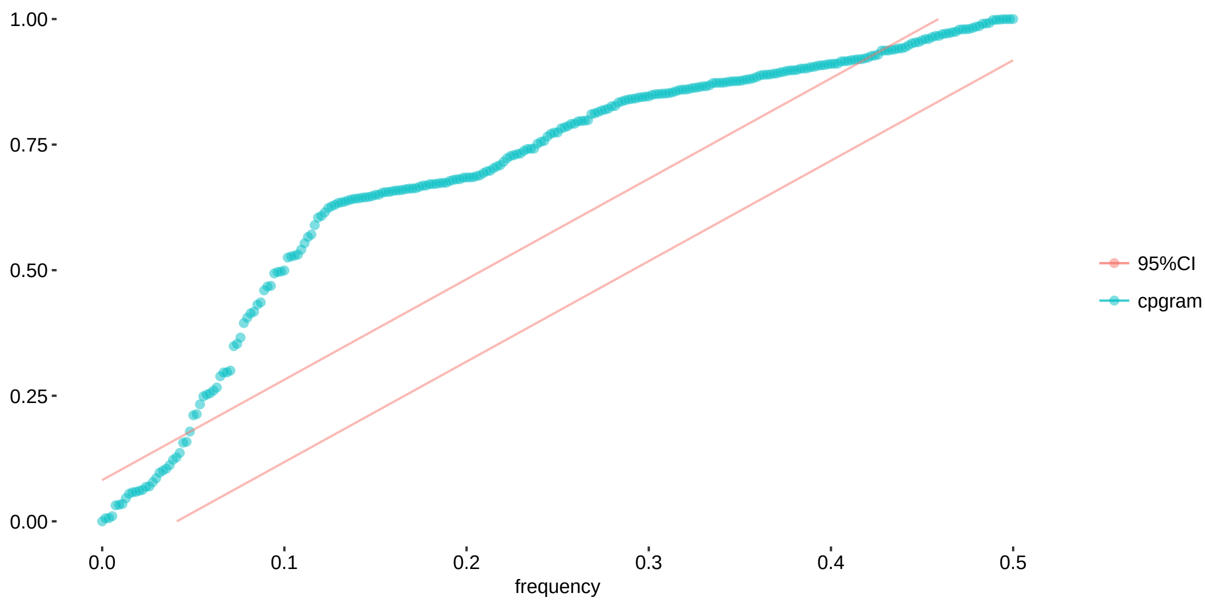

Figure 21: Cumulated periodogram of the residuals.

Combined plot

The one step ahead residuals have been placed in figure 22 have been placed above the three sub-plots in figure 15.

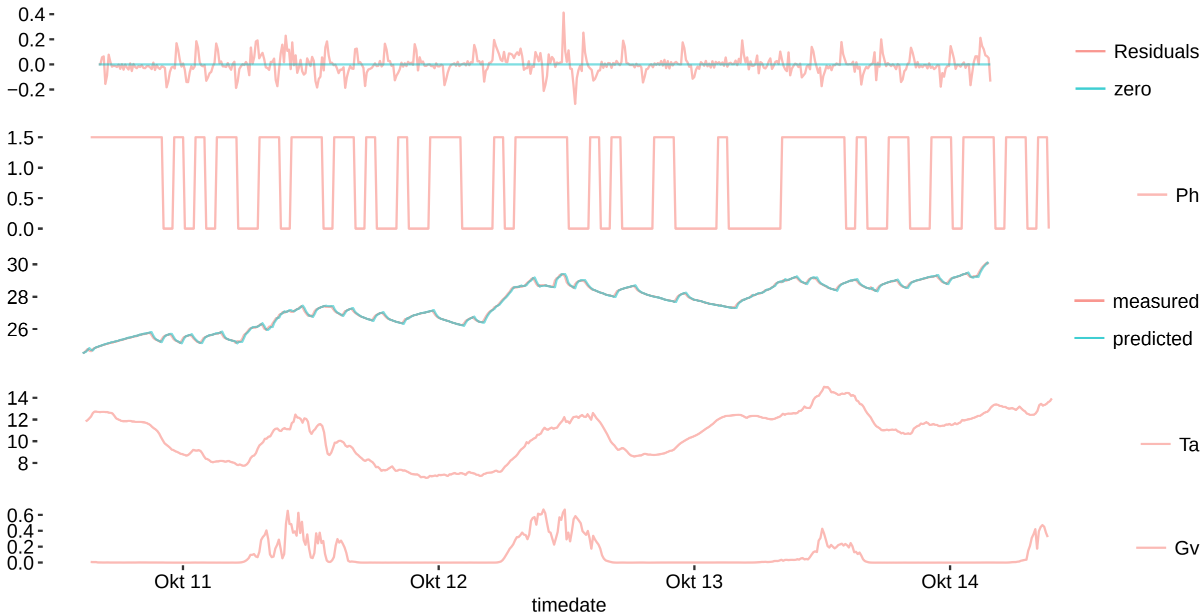

Figure 22: Conbined plot with the residuals, the heater, measured and predicted temperature, ect..

It is much easier to see a systematic behaviour in the residuals in figure 22 compared to figure 16. It is possible to see a cyclic pattern in the residuals when there is no Gv and only the Ph contributes to the temperature of the room.

5. Step

The model of the system (eqn. 3 and eqn. 4) have been updated to contain an new parameter R_{im} and a new state dT_m. The measurement equation remains the same as in eqn. 4. The new updated model is given in eqn. 5.

$$
\begin{aligned}
dT_i &= \left(\frac{1}{R_{im}C_i} (T_m - T_i) + \frac{1}{R_{ia}C_i} (T_a - T_i) + \frac{1}{C_i}\Phi_h \right) dt + \sigma_i d\omega_i \\
dT_m &= \left(\frac{1}{R_{im}C_m} (T_a - T_m) \right) dt + \sigma_m d\omega_m
\end{aligned}
\tag{5}
$$

The new $TiTm$ model has been estimated again and has following estimated parameters:

```
## Coefficients:
##        Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.4492e+01  3.7617e-02  6.5109e+02  0.0000e+00  4.4632e-03     4e-04
## Tm0  2.4381e+01  1.2695e-01  1.9205e+02  0.0000e+00  5.9287e-04     4e-04
## Ci   1.2220e+00  1.0998e-01  1.1111e+01  0.0000e+00 -3.5434e-05     0e+00
## Cm   1.3167e+01  6.5997e+00  1.9951e+00  4.6544e-02 -7.6129e-06     0e+00
## e11 -7.1306e+00  1.2886e-01 -5.5338e+01  0.0000e+00 -1.7809e-04     0e+00
## p11 -8.8369e+00  1.9800e-01 -4.4630e+01  0.0000e+00  4.9749e-05     0e+00
## p22 -1.5431e+00  8.5533e-02 -1.8042e+01  0.0000e+00 -1.2287e-04     0e+00
## Ria  6.6759e+01  9.2717e+01  7.2003e-01  4.7182e-01  8.4109e-07     0e+00
## Rim  2.1345e-01  1.8749e-02  1.1385e+01  0.0000e+00 -4.3269e-05     0e+00
##
## Correlation of coefficients:
##       Ti0   Tm0    Ci    Cm   e11   p11   p22   Ria
## Tm0  0.24
## Ci   0.00  0.07
## Cm   0.00 -0.13  0.46
## e11 -0.02 -0.02 -0.36 -0.12
## p11  0.02  0.01 -0.14  0.02  0.17
## p22  0.00  0.07  0.66  0.00 -0.37 -0.08
## Ria  0.01 -0.38  0.23  0.53 -0.11 -0.06 -0.03
## Rim -0.02 -0.04 -0.27 -0.18  0.51  0.02 -0.03 -0.17
##
## [1] "loglikelihood = 771.000411040433"
```

- An effect of adding the new state to the system is the increased log-likelihood value.
- The Ria parameter has a quiet large p-value but is still significant. The standard error is also quiet large compared to the estimated value of the parameter.

6. Step

The similar one step ahead prediction errors as in step 4. plotted and analysed in step 6.

Time Series of the residuals

Figure 23 shows a time series plot of residuals from the one step ahead predictions. It is clear to see that the residuals if not white noise which indicates that there is still some systematic behaviour within the residuals.

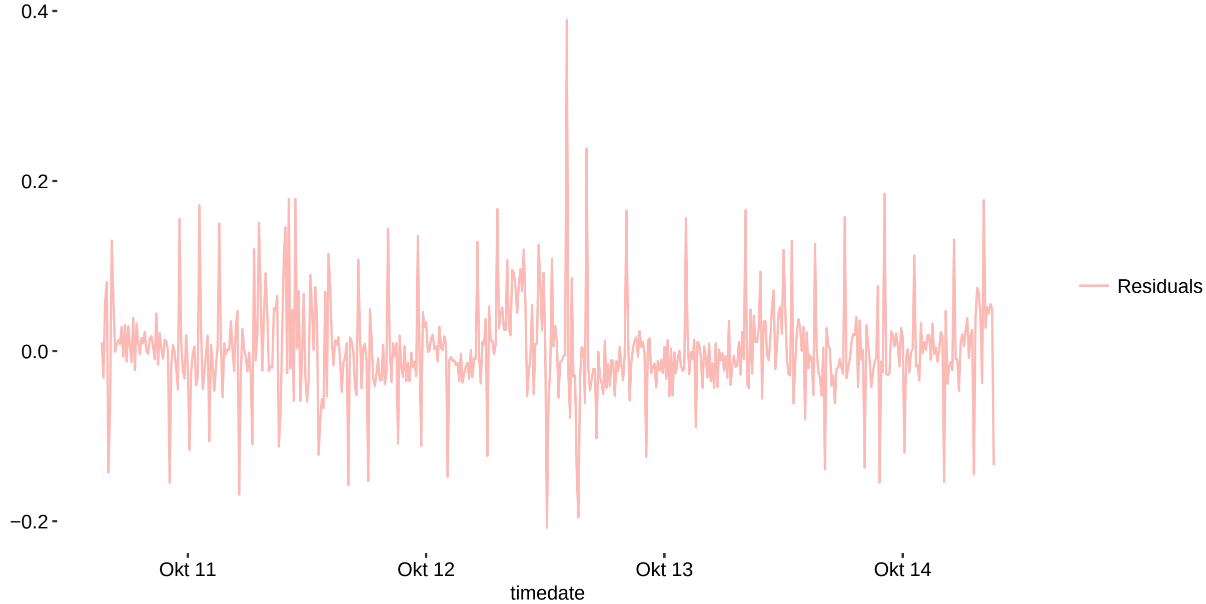

Figure 23: Time serie plot of the residuals.

Distribution of the residuals

Figure 24 shows a histogram of the residuals. If you compare the residuals from the simple model (eqn. 3 and figure 17) by the distribution of the residuals in from the new updated model (eqn. 5, figure 24), the residuals tends to be more in the positive region of the histogram. The residuals are still not Gaussian distributed and hereby not white noise.

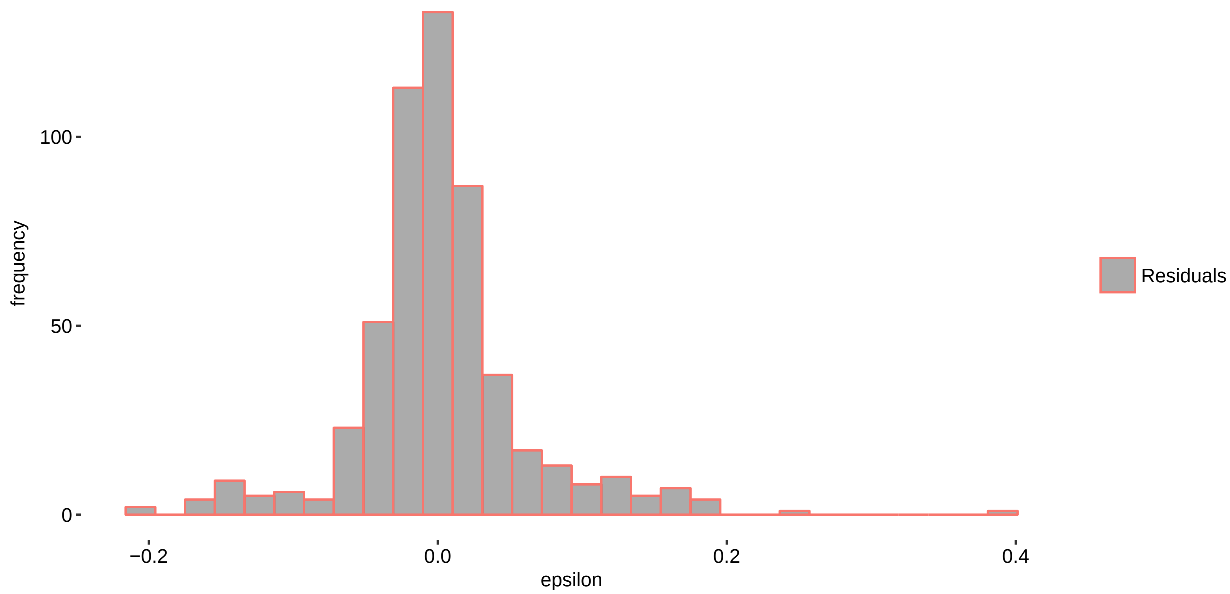

Figure 24: Histogram of the residuals.

ACF and PACF of the residuals

Figure 25 and figure 26 shows the ACF and the PACF of the residuals.

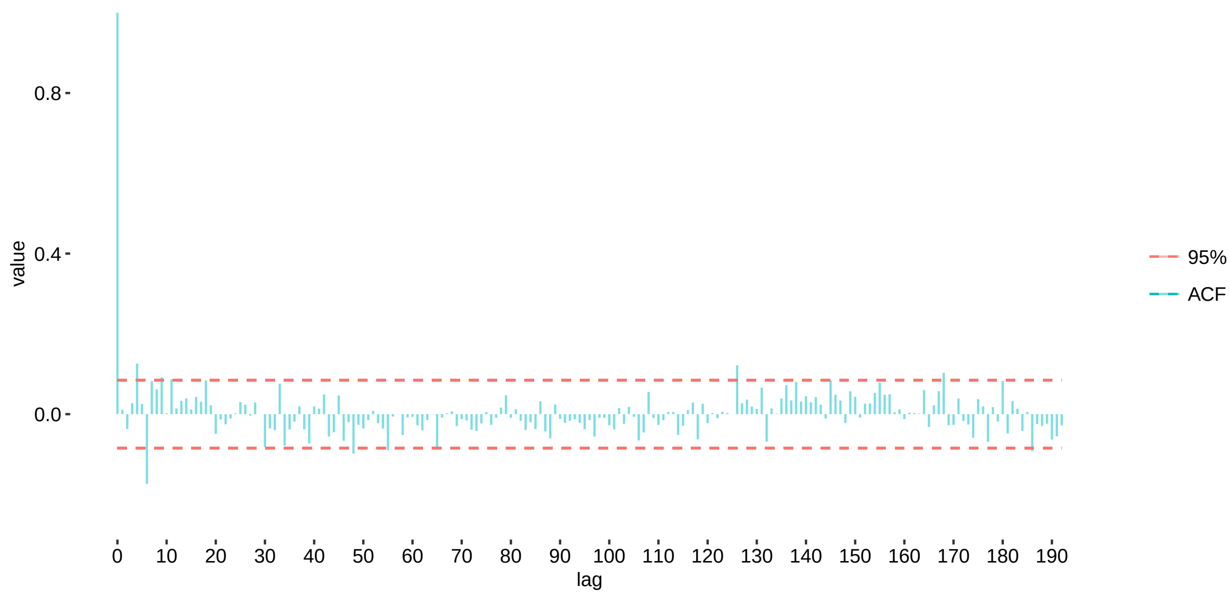

Figure 25: ACF of the residuals.

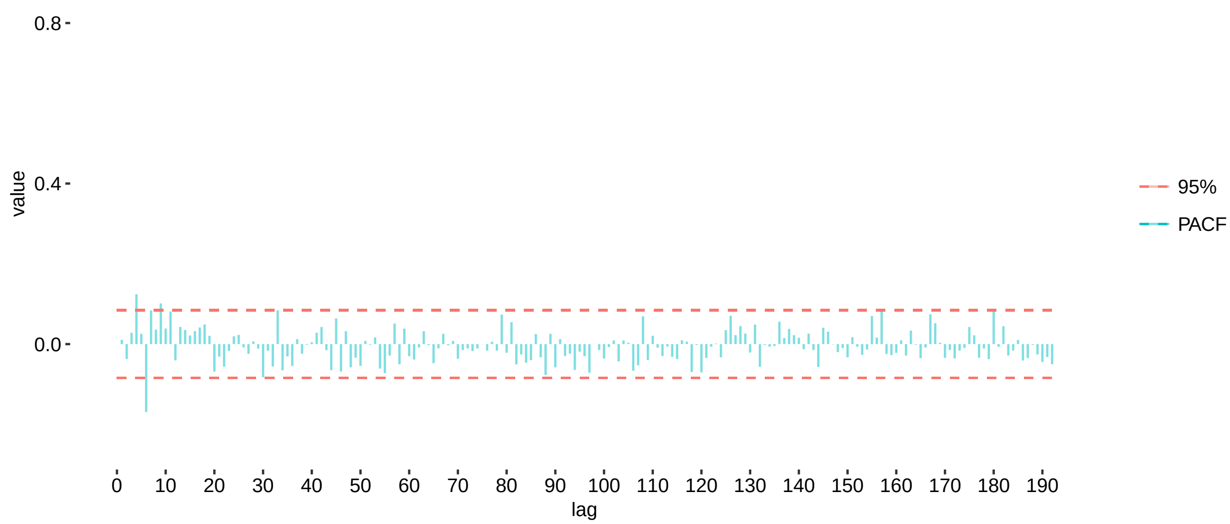

Figure 26: PACF of the residuals.

The ACF (figure 25) and the PACF (figure 26) of the residuals from model $TiTm$ much less correlation in the residuals, which indicates that the $TiTm$ model describes more systematic behaviour than the simple Ti model (ACF figure 18 and PACF figure 19).

Periodogram of the residuals

Figure 27 shows the periodogram of the residuals.

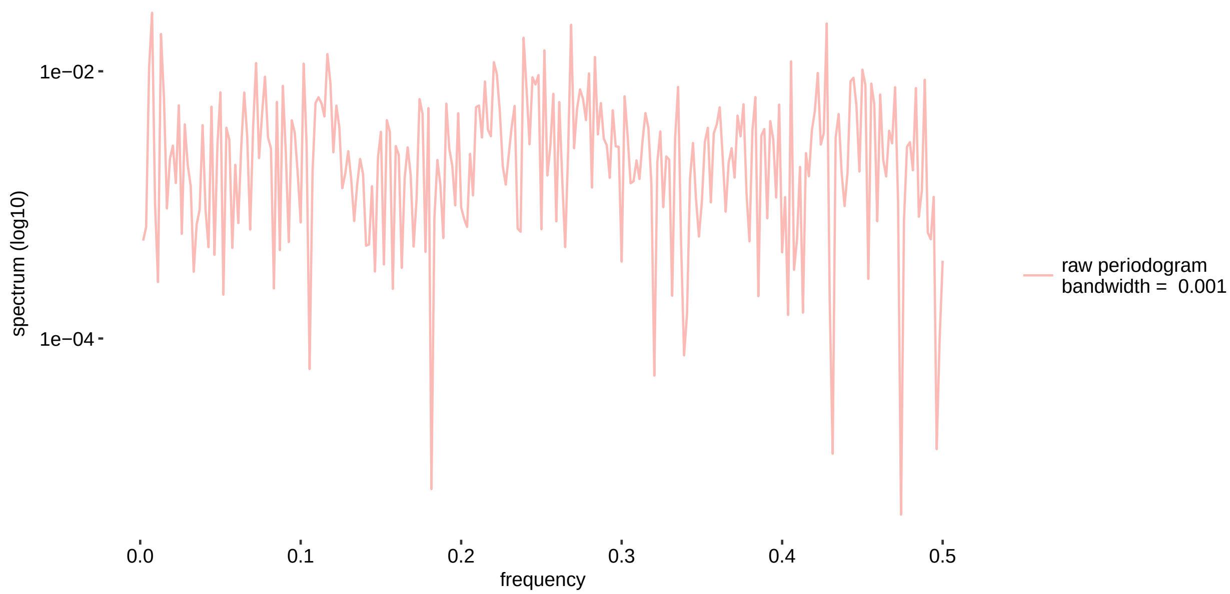

Figure 27: Periodogram of the residuals.

Cumulated periodogram of the residuals

Figure 28 shows the cumulated periodogram of the residuals.

The all the residuals from the one step ahead prediction errors of the new $TiTm$ model are within the 95% confident bands. This plot illustrates a clear improvement from the simple model Ti (figure 21) to the extended model $TiTm$ (figure 28). The residuals are closer to the diagonal which is an identification of white noise.

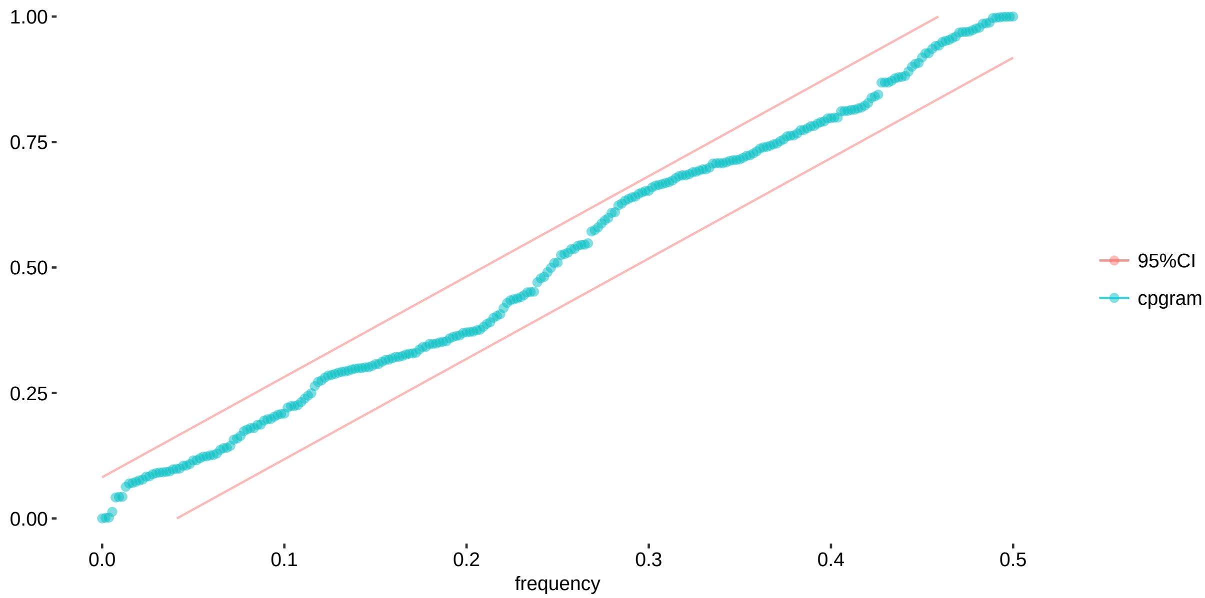

Figure 28: Cumulated periodogram of the residuals.

Combined plot

The one step ahead residuals have been placed above the three sub-plots (figure 15) in figure 29.

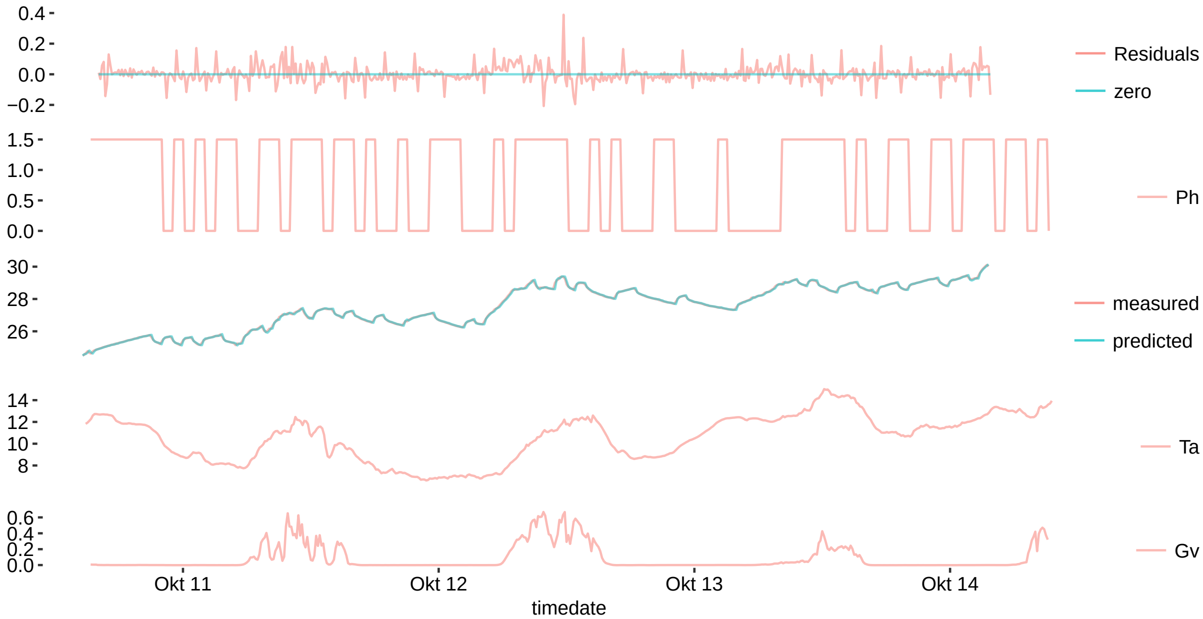

Figure 29: Conbined plot with the residuals, the heater, measured and predicted temperature, ect..

7. Step

There has been performed an log-likelihood-ratio test in step 7 of the two models Ti and $TiTm$.

The p-value of the log-likelihood-ratio test is: $p = 0$, which tells that the $TiTm$ model is significantly better than the Ti model.

Consider the following

- Discuss the white-noise properties of the (one-step ahead) residuals for model T_i.
  - By plotting the distribution, ACF and PACF and the cumulated periodogram of the residuals (figure 17, 18 and 20) it is possible to conclude that there is systematic behaviour left in the residuals.
- What useful information can be obtained from the time series plots of the residuals and the inputs for model T_i?
  - The realization of the residuals in figure 16 shows the first property of the residuals. There is to some extent systematic periodic behaviour left in the residuals.
  - The combined plot (figure 22) of the residuals and the other time series provides an good overview of the system. It is possible to see when the system struggles to predict the next input for the given state of the system.
  - It is possible to see where the one step ahead prediction struggles.
- Discuss the white-noise properties of the one-step ahead residuals for model T_iT_m.

- The residuals for the $TiTm$ model is much closer to white noise. This can be concluded by considering the plot of the ACF and PACF plots and the cumulated periodogram plot respectively in figure 25 and 27.

- What useful information can be obtained from the time series plots of the residuals and inputs for model T_iT_m?
  - The same information as for the Ti model above.
  - It is possible to conclude that the solar radiation has an effect on the one step ahead predictions for both models. This conclusion can easily be seen in both figure 22 and figure 29.
  - It will therefore make sense to include the solar radiation in both models.of the system.

- Based on the likelihood-ratio test is model T_iT_m then to be preferred over model T_i?
  - Yes.
  - The p-value of the log-likelihood-ratio test is: $p = 0$. This tells that the $TiTm$ model is significantly better than the Ti model.

Question 2b

The next improvement is to include the solar radiation Gv in the model, eqn. 6 and still use the same measurement eqn. 4. I have chosen the linear interpolation.

$$
\begin{aligned}
dT_i &= \left(\frac{1}{R_{im}C_i} \left(T_m - T_i\right) + \frac{1}{R_{ia}C_i} \left(T_a - T_i\right) + \frac{pA_w}{C_i} G_v + \frac{1}{C_i} \Phi_h \right) dt + \sigma_i d\omega_i \\
dT_m &= \left(\frac{1}{R_{im}C_m} \left(T_a - T_m\right) + \frac{(1-p)\,A_w}{C_m} G_v \right) dt + \sigma_m d\omega_m
\end{aligned}
\tag{6}
$$

By implemented the solar radiation (Gv) in the system, two additional parameters have been introduced:

- The linear interpretation p identifies the ratio in which the inside air temperature or the interior thermal medium will absorb the solar radiation Gv.
- $Aw = 7.5 + 4.8 = 12.3[m^2]$ which is the effective window area of the building.

The new models has been estimated and its output is as follows:

```
## Coefficients:
##        Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.4489e+01  3.6668e-02  6.6786e+02  0.0000e+00  4.3000e-05     4e-04
## Tm0  2.4477e+01  9.8163e-02  2.4935e+02  0.0000e+00 -2.9409e-05     4e-04
## Aw   1.2300e+01          NA          NA          NA          NA        NA
## Ci   9.8644e-01  6.8555e-02  1.4389e+01  0.0000e+00 -1.1716e-06     0e+00
## Cm   1.8582e+01  2.6874e+00  6.9145e+00  1.3832e-11 -2.9918e-06     0e+00
## e11 -6.8293e+00  1.0673e-01 -6.3988e+01  0.0000e+00 -1.5967e-05     0e+00
## p    1.8549e-01  2.9369e-02  6.3157e+00  5.7724e-10 -1.4033e-06     0e+00
## p11 -9.1487e+00  2.2310e-01 -4.1007e+01  0.0000e+00  4.9089e-05     0e+00
## p22 -1.9680e+00  8.6330e-02 -2.2796e+01  0.0000e+00 -1.6512e-06     0e+00
## Ria  1.9448e+01  6.6528e+00  2.9233e+00  3.6122e-03  4.7629e-07     0e+00
## Rim  2.3250e-01  1.3792e-02  1.6857e+01  0.0000e+00 -2.2276e-06     0e+00
##
## Correlation of coefficients:
##     Ti0   Tm0   Ci    Cm    e11   p     p11   p22   Ria
## Tm0  0.14
## Ci   0.02 -0.05
## Cm  -0.03 -0.21  0.25
## e11 -0.01  0.00 -0.37 -0.09
## p   -0.07  0.00  0.14  0.07  0.00
## p11  0.04  0.01 -0.31 -0.32  0.17  0.04
## p22  0.05 -0.03  0.64  0.12 -0.47 -0.07 -0.20
## Ria -0.01 -0.53  0.11  0.40 -0.02  0.04 -0.04  0.03
## Rim -0.02 -0.02 -0.18  0.02  0.44 -0.19 -0.06 -0.24 -0.02
##
## [1] "loglikelihood = 801.930993840359"
```

Consider . . .

It has been decided only to include a plot of the squared residuals against Gv and comparing the likelihoods of the two models, instead of an complete residual analysis as above.

- Findings
  - The parameter Aw is given as a constant, which causes the 'NA''s in the summary output. The value of Aw is 12.3.
  - The parameter p is estimated to 0.185, which tells that the 19% of the solar radiation will be absorbed in the inside air temperature. The remaining 81% will be absorbed by the interior walls and furnitures.
  - The two realizations in figure 30 illustrates the squared residuals against Gv of the model which not include the solar radiation and the model which do include the solar radiation, eqn. 5 and eqn. 6 respectively.

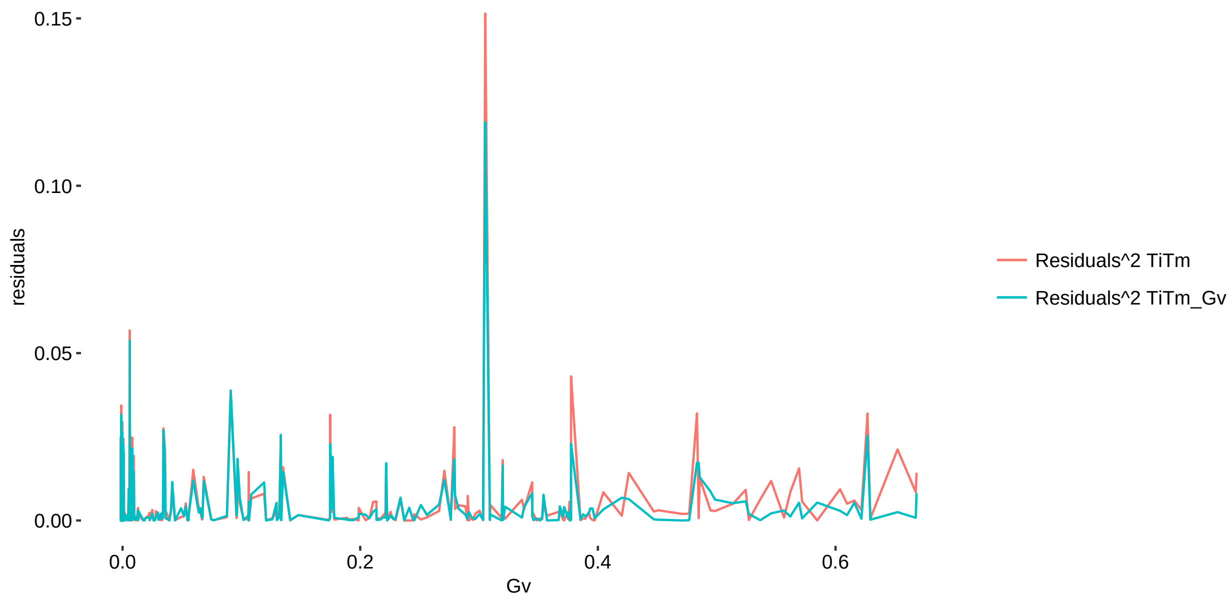

Figure 30: Ralizations of squared residuals as function of Gv.

The amplitudes of the squared residuals of the model (eqn. 6) which include Gv is smaller than the residuals from the model (eqn. 5) which exclude Gv.
The summed squared residuals, of the model (eqn. 6), is 11.289% lower than the model (eqn. 5), which identicates a better model. This can also be seen their log-likelihoods.

- Likelihood-ratio test
  - The model (eqn. 6) has increased the log-likelihoods of 3.857% from the model (eqn. 5).
  - Likelihood-ratio test
    A likelihood-ratio test has been performed in order to validate weather the model (eqn. 6) is better or worse than the model (eqn. 5).

    The p-value from the test is: $p = 0$ which tells that the model, which include the solar radiation, is significant better performing than the model which not include the solar radiation.
- Conclusion: The solar radiation should be included in the model.

Question 2c

I decided to use the same model (eqn. 6) and change the input data. The difference between privous two qestions and this is the way to select columns. Examples on selecting correct columns from the csv below:

- Question 2a and 2b: *data ← data[,c(1,2,4,5,6,3)]*
- Question 2c: *data ← data[,c(1,3,4,5,7,2)]*
- Column names in both cases: *names(data) ← c("timedate","yTi","Ta","Gv","Ph","Tn")*

Estimating the same model has been done after reading the data for the West room. I simply copyied the code chunk from Question 2b. The output from estimated model for the West room is given below.

```
## Coefficients:
##         Estimate  Std. Error     t value    Pr(>|t|)     dF/dPar dPen/dPar
## Ti0  2.7380e+01  5.4465e-02  5.0270e+02  0.0000e+00 -2.3262e-05     7e-04
## Tm0  2.6886e+01  3.5651e-01  7.5415e+01  0.0000e+00  6.2791e-05     6e-04
## Aw   1.2300e+01          NA          NA          NA          NA        NA
## Ci   7.0126e-01  5.4524e-02  1.2862e+01  0.0000e+00  1.1018e-05     0e+00
## Cm   1.6873e+01  7.2707e+00  2.3206e+00  2.0685e-02 -2.5728e-06     0e+00
## e11 -6.3590e+00  1.5539e-01 -4.0922e+01  0.0000e+00 -1.5480e-05     0e+00
## p    1.5783e-01  3.1869e-02  4.9523e+00  9.9102e-07  6.4493e-06     0e+00
## p11 -8.9078e+00  2.8365e-01 -3.1404e+01  0.0000e+00  2.0439e-04     0e+00
## p22 -1.0290e+00  9.2965e-02 -1.1069e+01  0.0000e+00  1.1677e-05     0e+00
## Ria  3.0916e+01  3.9330e+01  7.8605e-01  4.3219e-01  1.5932e-06     0e+00
## Rim  4.6024e-01  3.9437e-02  1.1670e+01  0.0000e+00  4.2493e-06     0e+00
##
## Correlation of coefficients:
##      Ti0   Tm0   Ci    Cm    e11   p     p11   p22   Ria
## Tm0  0.17
## Ci   0.00 -0.12
## Cm  -0.07 -0.54  0.29
## e11 -0.02  0.03 -0.65 -0.15
## p   -0.06 -0.12 -0.03  0.02  0.12
## p11  0.19  0.08 -0.20 -0.26  0.14 -0.01
## p22 -0.01  0.07  0.76  0.07 -0.55 -0.14 -0.24
## Ria -0.02 -0.83  0.19  0.64 -0.10  0.10  0.05 -0.06
## Rim  0.00  0.10 -0.57 -0.22  0.62 -0.08  0.11 -0.32 -0.23
##
## [1] "loglikelihood = 533.898842680385"
```

One step ahead prediction errors

We are interested in the residuals of the one step ahead prediction as before, but we will only consider the realization of the residuals and the cumulated periodogram of the residuals.

Time Series of the residuals

Figure 31 shows a time series plot of residuals from the one step ahead predictions. It is possible to see some seasonal (season of one day) trend within the residuals which identicates there is systematic behaviour left in the residuals.

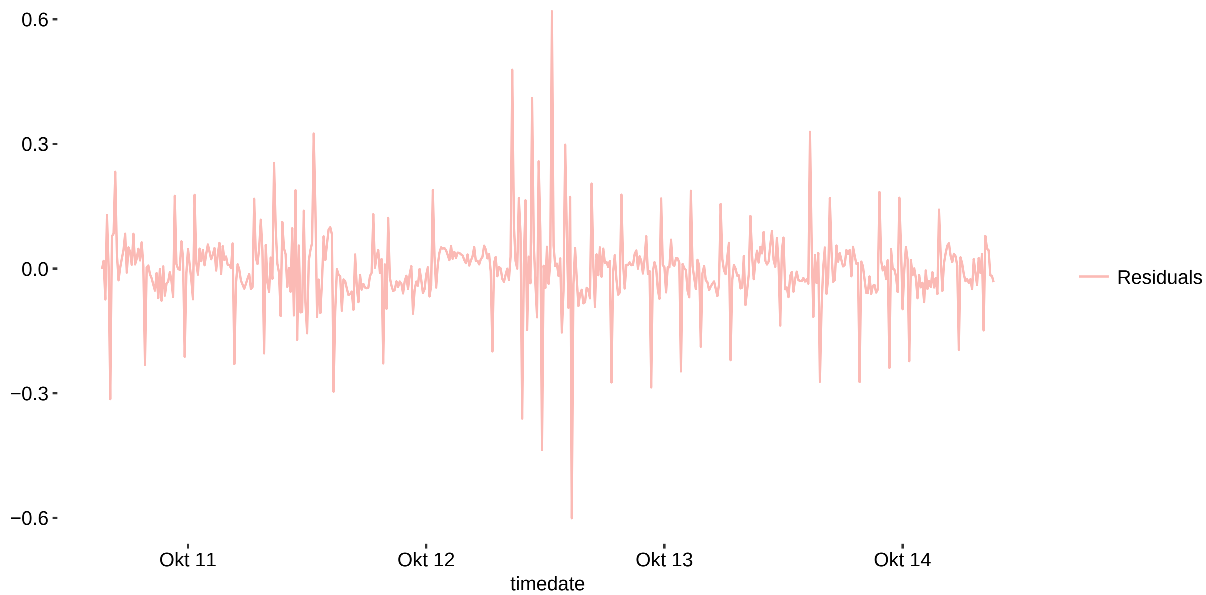

Figure 31: Time serie plot of the residuals.

Cumulated periodogram of the residuals

Figure 32 shows the cumulated periodogram of the residuals. The residuals are within the confident bands, but they are wobbling around the diagonal, which can support the fact about a seasonal trend.

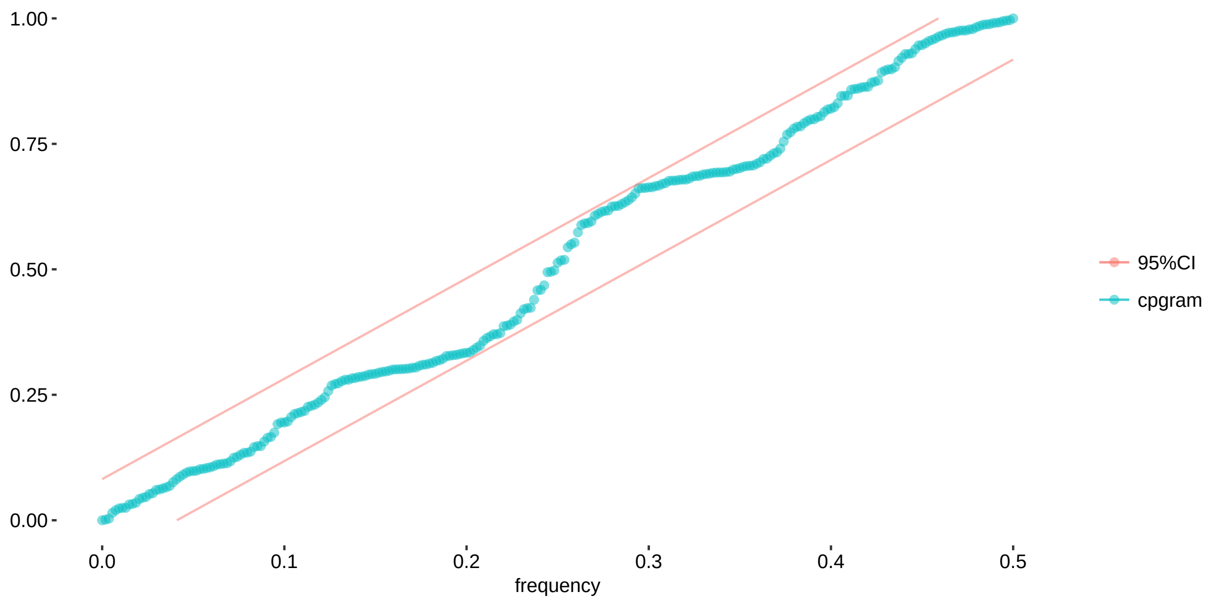

Figure 32: Cumulated periodogram of the residuals.

It is assumed that the model performs well enough. As future work the modelling process of the West room could be done in same steps as for the East room.

Estimated heat capacities

The estimated head capaities for the East room and the West room are given in eqn. 7 and eqn. 8 resoectively.

$$\begin{aligned} C_{east,i} &= 0.986 \\ C_{east,m} &= 18.582 \end{aligned} \tag{7}$$

$$\begin{aligned} C_{west,i} &= 0.701 \\ C_{west,m} &= 16.873 \end{aligned} \tag{8}$$

The specific heat property of air is 1.005 $\left[\frac{kJ}{kg\cdot{}^\circ C}\right]$ and the specific heat capacity for conrete (I assume the same heat capacity for flagstone) is 0.88 $\left[\frac{kJ}{kg\cdot{}^\circ C}\right]$[2].

- The estimates of C_i diviates with 28.91% from the east room to the west room. The east room ($C_{east,i}$) diviates with 1.881% from its specific heat capcity and the west room ($C_{west,i}$) diviates with 43.313% from its specific heat capcity.
- The estimates of C_m diviates with 9.198% from the east room to the west room.

Let us assume that the estimated value of $C_{west,m}$ is a correct estimate of C_m. If we then subtract $C_{east,m}$ from $C_{west,m}$, then this should represent the combined heat capcaity contribution of the concrete and flagstones: $C_{east,m} - C_{west,m} = 1.709$.

This value is ≈ 2 the value of the as specific heat capacity for concrete/flagstone. This can be caused one of the following, either the heat capacity for flagstone is much different than the heat capacity for concrete or the estimates of $C_{east,m}$ and $C_{west,m}$ are far from correct.

I do not have competencies within the field of constructions and those kind of materials. I will state the the estimated values of $C_{east,m}$ and $C_{west,m}$ reasonable, according to the outputs from each estimation on page and page **??**. TODO

Question 2d

This question is not considered due to other compulsory assignments.

[2]url: https://www.engineeringtoolbox.com/specific-heat-capacity-d_391.html