Assignment #4

Peyman Kor

11/28/2019

#4.1

library(tidyverse)

## -- Attaching packages ------------------------------------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v purrr 0.3.2  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 0.8.3 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts ---------------------------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

data <- read.csv('A4\_Kulhuse.csv')  
head(data)

## Temp Sal Depth pH Chl ODOsat ODO Battery DateTime  
## 1 18.22 18.03 5.191 8.25 2.91 108.6 9.19 12.7 2017-08-24 10:00:00  
## 2 18.30 17.96 5.213 8.26 4.11 107.7 9.10 12.7 2017-08-24 10:30:00  
## 3 18.29 18.00 5.234 8.26 4.01 108.5 9.17 12.7 2017-08-24 11:00:00  
## 4 18.30 18.00 5.258 8.26 4.33 108.8 9.20 12.7 2017-08-24 11:30:00  
## 5 18.34 18.02 5.277 8.26 3.74 110.1 9.29 12.7 2017-08-24 12:00:00  
## 6 18.35 18.04 5.282 8.26 4.45 110.9 9.36 12.7 2017-08-24 12:30:00

Now, let’s have alook on the data for possible “NA” values,

data %>% filter(is.na(Sal))

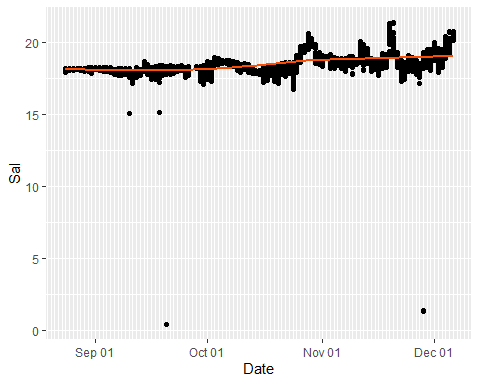
## Temp Sal Depth pH Chl ODOsat ODO Battery DateTime  
## 1 NA NA NA NA NA NA NA NA 2017-09-26 12:00:00  
## 2 NA NA NA NA NA NA NA NA 2017-09-26 12:30:00  
## 3 NA NA NA NA NA NA NA NA 2017-09-26 13:00:00  
## 4 NA NA NA NA NA NA NA NA 2017-09-26 13:30:00  
## 5 NA NA NA NA NA NA NA NA 2017-09-26 14:00:00  
## 6 NA NA NA NA NA NA NA NA 2017-09-26 14:30:00  
## 7 NA NA NA NA NA NA NA NA 2017-09-26 15:00:00  
## 8 NA NA NA NA NA NA NA NA 2017-09-26 15:30:00  
## 9 NA NA NA NA NA NA NA NA 2017-09-26 16:00:00  
## 10 NA NA NA NA NA NA NA NA 2017-09-26 16:30:00  
## 11 NA NA NA NA NA NA NA NA 2017-09-26 17:00:00  
## 12 NA NA NA NA NA NA NA NA 2017-09-26 17:30:00  
## 13 NA NA NA NA NA NA NA NA 2017-09-26 18:00:00  
## 14 NA NA NA NA NA NA NA NA 2017-09-26 18:30:00  
## 15 NA NA NA NA NA NA NA NA 2017-09-26 19:00:00  
## 16 NA NA NA NA NA NA NA NA 2017-09-26 19:30:00  
## 17 NA NA NA NA NA NA NA NA 2017-09-26 20:00:00  
## 18 NA NA NA NA NA NA NA NA 2017-09-26 20:30:00  
## 19 NA NA NA NA NA NA NA NA 2017-09-26 21:00:00  
## 20 NA NA NA NA NA NA NA NA 2017-09-26 21:30:00  
## 21 NA NA NA NA NA NA NA NA 2017-09-26 22:00:00  
## 22 NA NA NA NA NA NA NA NA 2017-09-26 22:30:00  
## 23 NA NA NA NA NA NA NA NA 2017-09-26 23:00:00  
## 24 NA NA NA NA NA NA NA NA 2017-09-26 23:30:00  
## 25 NA NA NA NA NA NA NA NA 2017-09-27 00:00:00  
## 26 NA NA NA NA NA NA NA NA 2017-09-27 00:30:00  
## 27 NA NA NA NA NA NA NA NA 2017-09-27 01:00:00  
## 28 NA NA NA NA NA NA NA NA 2017-09-27 01:30:00  
## 29 NA NA NA NA NA NA NA NA 2017-09-27 02:00:00  
## 30 NA NA NA NA NA NA NA NA 2017-09-27 02:30:00  
## 31 NA NA NA NA NA NA NA NA 2017-09-27 03:00:00  
## 32 NA NA NA NA NA NA NA NA 2017-09-27 03:30:00  
## 33 NA NA NA NA NA NA NA NA 2017-09-27 04:00:00  
## 34 NA NA NA NA NA NA NA NA 2017-09-27 04:30:00  
## 35 NA NA NA NA NA NA NA NA 2017-09-27 05:00:00  
## 36 NA NA NA NA NA NA NA NA 2017-09-27 05:30:00  
## 37 NA NA NA NA NA NA NA NA 2017-09-27 06:00:00  
## 38 NA NA NA NA NA NA NA NA 2017-09-27 06:30:00  
## 39 NA NA NA NA NA NA NA NA 2017-09-27 07:00:00  
## 40 NA NA NA NA NA NA NA NA 2017-09-27 07:30:00  
## 41 NA NA NA NA NA NA NA NA 2017-09-27 08:00:00  
## 42 NA NA NA NA NA NA NA NA 2017-09-27 08:30:00  
## 43 NA NA NA NA NA NA NA NA 2017-09-27 09:00:00  
## 44 NA NA NA NA NA NA NA NA 2017-09-27 09:30:00  
## 45 NA NA NA NA NA NA NA NA 2017-09-27 10:00:00  
## 46 NA NA NA NA NA NA NA NA 2017-09-27 10:30:00  
## 47 NA NA NA NA NA NA NA NA 2017-09-27 11:00:00  
## 48 NA NA NA NA NA NA NA NA 2017-09-27 11:30:00  
## 49 NA NA NA NA NA NA NA NA 2017-09-27 12:00:00  
## 50 NA NA NA NA NA NA NA NA 2017-09-27 12:30:00  
## 51 NA NA NA NA NA NA NA NA 2017-09-27 13:00:00  
## 52 NA NA NA NA NA NA NA NA 2017-09-27 13:30:00  
## 53 NA NA NA NA NA NA NA NA 2017-09-27 14:00:00  
## 54 NA NA NA NA NA NA NA NA 2017-09-27 14:30:00  
## 55 NA NA NA NA NA NA NA NA 2017-09-27 15:00:00  
## 56 NA NA NA NA NA NA NA NA 2017-09-27 15:30:00  
## 57 NA NA NA NA NA NA NA NA 2017-09-27 16:00:00  
## 58 NA NA NA NA NA NA NA NA 2017-09-27 16:30:00  
## 59 NA NA NA NA NA NA NA NA 2017-09-27 17:00:00  
## 60 NA NA NA NA NA NA NA NA 2017-09-27 17:30:00  
## 61 NA NA NA NA NA NA NA NA 2017-09-27 18:00:00  
## 62 NA NA NA NA NA NA NA NA 2017-09-27 18:30:00  
## 63 NA NA NA NA NA NA NA NA 2017-09-27 19:00:00  
## 64 NA NA NA NA NA NA NA NA 2017-09-27 19:30:00  
## 65 NA NA NA NA NA NA NA NA 2017-09-27 20:00:00  
## 66 NA NA NA NA NA NA NA NA 2017-09-27 20:30:00  
## 67 NA NA NA NA NA NA NA NA 2017-09-27 21:00:00  
## 68 NA NA NA NA NA NA NA NA 2017-09-27 21:30:00  
## 69 NA NA NA NA NA NA NA NA 2017-09-27 22:00:00  
## 70 NA NA NA NA NA NA NA NA 2017-09-27 22:30:00  
## 71 NA NA NA NA NA NA NA NA 2017-09-27 23:00:00  
## 72 NA NA NA NA NA NA NA NA 2017-09-27 23:30:00  
## 73 NA NA NA NA NA NA NA NA 2017-09-28 00:00:00  
## 74 NA NA NA NA NA NA NA NA 2017-09-28 00:30:00  
## 75 NA NA NA NA NA NA NA NA 2017-09-28 01:00:00  
## 76 NA NA NA NA NA NA NA NA 2017-09-28 01:30:00  
## 77 NA NA NA NA NA NA NA NA 2017-09-28 02:00:00  
## 78 NA NA NA NA NA NA NA NA 2017-09-28 02:30:00  
## 79 NA NA NA NA NA NA NA NA 2017-09-28 03:00:00  
## 80 NA NA NA NA NA NA NA NA 2017-09-28 03:30:00  
## 81 NA NA NA NA NA NA NA NA 2017-09-28 04:00:00  
## 82 NA NA NA NA NA NA NA NA 2017-09-28 04:30:00  
## 83 NA NA NA NA NA NA NA NA 2017-09-28 05:00:00  
## 84 NA NA NA NA NA NA NA NA 2017-09-28 05:30:00  
## 85 NA NA NA NA NA NA NA NA 2017-09-28 06:00:00  
## 86 NA NA NA NA NA NA NA NA 2017-09-28 06:30:00  
## 87 NA NA NA NA NA NA NA NA 2017-09-28 07:00:00  
## 88 NA NA NA NA NA NA NA NA 2017-09-28 07:30:00  
## 89 NA NA NA NA NA NA NA NA 2017-09-28 08:00:00  
## 90 NA NA NA NA NA NA NA NA 2017-09-28 08:30:00  
## 91 NA NA NA NA NA NA NA NA 2017-09-28 09:00:00  
## 92 NA NA NA NA NA NA NA NA 2017-09-28 09:30:00  
## 93 NA NA NA NA NA NA NA NA 2017-09-28 10:00:00  
## 94 NA NA NA NA NA NA NA NA 2017-09-28 10:30:00  
## 95 NA NA NA NA NA NA NA NA 2017-09-28 11:00:00  
## 96 NA NA NA NA NA NA NA NA 2017-09-28 11:30:00  
## 97 NA NA NA NA NA NA NA NA 2017-09-28 12:00:00  
## 98 NA NA NA NA NA NA NA NA 2017-09-28 12:30:00  
## 99 NA NA NA NA NA NA NA NA 2017-09-28 13:00:00  
## 100 NA NA NA NA NA NA NA NA 2017-09-28 13:30:00  
## 101 NA NA NA NA NA NA NA NA 2017-09-28 14:00:00  
## 102 NA NA NA NA NA NA NA NA 2017-09-28 14:30:00  
## 103 NA NA NA NA NA NA NA NA 2017-09-28 15:00:00  
## 104 NA NA NA NA NA NA NA NA 2017-09-28 15:30:00  
## 105 NA NA NA NA NA NA NA NA 2017-09-28 16:00:00  
## 106 NA NA NA NA NA NA NA NA 2017-10-01 08:00:00  
## 107 NA NA NA NA NA NA NA NA 2017-10-02 17:30:00  
## 108 NA NA NA NA NA NA NA NA 2017-10-11 14:30:00  
## 109 NA NA NA NA NA NA NA NA 2017-10-11 21:30:00  
## 110 NA NA NA NA NA NA NA NA 2017-10-12 05:30:00  
## 111 NA NA NA NA NA NA NA NA <NA>

We see that the column Salinity has 111 rows whcih are NA, also we could see other columns as well are NA, so could safely removes these 111 rows beacuse they contain no information.

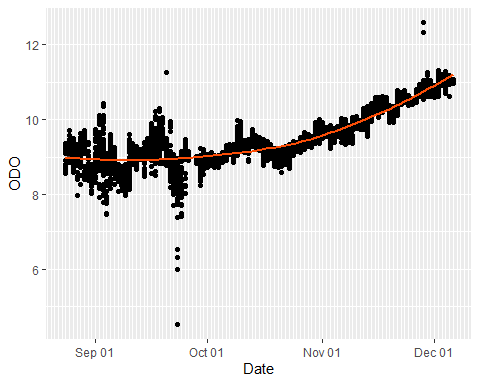
data\_nona <- data %>% filter(!is.na(Sal))

Having this now, we visulize the Salinity versus time:

library(ggplot2)  
data\_nona %>% mutate(Date = as.Date(DateTime)) %>%   
 ggplot(mapping = aes(Date,Sal)) +  
 geom\_point() +  
 scale\_x\_date(date\_minor\_breaks = "1 day ",date\_labels = "%b %d") +  
 stat\_smooth(  
 color = "#FC4E07",  
 method = "loess"  
 )



library(ggplot2)  
data\_nona %>% mutate(Date = as.Date(DateTime)) %>%   
 ggplot(mapping = aes(Date,ODO)) +  
 geom\_point() +  
 scale\_x\_date(date\_minor\_breaks = "1 day ",date\_labels = "%b %d") +  
 stat\_smooth(  
 color = "#FC4E07",  
 method = "loess"  
 )



# 4.2

The salinity assumed to be random walk process as the follow:

So, the following is the state space model for salinity, where the and are the white noise with the standard deviation of and .

On the other hand, the salinity could be written as the ARIMA process (0,1,1) as the follow:

In the state space model as the matrix,

Where the

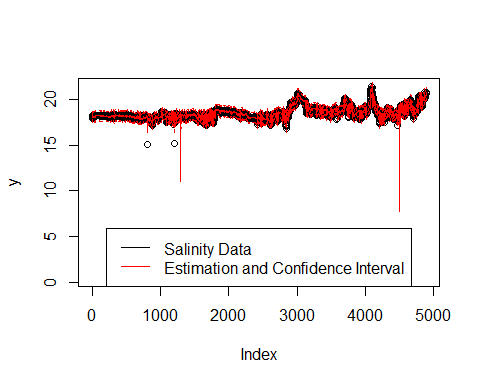
#3

library(FKF)

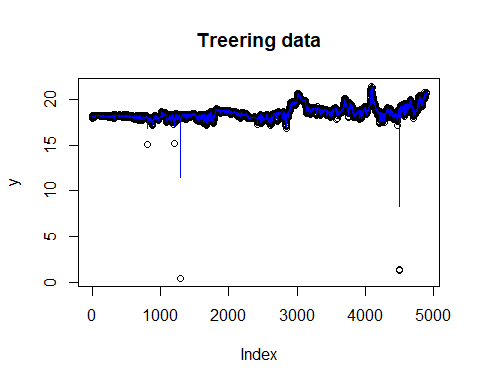
## Loading required package: RUnit

y <- data\_nona$Sal  
dt <- ct <- matrix(0)  
Zt <- Tt <- matrix(1)  
a0 <- y[1] # Estimation of the first year flow  
P0 <- matrix(0.01) # Variance of 'a0'  
fit.fkf <- optim(c(HHt = 0.01 ,  
 GGt = 0.005 ),  
 fn = function(par, ...)  
 -fkf(HHt = matrix(par[1]), GGt = matrix(par[2]), ...)$logLik,  
 yt = rbind(y), a0 = a0, P0 = P0, dt = dt, ct = ct,  
 Zt = Zt, Tt = Tt)  
## Filter Nile data with estimated parameters:plot.fkf 7  
fkf.obj <- fkf(a0, P0, dt, ct, Tt, Zt, HHt = matrix(fit.fkf$par[1]),  
 GGt = matrix(fit.fkf$par[2]), yt = rbind(y))

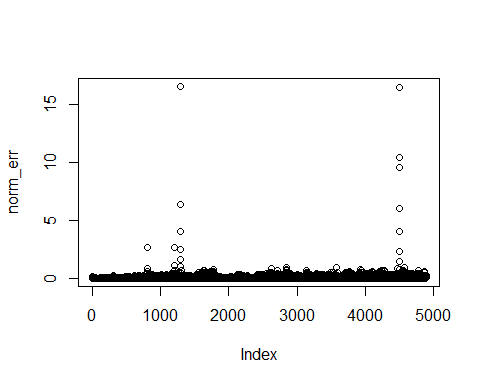
plot(y)  
with(fkf.obj, matlines((at[1,]) + cbind(0,-1.96\*sqrt(Pt[1,1,]),1.96\*sqrt(Pt[1,1,])),type="l", lty=c(1,2,2), col=2))  
legend("bottom", c("Salinity Data", "Estimation and Confidence Interval"),  
 col = c("black", "red"), lty = 1)



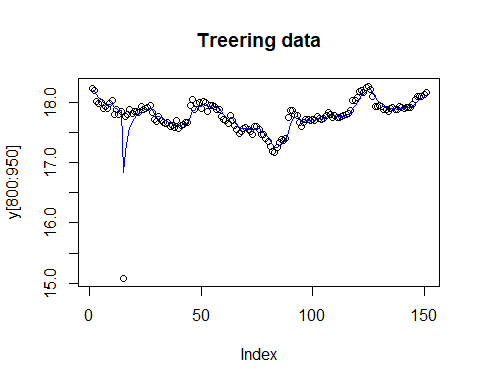
plot(y, main = "Treering data")  
lines(ts(fkf.obj$att[1, ], start = start(y), frequency = frequency(y)), col = "blue")



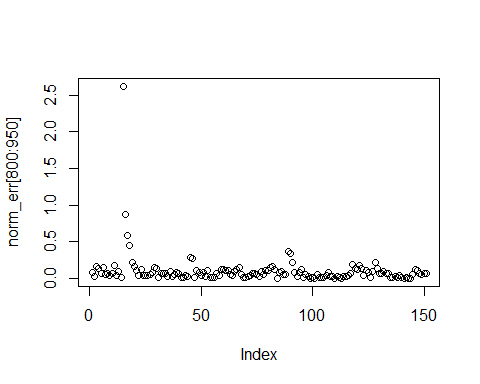
prediction <- fkf.obj$att[1, ]  
error <- abs(y-prediction)  
norm\_err <- error/sd(prediction)  
plot(norm\_err)



plot(y[800:950], main = "Treering data")  
lines(ts(fkf.obj$att[1, ][800:950], start = start(y), frequency = frequency(y)), col = "blue")



prediction <- fkf.obj$att[1, ]  
error <- abs(y-prediction)  
norm\_err <- error/sd(prediction)  
plot(norm\_err[800:950])



prediction[4889]

## [1] 20.69859

# 4

sd\_pre <- sd(prediction)  
sum(abs(prediction-y) > 6\*sd\_pre)

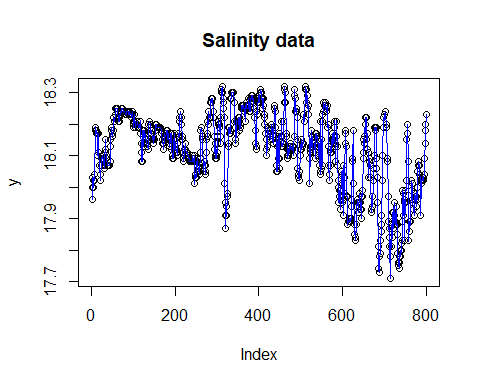
## [1] 6

# 5

y <- data\_nona$Sal[1:800]  
dt <- ct <- matrix(0)  
Zt <- Tt <- matrix(1)  
a0 <- y[1] # Estimation of the first year flow  
P0 <- matrix(0.01) # Variance of 'a0'  
fit.fkf <- optim(c(HHt = 0.0 ,  
 GGt = 0.005 ),  
 fn = function(par, ...)  
 -fkf(HHt = matrix(par[1]), GGt = matrix(par[2]), ...)$logLik,  
 yt = rbind(y), a0 = a0, P0 = P0, dt = dt, ct = ct,  
 Zt = Zt, Tt = Tt)  
  
fit.fkf$par

## HHt GGt   
## 1.910121e-03 -1.175588e-05

fkf.obj1 <- fkf(a0, P0, dt, ct, Tt, Zt, HHt = matrix(fit.fkf$par[1]),  
 GGt = matrix(fit.fkf$par[2]), yt = rbind(y))  
plot(y, main = "Salinity data")  
lines(ts(fkf.obj1$att[1, ], start = start(y), frequency = frequency(y)), col = "blue")



plot(data\_nona$Sal[800:950], main = "Salinity data")  
lines(ts(fkf.obj1$att[1, ], start = start(y), frequency = frequency(y)), col = "blue")

