Advanced Time Series Analysis: Computer Exercise 3

Anders Launer Bæk (s160159)

22 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 relizations 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
```

```r
        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 relizations 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 $\sigma = [0.0, 0.1, 0.2, 0.3, 0.4]$.

$\sigma = 0.00$

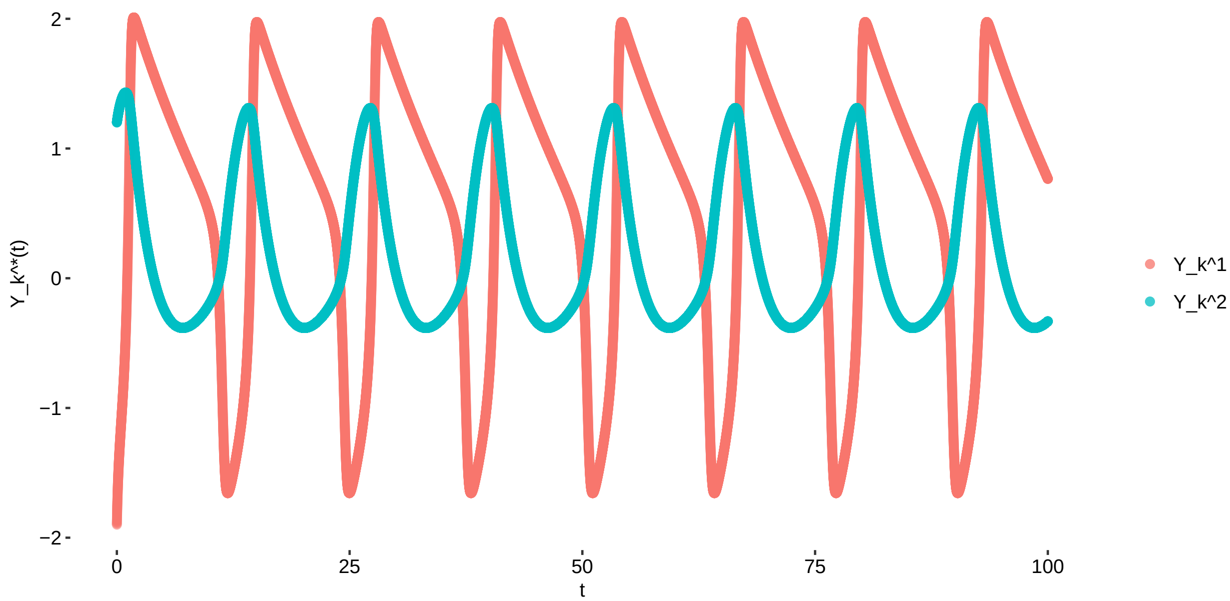


Figure 1: Plot of the simulation realizations with sigma = 0.0.

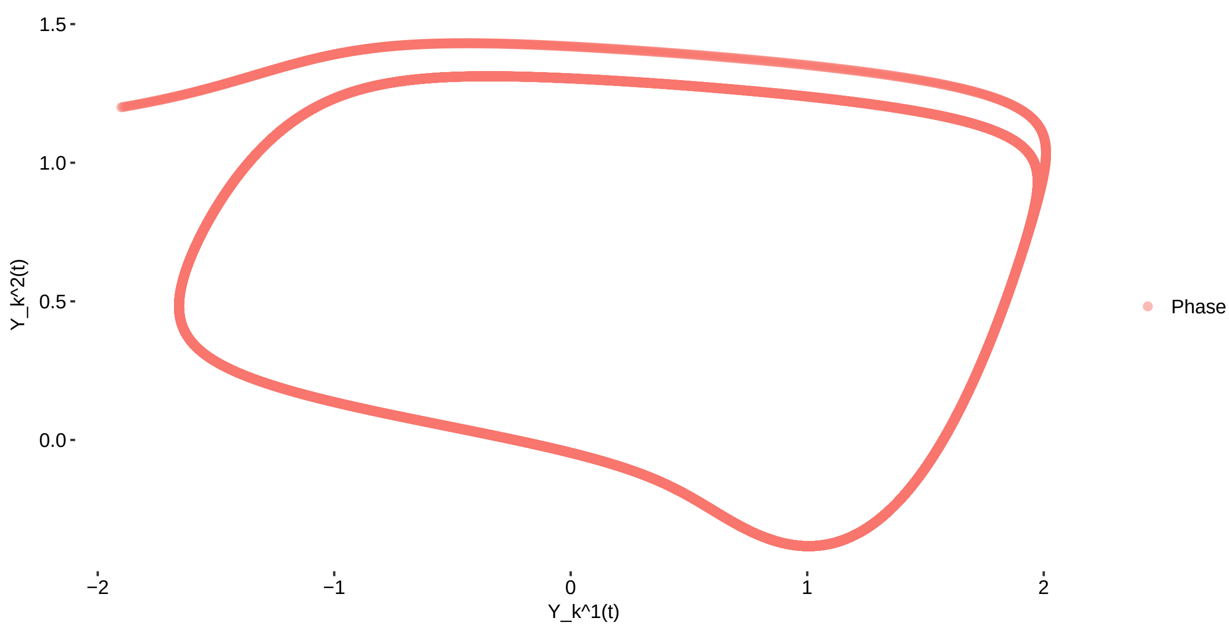


Figure 2: Phase plot of the simulation with sigma = 0.0.

$\sigma = 0.0$ 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 intial parameters of Y_0^1 and Y_0^2.

$\sigma = 0.10$

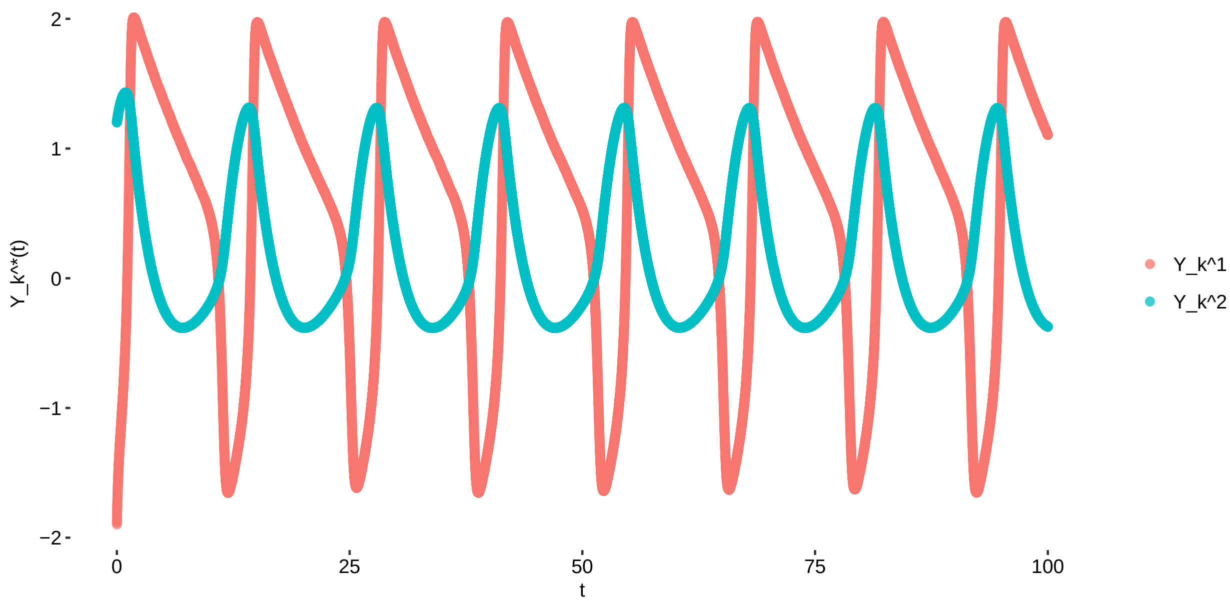


Figure 3: Plot of the simulation realizations with sigma = 0.10.

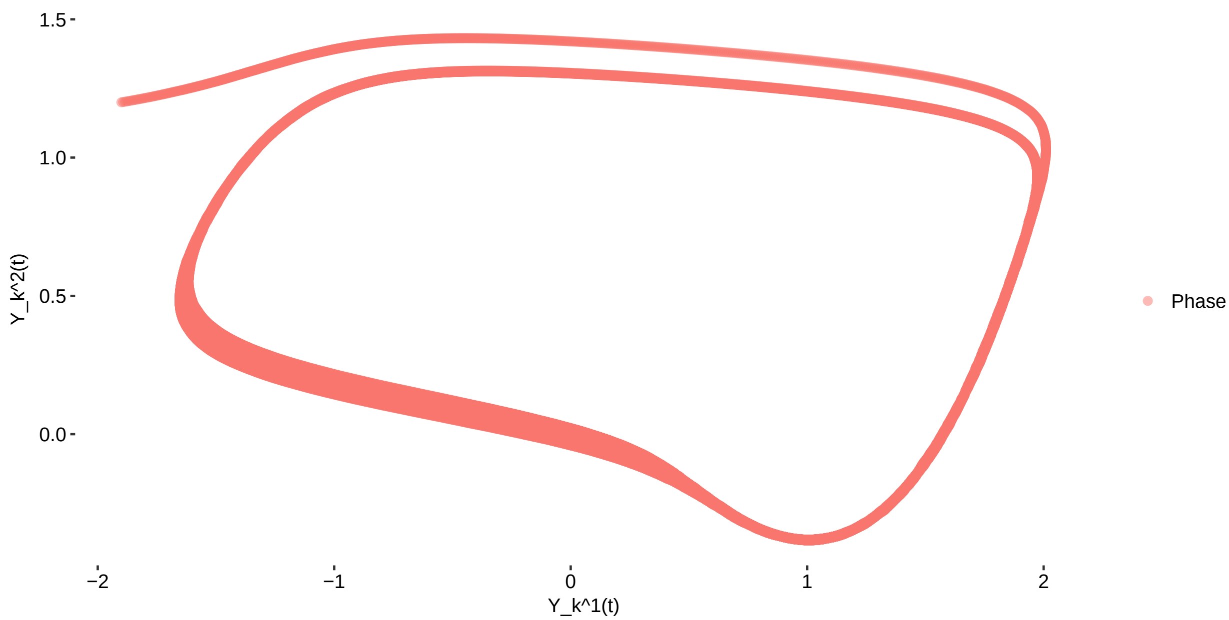


Figure 4: Phase plot of the simulation with sigma = 0.10.

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 $[-1.75; 0.5]$.

$\sigma = 0.20$

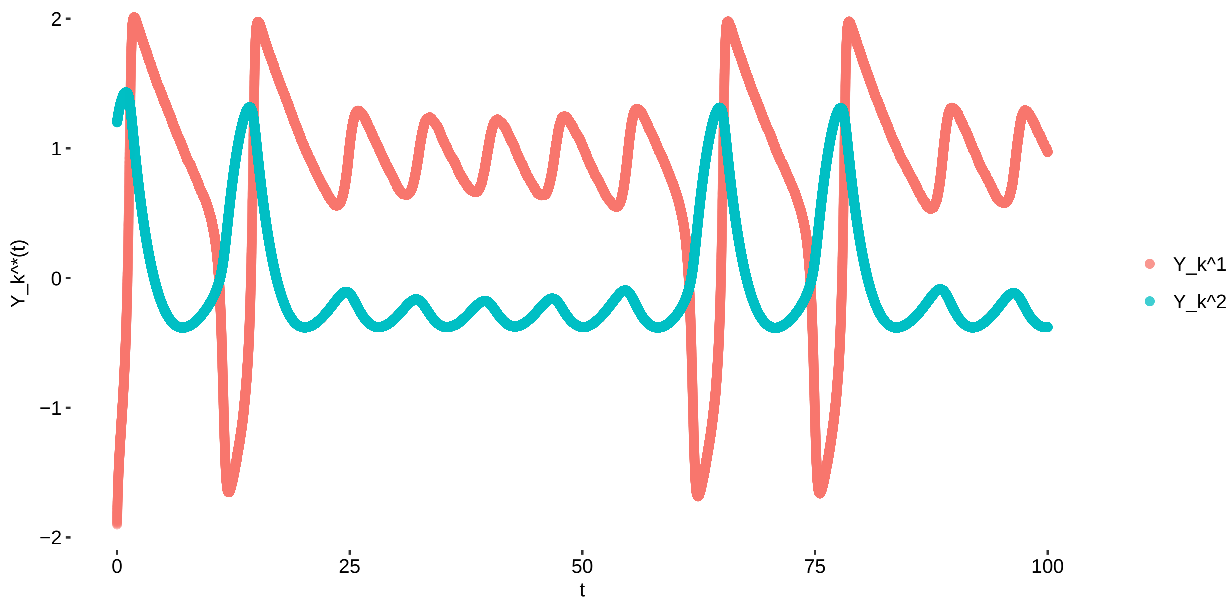


Figure 5: Plot of the simulation realizations with sigma = 0.20.

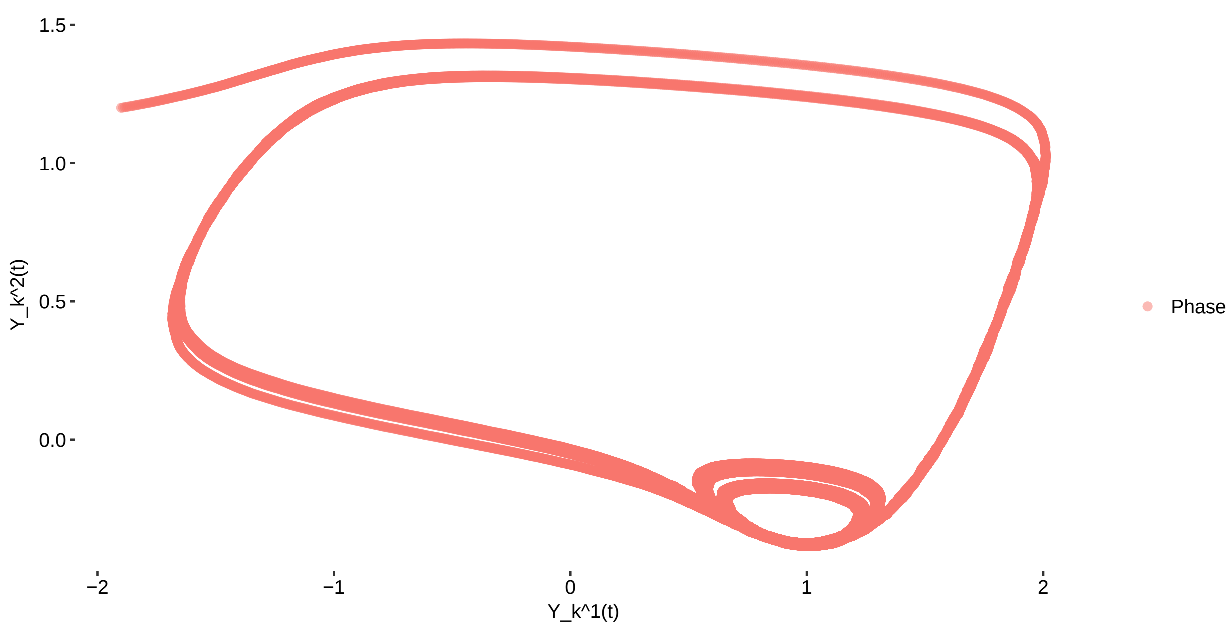


Figure 6: Phase plot of the simulation with sigma = 0.20.

Changing σ to 0.20 does make an effect in both the realizations and in the phase plot. The time sereis still needs ≈ 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 agian. The cahnge in σ from 0.10 to 0.20 increase the effect of the wiener process.

$\sigma = 0.30$

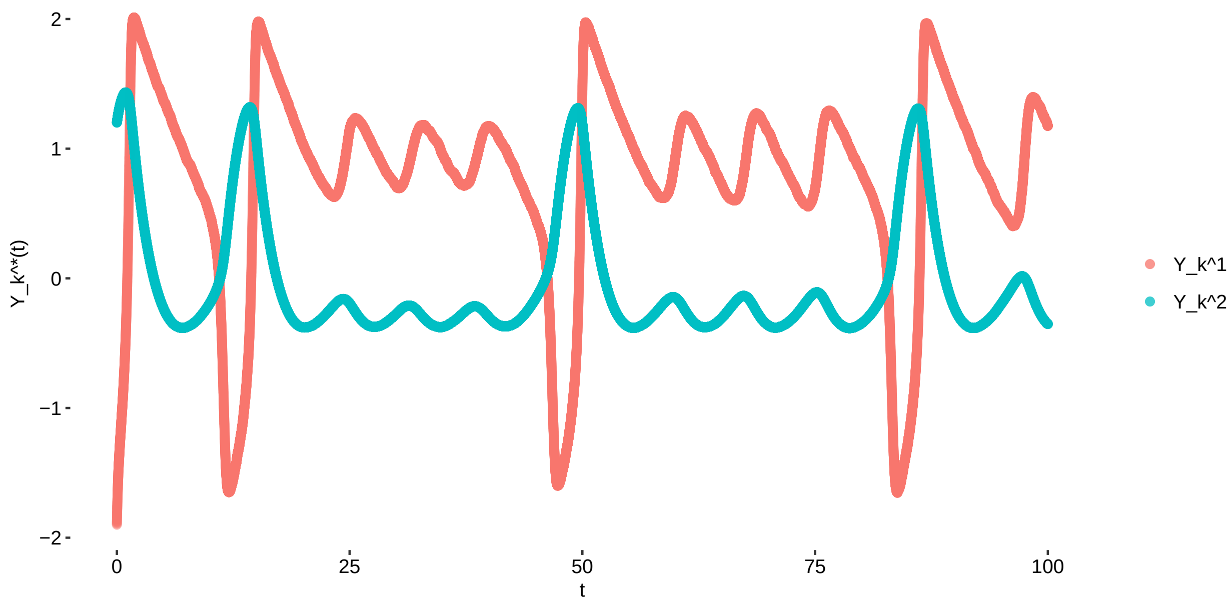


Figure 7: Plot of the simulation realizations with sigma = 0.30.

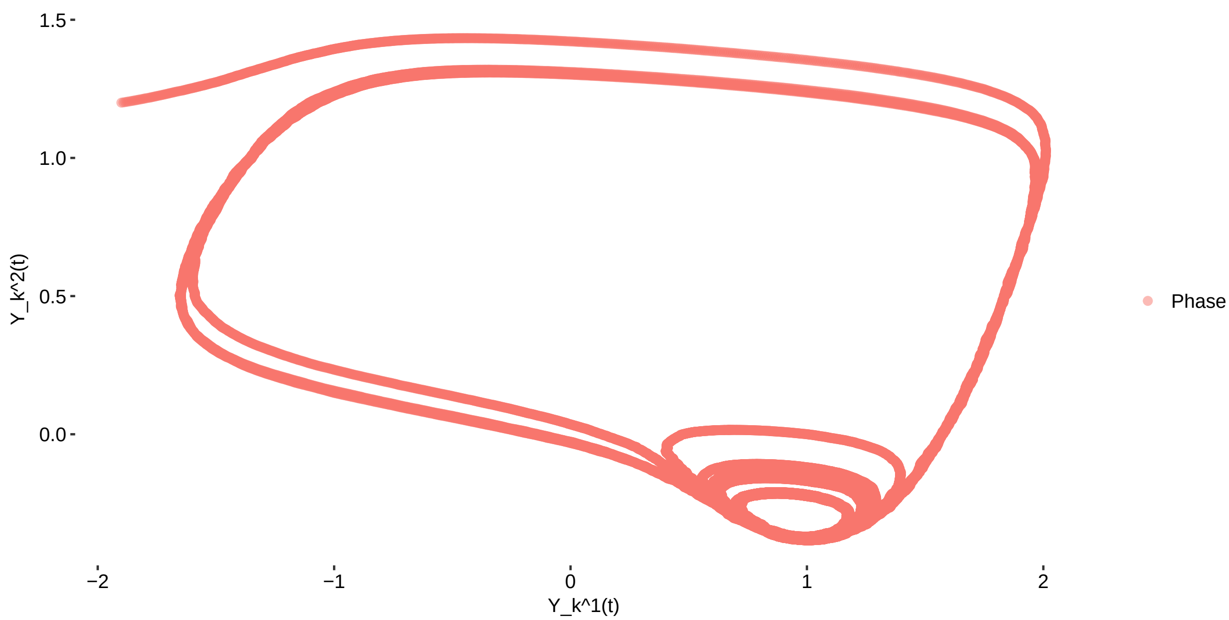


Figure 8: Phase plot of the simulation with sigma = 0.30.

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

$\sigma = 0.40$

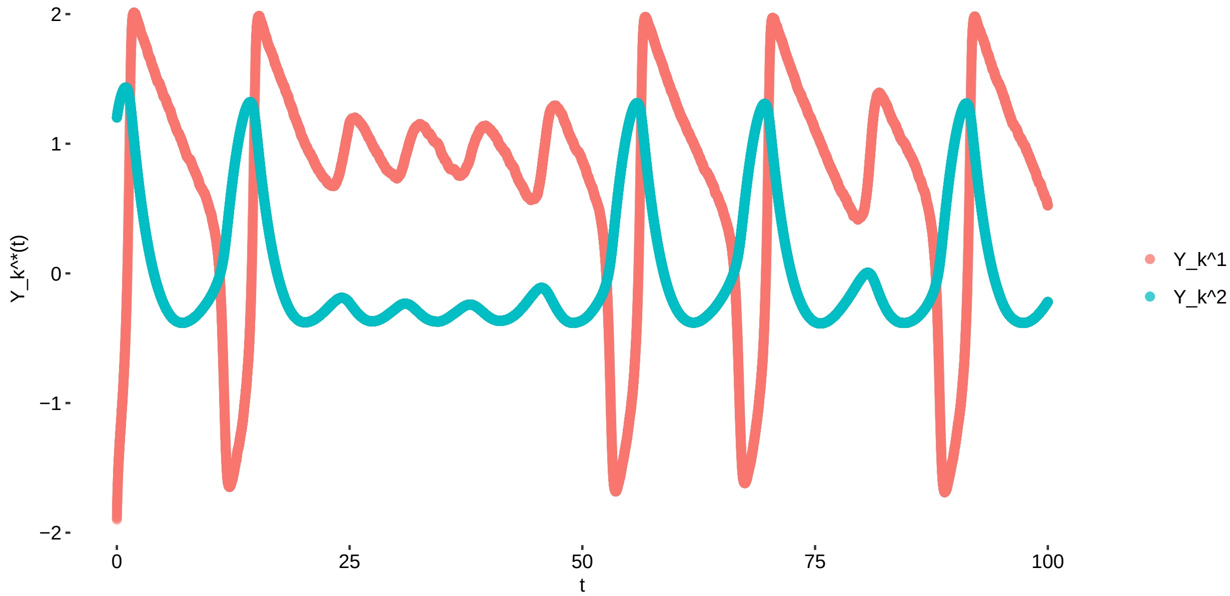


Figure 9: Plot of the simulation realizations with sigma = 0.40.

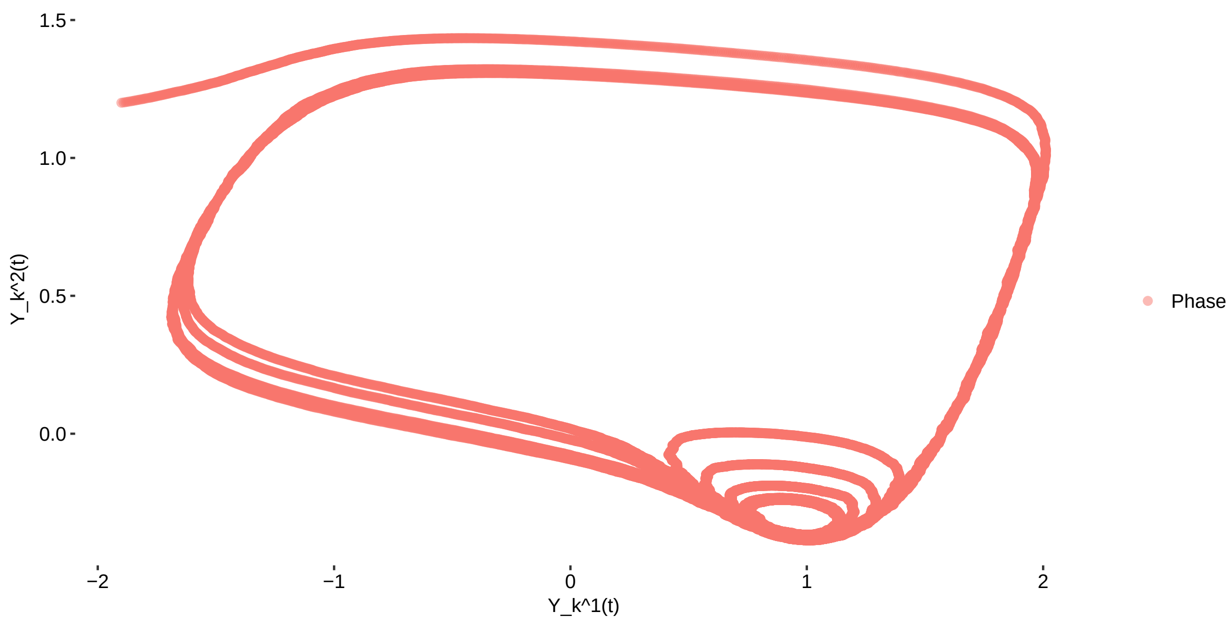


Figure 10: Phase plot of the simulation with sigma = 0.40.

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 influce of the amplitude of the time series. The amplitudes are more similar to the normal pattern when the abnormal pattern is short. For longer peroid of abnormal pattern entail smaller amplitudes. The smaller amplitudes creates more cycles in the lower rigth 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 nnumber of cycles in the abnormal pattern. Those effecting the "number of cycles" in the abnormal pattern where greater values of σ ential 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: $\sigma = [0.0, 0.1, 0.2, 0.3, 0.4]$.

$\sigma = 0.10$

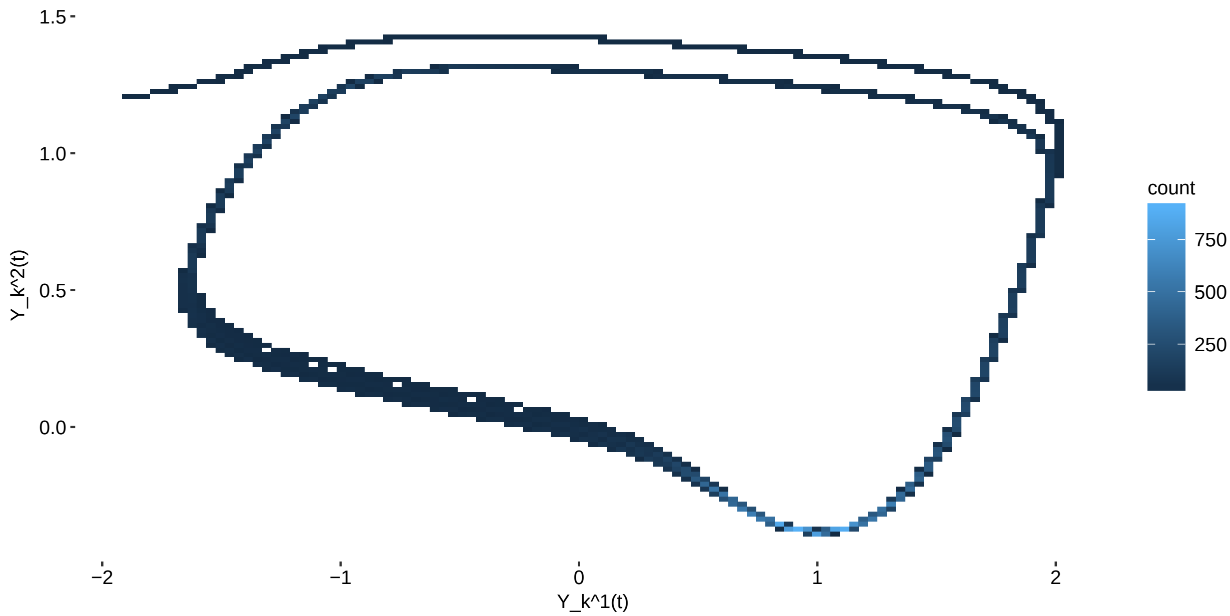


Figure 11: Phase plot of the simulation with sigma = 0.10.

$\sigma = 0.20$

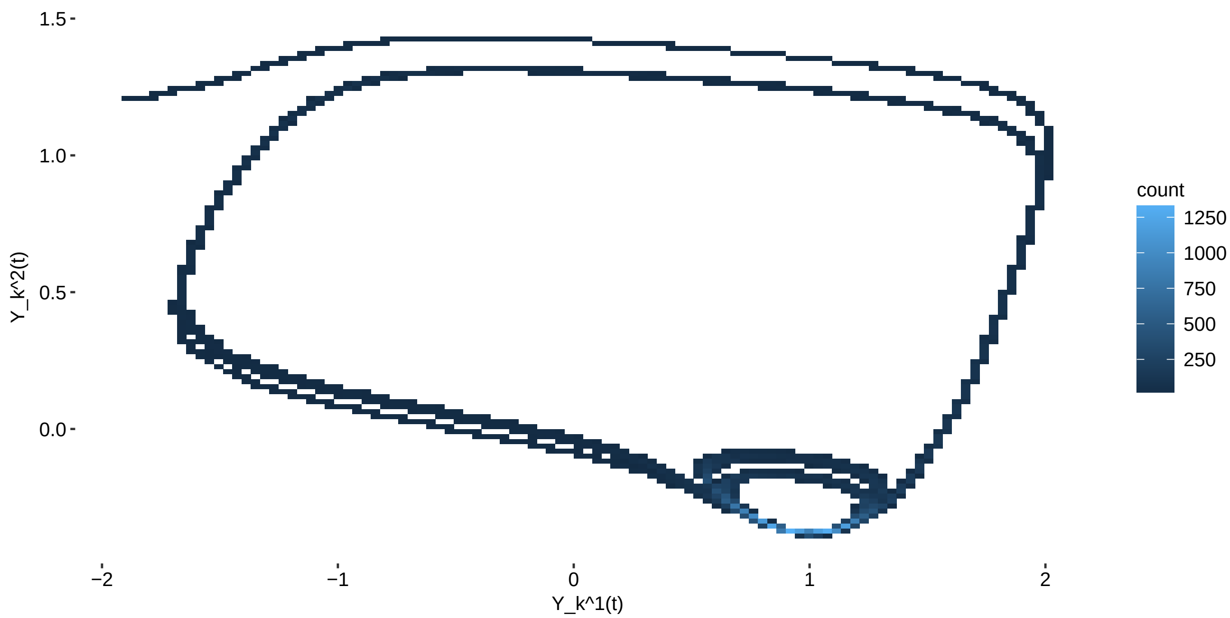


Figure 12: Phase plot of the simulation with sigma = 0.20.

$\sigma = 0.30$

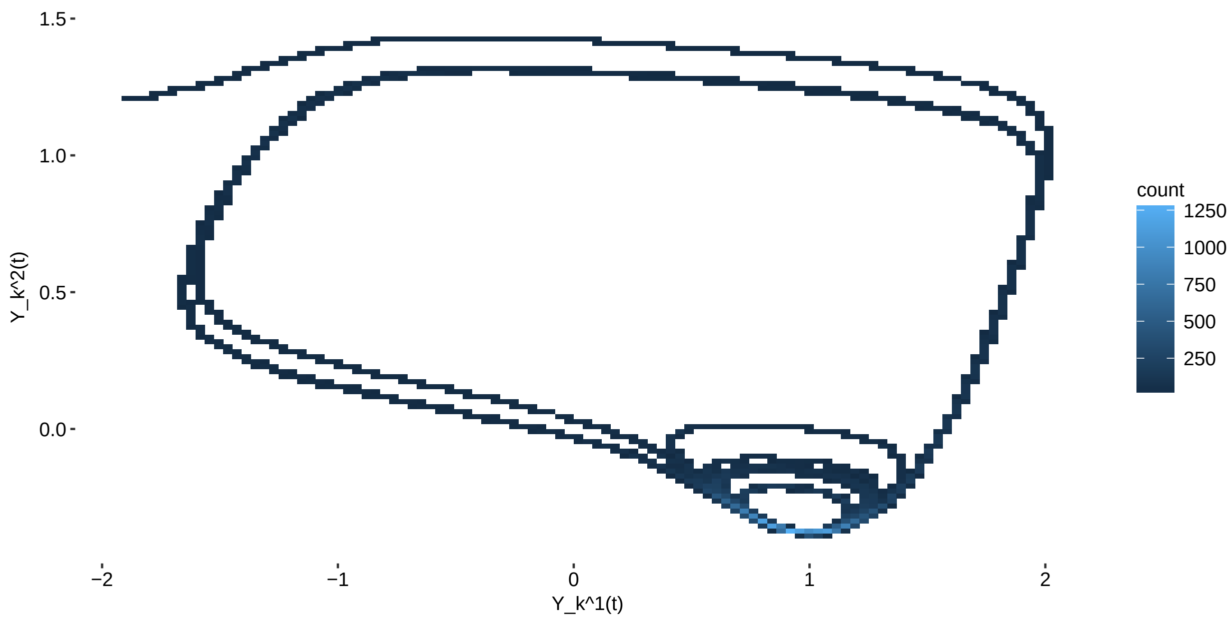


Figure 13: Phase plot of the simulation with sigma = 0.30.

$\sigma = 0.40$

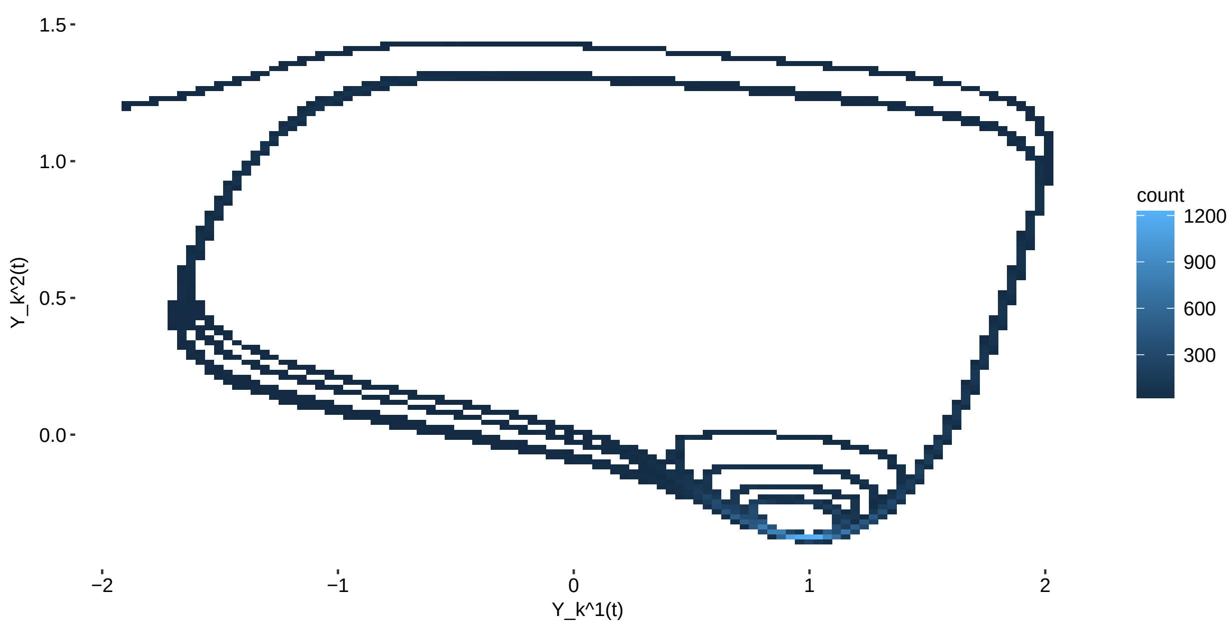


Figure 14: Phase plot of the simulation with sigma = 0.40.

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 $[°C]$.
- TiW The indoor air temperature of the West room $[°C]$.
- Ta The ambient temperature $[°C]$.
- 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 perid is $10\,[min]$.

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 subplots:

- A step seqence 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 solor radiation, on surface facing south, increases the temperatures and the pattern for the heater in figure 15.

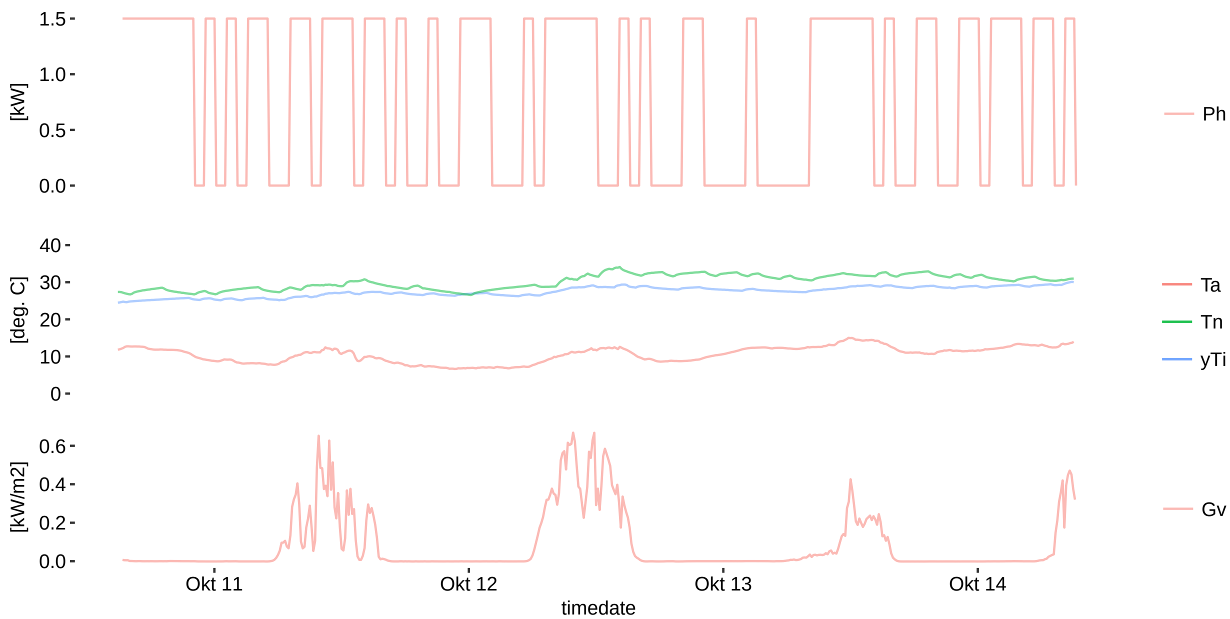


Figure 15: Plot of the simulation realizations with sigma = 0.0.

2. Step

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

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

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

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

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 agian.

```
## 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 inital upper limit value for Ria.

By fixing the initiallimit issues also increased the loglikelihood significantly and decreased the correlations among the parameters which is desirable.

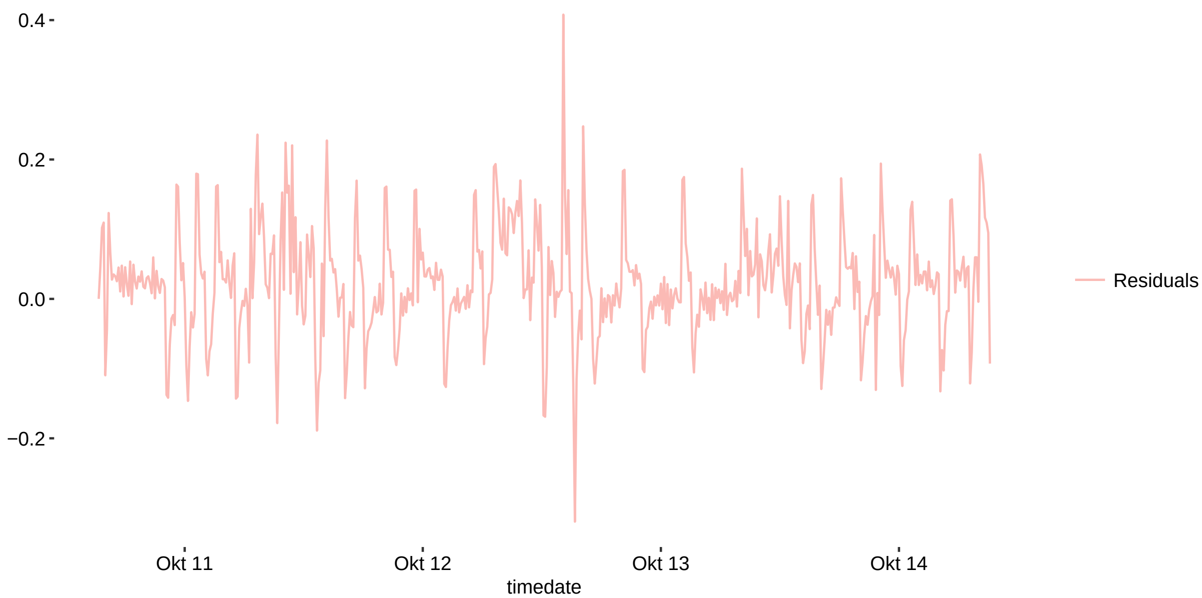
[1] P.15 http://ctsm.info/pdfs/ctsmr-reference.pdf

4. Step

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

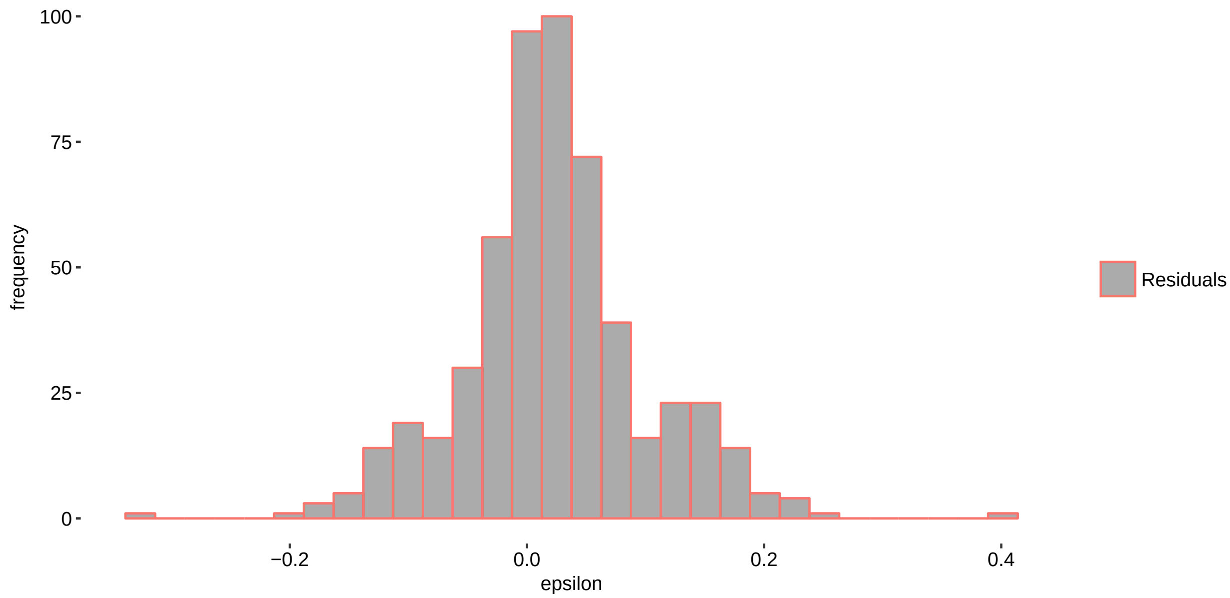
```
# Calculate the one-step predictions of the state
pred <- predict(fit.Ti)
# Extract the estimated value of yTi
data$yTiHat <- pred[[1]]$output$pred$yTi
# Calculate the residuals and add them to the data frame
data$residuals <- data$yTi - data$yTiHat
```

Time Series of the residuals

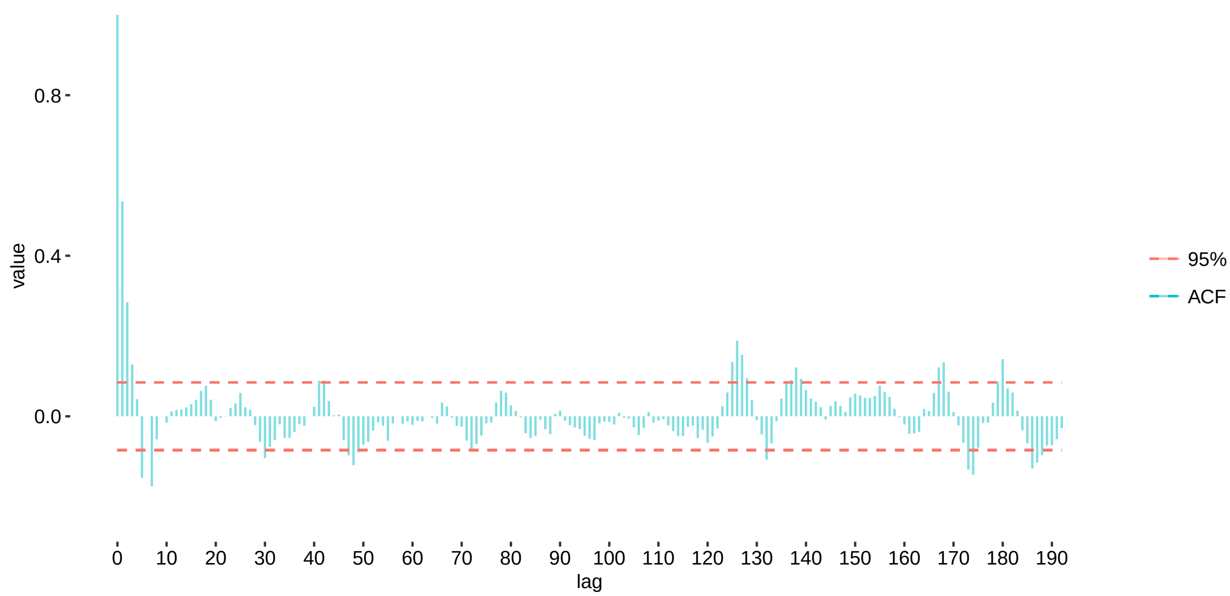


Distribution of the residuals

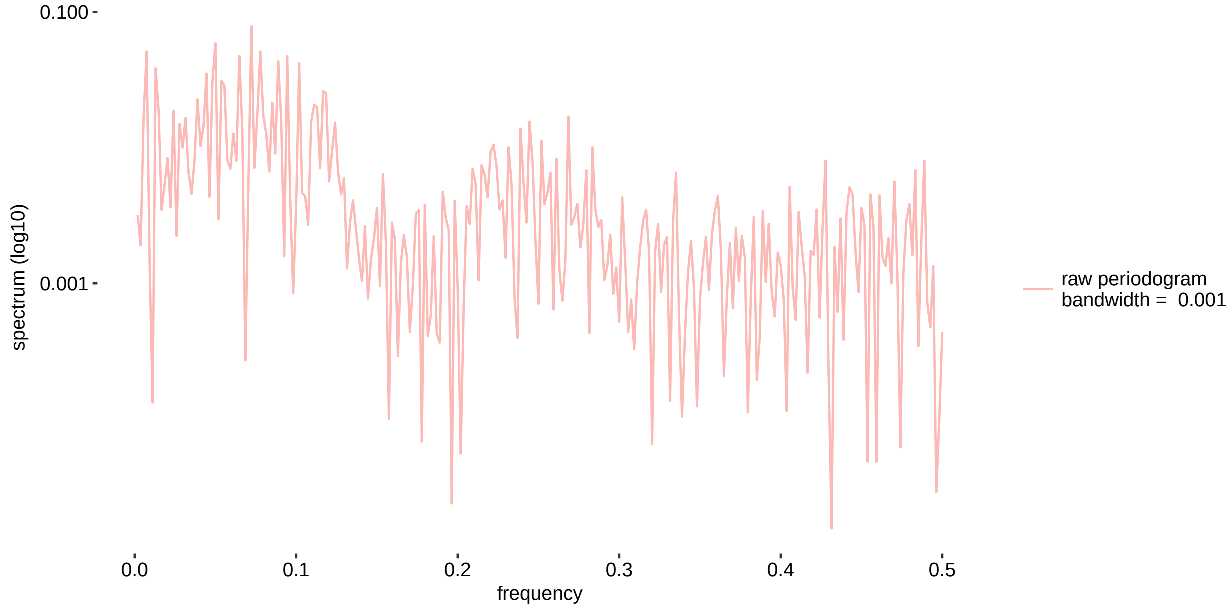
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



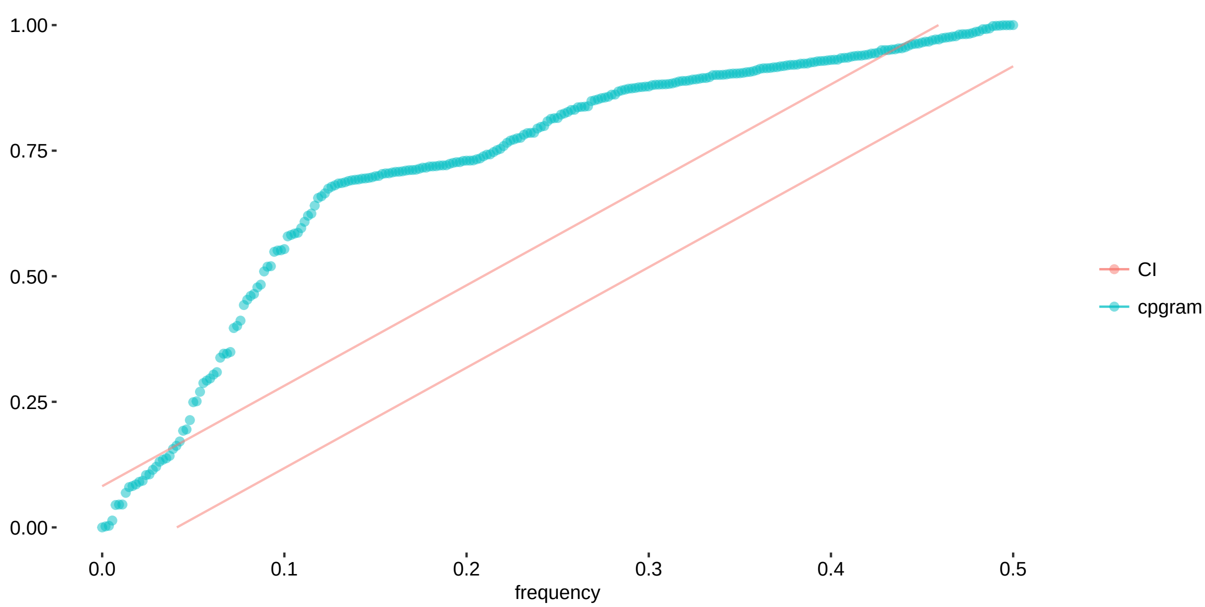
ACF of the residuals



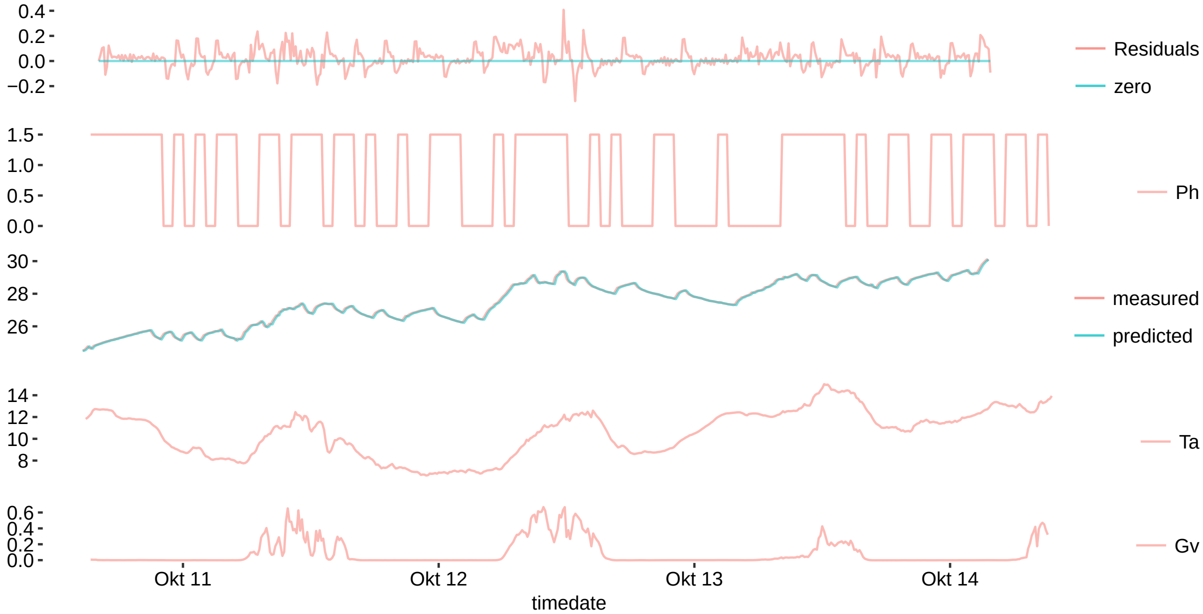
Periodogram of the residuals



Commulated Periodogram of the residuals



combined



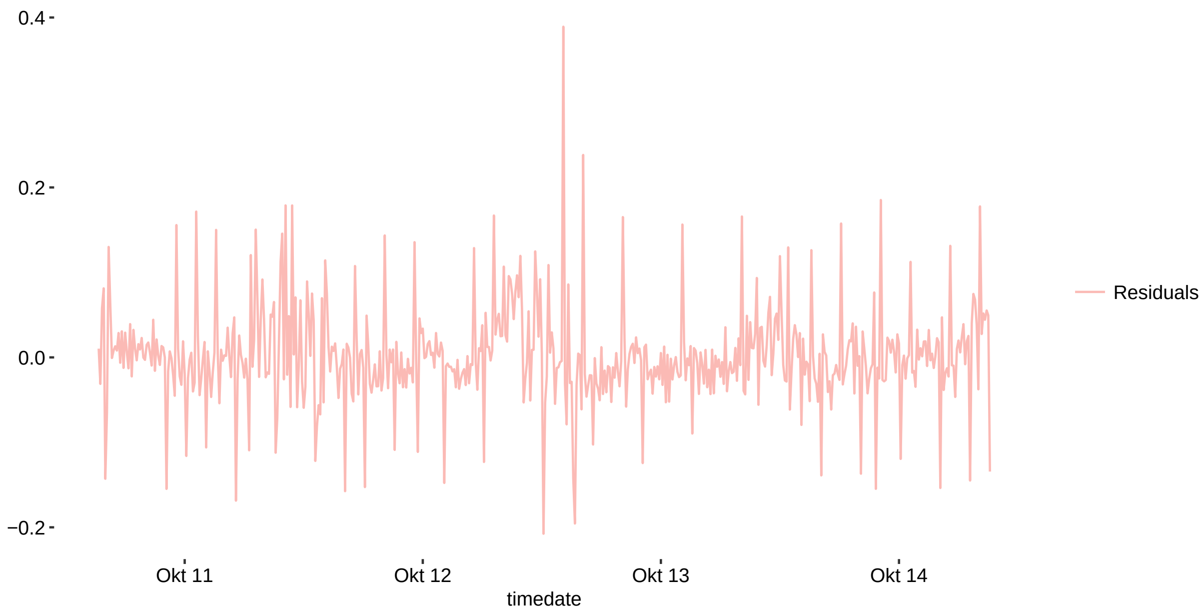
5. Step

figure **??** new

```
## 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"
```

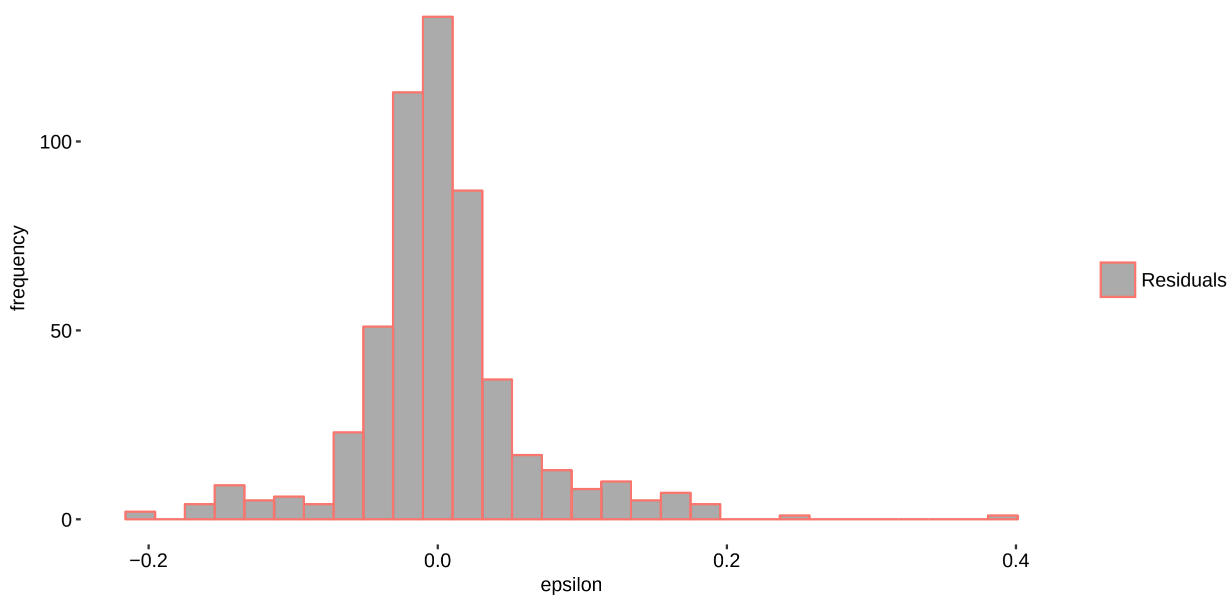
6. Step

Time Series of the residuals

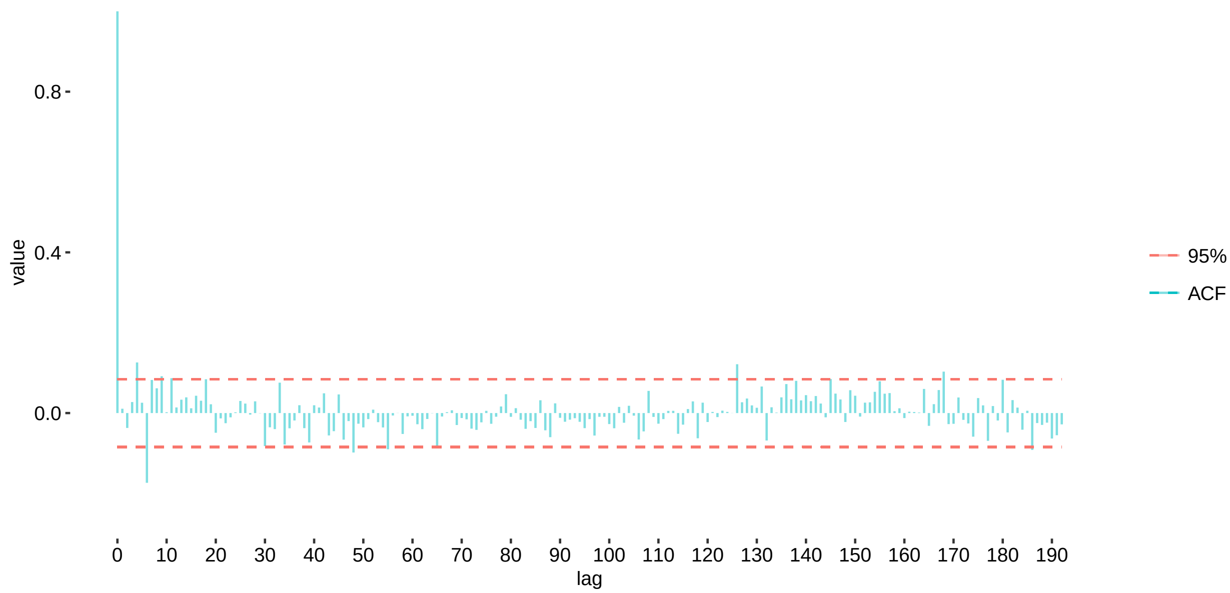


Distribution of the residuals

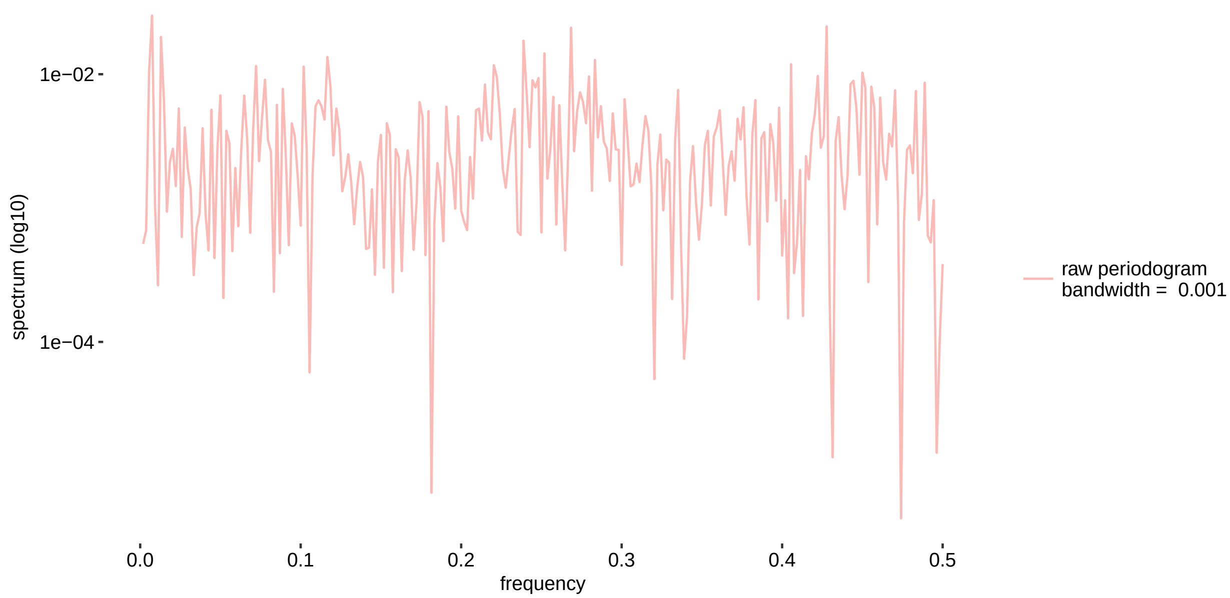
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



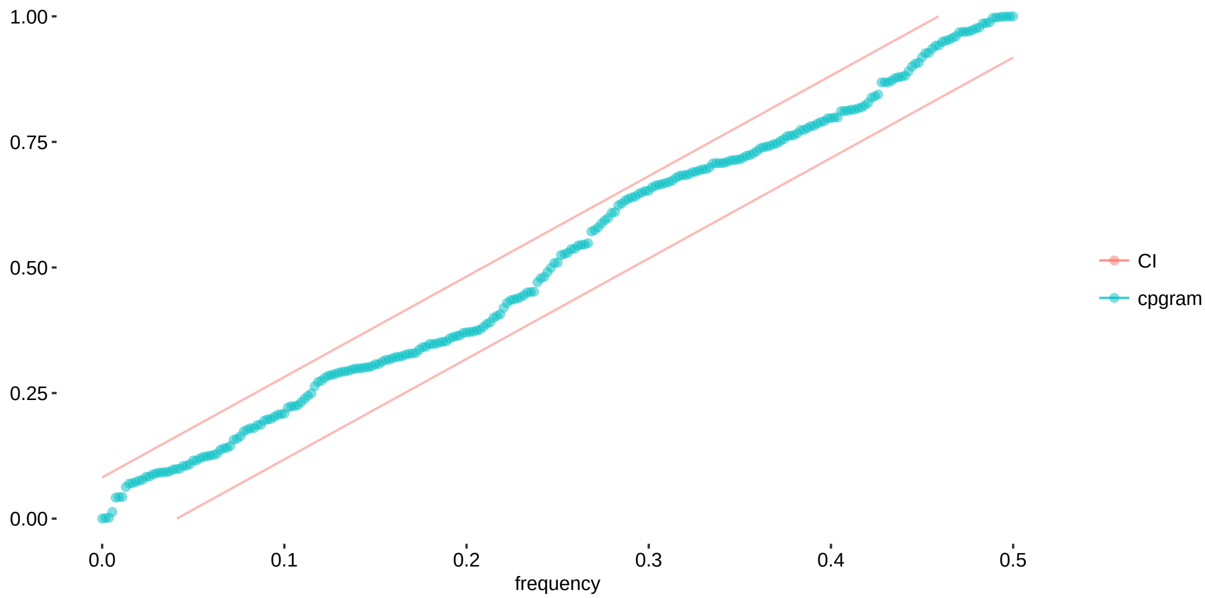
ACF of the residuals



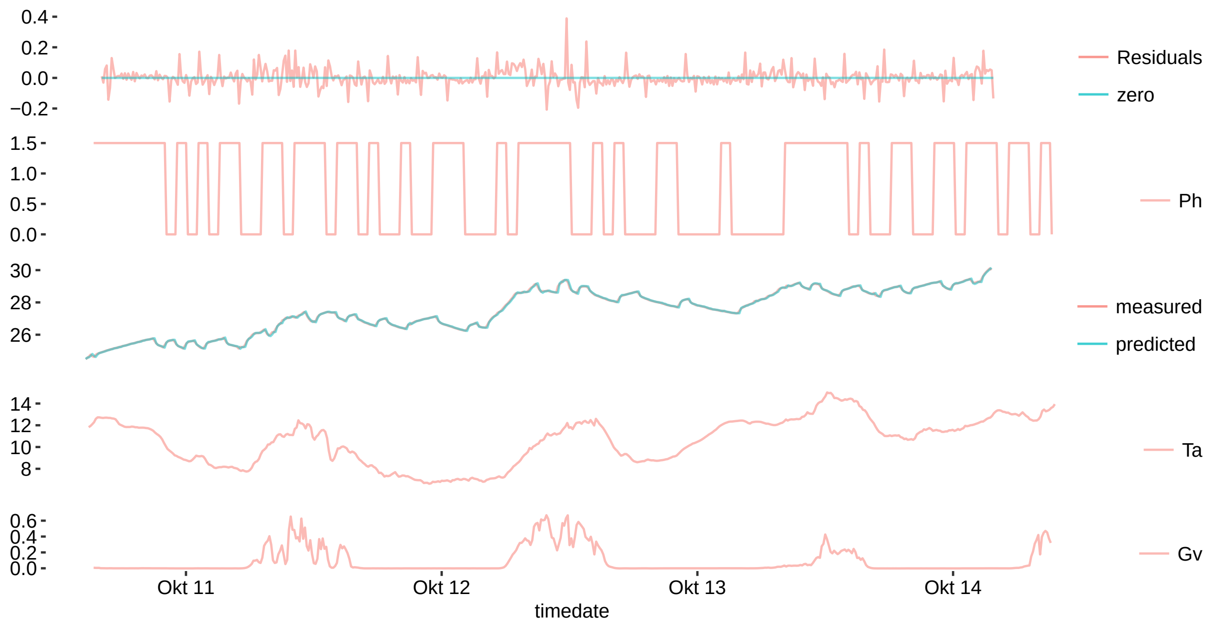
Periodogram of the residuals



Commulated Periodogram of the residuals



combined



7. Step

The P value from 0

Consider the following

- Discuss the white-noise properties of the (one-step ahead) residuals for model T_i.
  - –
  - –

- What useful information can be obtained from the time series plots of the residuals and the inputs for model T_i?
  -
  -
- Discuss the white-noise properties of the one-step ahead residuals for model T_iT_m.
  -
  -
- What useful information can be obtained from the time series plots of the residuals and inputs for model T_iT_m?
  -
  -
- Based on the likelihood-ratio test is model T_iT_m then to be preferred over model T_i?
  -
  -

Question 2b

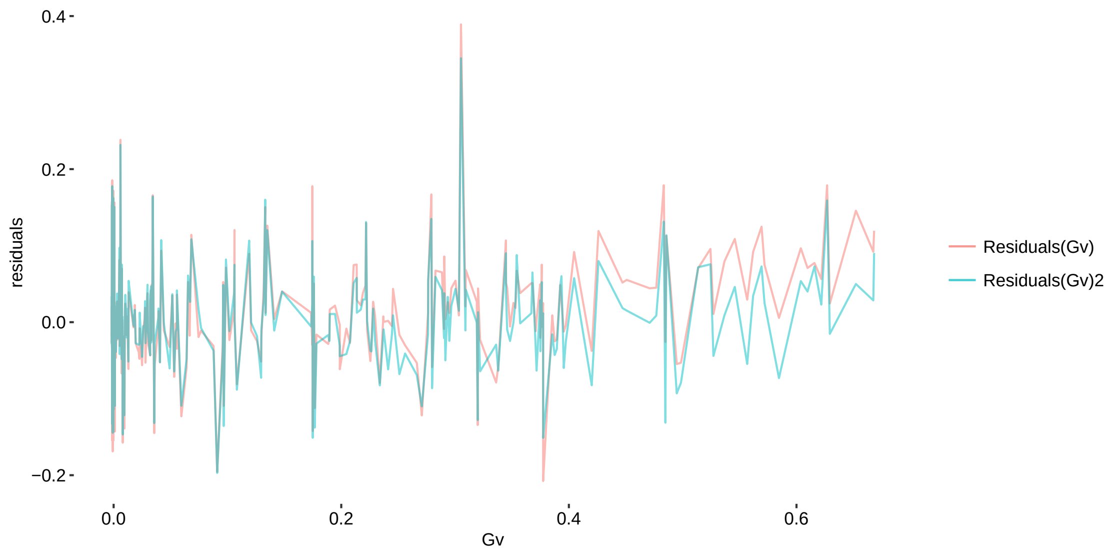
Describe the good way to implement it…!

what is Aw, Gv and p.. must

```
## 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

- Findings
  - Plot of the resi



  - Changes in squard residuals:
    * Question 2a: 1.8746528
    * Question 2b: 1.6630286
- Likelihood
  - Question 2a: 771.000411
  - Question 2b: 801.9309938
- Likelihood ratio test

  The P value from $3.6859404 \times 10^{-14}$
- Conclusion:

Question 2c

Describe how i changed input

```
## 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"
```

Findings

- Log like
  - Question 2b: 801.9309938
  - Question 2c: 533.8988427
- Cap
  - $C_i + C_m$
  - Question 2b: 19.5682341
  - Question 2c: 17.5739334

Question 2d

Not considered due to time.