10\_gls\_models

Jake Cavaiani

9/12/2022

#load libraries

library(here)

## here() starts at /Users/jakecavaiani/Documents/Storms\_clean\_repo

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5 ✓ purrr 0.3.4  
## ✓ tibble 3.1.6 ✓ dplyr 1.0.7  
## ✓ tidyr 1.1.3 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.1

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

library(nlme)

##   
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':  
##   
## collapse

library(forecast)

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

##   
## Attaching package: 'forecast'

## The following object is masked from 'package:nlme':  
##   
## getResponse

library(stats)  
library(readr)  
library(ggplot2)  
library(plotly)

##   
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':  
##   
## last\_plot

## The following object is masked from 'package:stats':  
##   
## filter

## The following object is masked from 'package:graphics':  
##   
## layout

library(GGally)

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

library(ggpmisc)

## Loading required package: ggpp

##   
## Attaching package: 'ggpp'

## The following object is masked from 'package:ggplot2':  
##   
## annotate

library(ggpubr)

##   
## Attaching package: 'ggpubr'

## The following object is masked from 'package:forecast':  
##   
## gghistogram

library(ggExtra)  
library(lubridate)

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

library(nlme)  
library(MuMIn)  
library(multcomp)

## Loading required package: mvtnorm

## Loading required package: survival

## Loading required package: TH.data

## Loading required package: MASS

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:plotly':  
##   
## select

## The following object is masked from 'package:dplyr':  
##   
## select

##   
## Attaching package: 'TH.data'

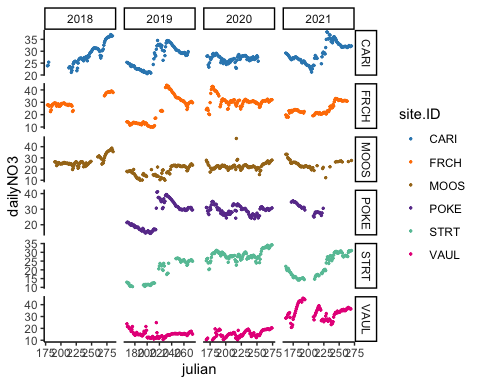
## The following object is masked from 'package:MASS':  
##   
## geyser

NO3\_data <- subset(mean\_daily\_long, response\_var == "dailyNO3" )  
  
which(mean\_daily$dailyNO3 > 60)

## [1] 1707

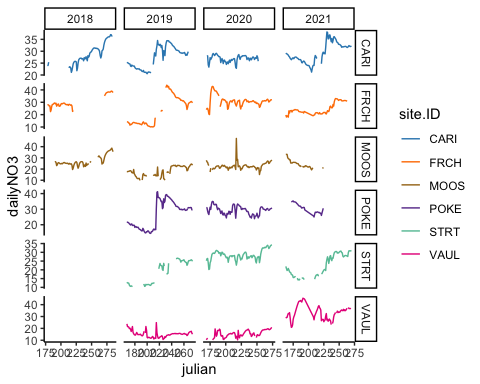
mean\_daily <- mean\_daily[-1707, ] # removing an errant point in FRCH 2021 nitrate that was over 60uM  
ggplot(mean\_daily, aes(x = julian, y = dailyNO3, color = site.ID)) +  
 geom\_point(size = 0.5) +  
 scale\_color\_manual(values=c("#3288BD","#FF7F00", "#A6761D", "#6A3D9A", "#66C2A5", "#E7298A")) +  
 facet\_grid(site.ID ~ year, scales = "free") +  
 theme\_classic()

## Warning: Removed 272 rows containing missing values (geom\_point).



ggplot(mean\_daily, aes(x = julian, y = dailyNO3, color = site.ID)) +  
 geom\_line() +  
 scale\_color\_manual(values=c("#3288BD","#FF7F00", "#A6761D", "#6A3D9A", "#66C2A5", "#E7298A")) +  
 facet\_grid(site.ID ~ year, scales = "free") +  
 theme\_classic()

## Warning: Removed 19 row(s) containing missing values (geom\_path).

 comparing each group to the base case (CARI) (and year of 2018) The intercept is the mean for CARI and the p-value each row for site is comparing that catchment mean across all years and site mean differs for caribou use a post hoc test to find if there is a significant interaction than we will run a post hoc test on the interaction

### Without 2018

The first thing I want to check here is if the gls does not the model because 2018 is missing 3 of the sites…so I want to investigate if the gls model structure works without 2018 data

Make sure that site.ID and year are a character

mean\_daily\_no\_2018 <- subset(mean\_daily, year != "2018")  
  
no3.mod.gls.no.2018 <- gls(dailyNO3 ~ site.ID\*year,  
 na.action = na.omit,  
 data = mean\_daily\_no\_2018)  
summary(no3.mod.gls.no.2018)

## Generalized least squares fit by REML  
## Model: dailyNO3 ~ site.ID \* year   
## Data: mean\_daily\_no\_2018   
## AIC BIC logLik  
## 10048.67 10151.09 -5005.336  
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 27.468525 0.5238140 52.43947 0.0000  
## site.IDFRCH -4.483416 0.7426916 -6.03671 0.0000  
## site.IDMOOS -9.491139 0.7527660 -12.60835 0.0000  
## site.IDPOKE -1.083696 0.7234338 -1.49799 0.1343  
## site.IDSTRT -8.588451 0.8081487 -10.62732 0.0000  
## site.IDVAUL -12.328836 0.7250403 -17.00435 0.0000  
## year2020 -0.604807 0.7735303 -0.78188 0.4344  
## year2021 1.920064 0.7334863 2.61772 0.0089  
## site.IDFRCH:year2020 8.561876 1.0637894 8.04847 0.0000  
## site.IDMOOS:year2020 5.386883 1.0719963 5.02509 0.0000  
## site.IDPOKE:year2020 3.453273 1.0492843 3.29107 0.0010  
## site.IDSTRT:year2020 10.162836 1.1093896 9.16075 0.0000  
## site.IDVAUL:year2020 0.998260 1.0602040 0.94157 0.3466  
## site.IDFRCH:year2021 -0.280166 1.0603230 -0.26423 0.7916  
## site.IDMOOS:year2021 4.489362 1.1374916 3.94672 0.0001  
## site.IDPOKE:year2021 2.360193 1.2150918 1.94240 0.0523  
## site.IDSTRT:year2021 1.951153 1.1087796 1.75973 0.0786  
## site.IDVAUL:year2021 16.172891 1.0415334 15.52796 0.0000  
##   
## Correlation:   
## (Intr) st.IDFRCH st.IDMOOS st.IDPOKE st.IDSTRT st.IDVAUL  
## site.IDFRCH -0.705   
## site.IDMOOS -0.696 0.491   
## site.IDPOKE -0.724 0.511 0.504   
## site.IDSTRT -0.648 0.457 0.451 0.469   
## site.IDVAUL -0.722 0.510 0.503 0.523 0.468   
## year2020 -0.677 0.478 0.471 0.490 0.439 0.489   
## year2021 -0.714 0.504 0.497 0.517 0.463 0.516   
## site.IDFRCH:year2020 0.492 -0.698 -0.343 -0.357 -0.319 -0.356   
## site.IDMOOS:year2020 0.489 -0.345 -0.702 -0.354 -0.317 -0.353   
## site.IDPOKE:year2020 0.499 -0.352 -0.347 -0.689 -0.324 -0.361   
## site.IDSTRT:year2020 0.472 -0.333 -0.329 -0.342 -0.728 -0.341   
## site.IDVAUL:year2020 0.494 -0.348 -0.344 -0.358 -0.320 -0.684   
## site.IDFRCH:year2021 0.494 -0.700 -0.344 -0.358 -0.320 -0.357   
## site.IDMOOS:year2021 0.460 -0.325 -0.662 -0.333 -0.298 -0.333   
## site.IDPOKE:year2021 0.431 -0.304 -0.300 -0.595 -0.279 -0.311   
## site.IDSTRT:year2021 0.472 -0.333 -0.329 -0.342 -0.729 -0.341   
## site.IDVAUL:year2021 0.503 -0.355 -0.350 -0.364 -0.326 -0.696   
## yr2020 yr2021 s.IDFRCH:2020 s.IDMOOS:2020 s.IDPOKE:2020  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021 0.484   
## site.IDFRCH:year2020 -0.727 -0.352   
## site.IDMOOS:year2020 -0.722 -0.349 0.525   
## site.IDPOKE:year2020 -0.737 -0.357 0.536 0.532   
## site.IDSTRT:year2020 -0.697 -0.337 0.507 0.503 0.514   
## site.IDVAUL:year2020 -0.730 -0.353 0.531 0.526 0.538   
## site.IDFRCH:year2021 -0.335 -0.692 0.489 0.241 0.247   
## site.IDMOOS:year2021 -0.312 -0.645 0.227 0.465 0.230   
## site.IDPOKE:year2021 -0.292 -0.604 0.212 0.211 0.410   
## site.IDSTRT:year2021 -0.320 -0.662 0.233 0.231 0.236   
## site.IDVAUL:year2021 -0.341 -0.704 0.248 0.246 0.251   
## s.IDSTRT:2020 s.IDVAUL:2020 s.IDFRCH:2021 s.IDMOOS:2021  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021   
## site.IDFRCH:year2020   
## site.IDMOOS:year2020   
## site.IDPOKE:year2020   
## site.IDSTRT:year2020   
## site.IDVAUL:year2020 0.509   
## site.IDFRCH:year2021 0.233 0.244   
## site.IDMOOS:year2021 0.217 0.228 0.446   
## site.IDPOKE:year2021 0.204 0.213 0.418 0.389   
## site.IDSTRT:year2021 0.531 0.233 0.458 0.427   
## site.IDVAUL:year2021 0.237 0.476 0.487 0.454   
## s.IDPOKE:2021 s.IDSTRT:2021  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021   
## site.IDFRCH:year2020   
## site.IDMOOS:year2020   
## site.IDPOKE:year2020   
## site.IDSTRT:year2020   
## site.IDVAUL:year2020   
## site.IDFRCH:year2021   
## site.IDMOOS:year2021   
## site.IDPOKE:year2021   
## site.IDSTRT:year2021 0.399   
## site.IDVAUL:year2021 0.425 0.466   
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -2.49528707 -0.57634733 -0.02535241 0.61499047 4.74493373   
##   
## Residual standard error: 5.185494   
## Degrees of freedom: 1638 total; 1620 residual

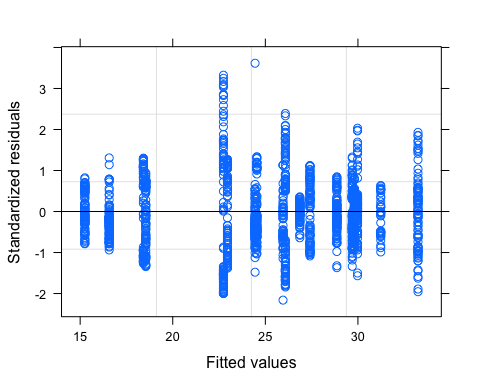
### CorAR1 structure

no3.mod.ar1.no.2018 <- gls(dailyNO3 ~ site.ID\*year,   
 na.action = na.omit,   
 data = mean\_daily\_no\_2018,   
 correlation = corAR1(form = ~ julian|site.ID/year))

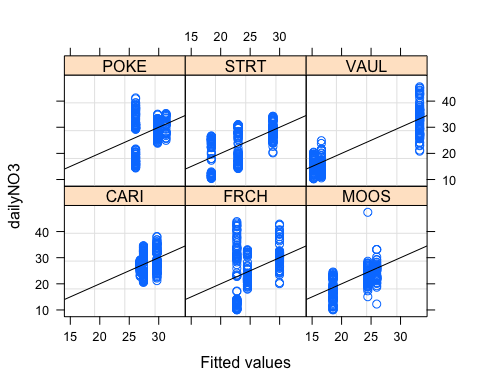
This works without 2018 data

# diagnostic plots

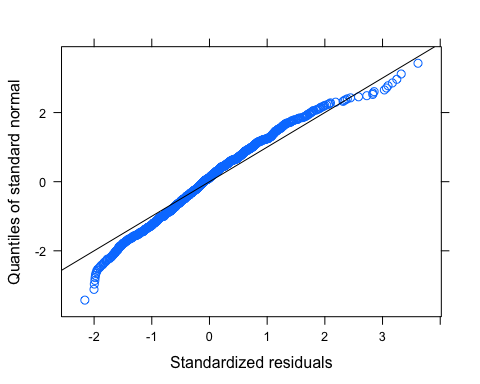
plot(no3.mod.ar1.no.2018, resid(., type = "p") ~ fitted(.), abline = 0)



plot(no3.mod.ar1.no.2018, dailyNO3 ~ fitted(.) | site.ID, abline = c(0,1))

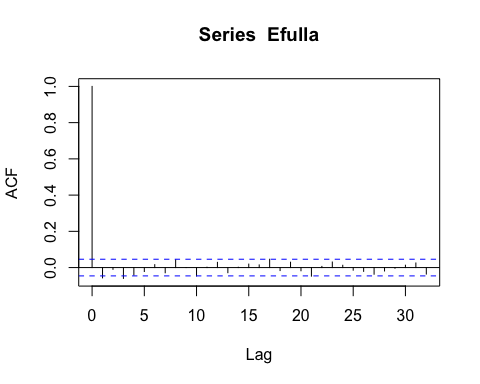


qqnorm(no3.mod.ar1.no.2018, abline = c(0,1))

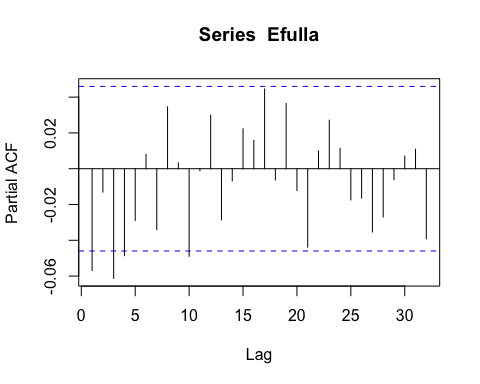


#ACF plots

Ear1<-residuals(no3.mod.ar1.no.2018, type="normalized")  
I1<-!is.na(mean\_daily\_no\_2018$dailyNO3)  
Efulla<-vector(length = length(mean\_daily\_no\_2018$dailyNO3))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



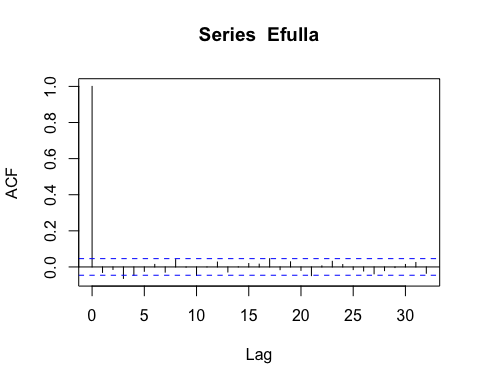
pacf(Efulla, na.action=na.pass)

 Significance at multiple lags….lets try the ARMA structure

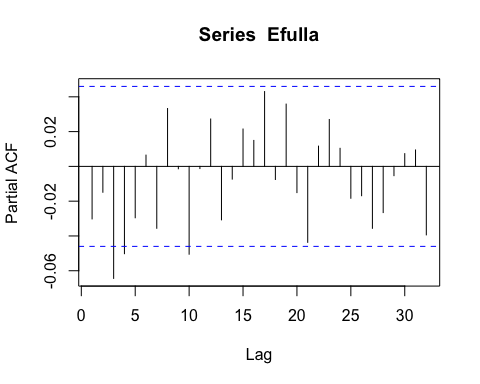
# corARMA structure

no3.mod.arma.1.1.no.2018 <- gls(dailyNO3 ~ site.ID\*year,   
 na.action = na.omit,   
 data = mean\_daily\_no\_2018,   
 correlation = corARMA(form = ~ julian|site.ID/year, p = 1, q = 1))

Ear1<-residuals(no3.mod.arma.1.1.no.2018, type="normalized")  
I1<-!is.na(mean\_daily\_no\_2018$dailyNO3)  
Efulla<-vector(length = length(mean\_daily\_no\_2018$dailyNO3))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)

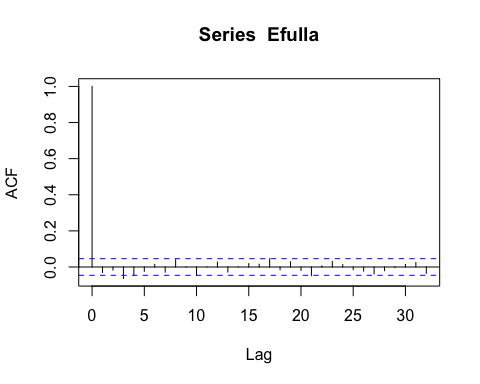


pacf(Efulla, na.action=na.pass)

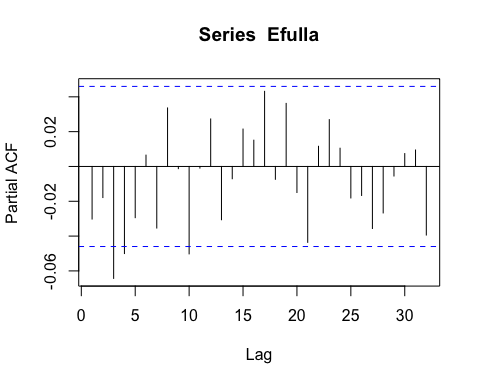
 Still some significance on multiple lags

mean\_daily\_no\_2018$site.ID <- as.factor(mean\_daily\_no\_2018$site.ID)  
no3.mod.arma.1.2.no.2018 <- gls(dailyNO3 ~ site.ID\*year,   
 na.action = na.omit,   
 data = mean\_daily\_no\_2018,   
 correlation = corARMA(form = ~ julian|site.ID/year, p = 1, q = 2))

Ear1<-residuals(no3.mod.arma.1.2.no.2018, type="normalized")  
I1<-!is.na(mean\_daily\_no\_2018$dailyNO3)  
Efulla<-vector(length = length(mean\_daily\_no\_2018$dailyNO3))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)

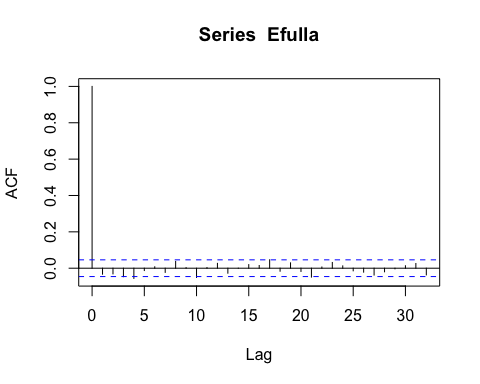


pacf(Efulla, na.action=na.pass)

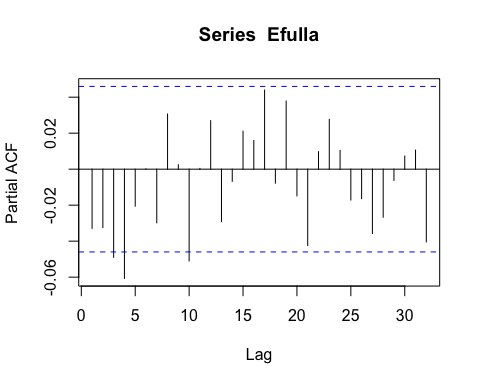


no3.mod.arma.2.1.no.2018 <- gls(dailyNO3 ~ site.ID\*year,   
 na.action = na.omit,   
 data = mean\_daily\_no\_2018,   
 correlation = corARMA(form = ~ julian|site.ID/year, p = 2, q = 1))

Ear1<-residuals(no3.mod.arma.2.1.no.2018, type="normalized")  
I1<-!is.na(mean\_daily\_no\_2018$dailyNO3)  
Efulla<-vector(length = length(mean\_daily\_no\_2018$dailyNO3))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)

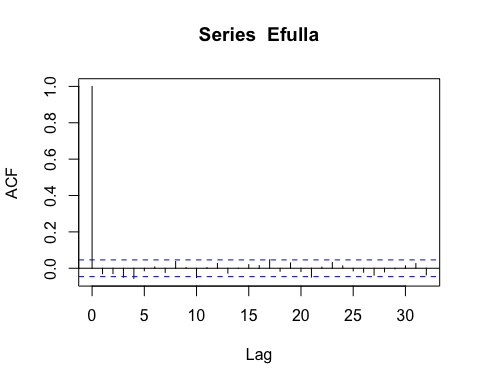


pacf(Efulla, na.action=na.pass)

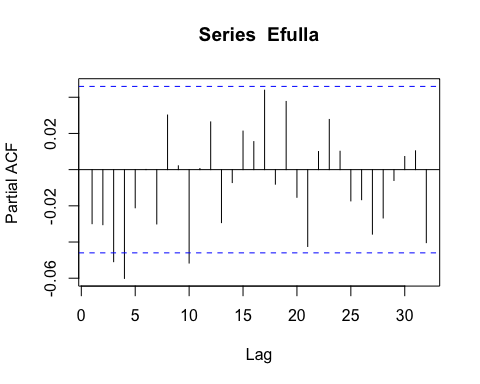


no3.mod.arma.2.2.no.2018 <- gls(dailyNO3 ~ site.ID\*year,   
 na.action = na.omit,   
 data = mean\_daily\_no\_2018,   
 correlation = corARMA(form = ~ julian|site.ID/year, p = 2, q = 2))

Ear1<-residuals(no3.mod.arma.2.2.no.2018, type="normalized")  
I1<-!is.na(mean\_daily\_no\_2018$dailyNO3)  
Efulla<-vector(length = length(mean\_daily\_no\_2018$dailyNO3))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)

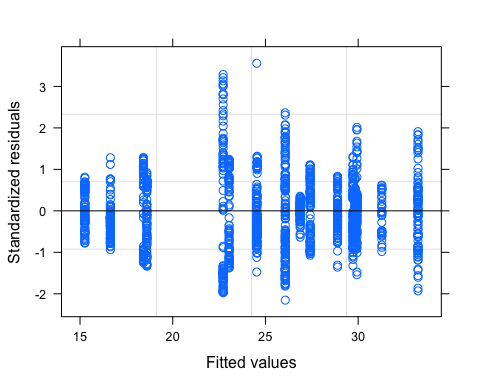


pacf(Efulla, na.action=na.pass)

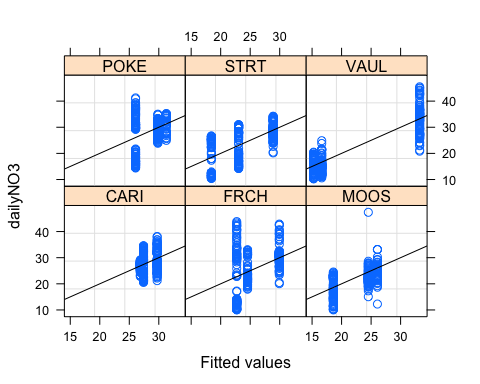


# diagnostic plots

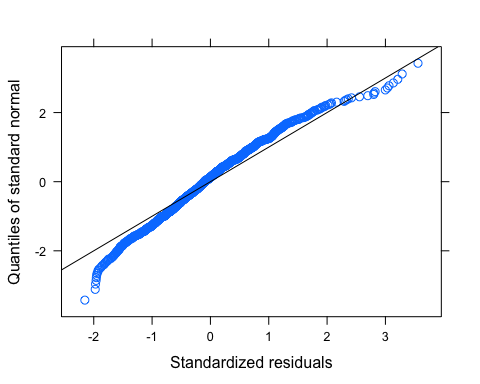
plot(no3.mod.arma.1.2.no.2018, resid(., type = "p") ~ fitted(.), abline = 0)



plot(no3.mod.arma.1.2.no.2018, dailyNO3 ~ fitted(.) | site.ID, abline = c(0,1))



qqnorm(no3.mod.arma.1.2.no.2018, abline = c(0,1))



summary(no3.mod.arma.1.2.no.2018)

## Generalized least squares fit by REML  
## Model: dailyNO3 ~ site.ID \* year   
## Data: mean\_daily\_no\_2018   
## AIC BIC logLik  
## 6426.307 6544.89 -3191.153  
##   
## Correlation Structure: ARMA(1,2)  
## Formula: ~julian | site.ID/year   
## Parameter estimate(s):  
## Phi1 Theta1 Theta2   
## 0.966274864 0.003620006 -0.031054604   
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 27.411277 3.891607 7.043691 0.0000  
## site.IDFRCH -4.693128 5.445419 -0.861849 0.3889  
## site.IDMOOS -8.807659 5.445409 -1.617447 0.1060  
## site.IDPOKE -1.345001 5.444920 -0.247019 0.8049  
## site.IDSTRT -8.989633 5.451072 -1.649149 0.0993  
## site.IDVAUL -10.774344 5.444925 -1.978787 0.0480  
## year2020 -0.539471 5.677337 -0.095022 0.9243  
## year2021 2.308523 5.445189 0.423956 0.6717  
## site.IDFRCH:year2020 7.753925 7.837118 0.989385 0.3226  
## site.IDMOOS:year2020 6.465708 7.837114 0.825011 0.4095  
## site.IDPOKE:year2020 4.309509 7.836767 0.549909 0.5825  
## site.IDSTRT:year2020 11.006539 7.841043 1.403709 0.1606  
## site.IDVAUL:year2020 -0.837980 7.837060 -0.106925 0.9149  
## site.IDFRCH:year2021 -0.475442 7.702434 -0.061726 0.9508  
## site.IDMOOS:year2021 5.159595 7.660829 0.673503 0.5007  
## site.IDPOKE:year2021 2.901186 8.120483 0.357268 0.7209  
## site.IDSTRT:year2021 2.309926 7.665783 0.301329 0.7632  
## site.IDVAUL:year2021 14.282422 7.665977 1.863092 0.0626  
##   
## Correlation:   
## (Intr) st.IDFRCH st.IDMOOS st.IDPOKE st.IDSTRT st.IDVAUL  
## site.IDFRCH -0.715   
## site.IDMOOS -0.715 0.511   
## site.IDPOKE -0.715 0.511 0.511   
## site.IDSTRT -0.714 0.510 0.510 0.510   
## site.IDVAUL -0.715 0.511 0.511 0.511 0.510   
## year2020 -0.685 0.490 0.490 0.490 0.489 0.490   
## year2021 -0.715 0.511 0.511 0.511 0.510 0.511   
## site.IDFRCH:year2020 0.497 -0.695 -0.355 -0.355 -0.355 -0.355   
## site.IDMOOS:year2020 0.497 -0.355 -0.695 -0.355 -0.355 -0.355   
## site.IDPOKE:year2020 0.497 -0.355 -0.355 -0.695 -0.355 -0.355   
## site.IDSTRT:year2020 0.496 -0.355 -0.355 -0.355 -0.695 -0.355   
## site.IDVAUL:year2020 0.497 -0.355 -0.355 -0.355 -0.355 -0.695   
## site.IDFRCH:year2021 0.505 -0.707 -0.361 -0.361 -0.361 -0.361   
## site.IDMOOS:year2021 0.508 -0.363 -0.711 -0.363 -0.363 -0.363   
## site.IDPOKE:year2021 0.479 -0.342 -0.342 -0.671 -0.342 -0.343   
## site.IDSTRT:year2021 0.508 -0.363 -0.363 -0.363 -0.711 -0.363   
## site.IDVAUL:year2021 0.508 -0.363 -0.363 -0.363 -0.362 -0.710   
## yr2020 yr2021 s.IDFRCH:2020 s.IDMOOS:2020 s.IDPOKE:2020  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021 0.490   
## site.IDFRCH:year2020 -0.724 -0.355   
## site.IDMOOS:year2020 -0.724 -0.355 0.525   
## site.IDPOKE:year2020 -0.724 -0.355 0.525 0.525   
## site.IDSTRT:year2020 -0.724 -0.355 0.525 0.525 0.525   
## site.IDVAUL:year2020 -0.724 -0.355 0.525 0.525 0.525   
## site.IDFRCH:year2021 -0.346 -0.707 0.491 0.251 0.251   
## site.IDMOOS:year2021 -0.348 -0.711 0.252 0.494 0.252   
## site.IDPOKE:year2021 -0.328 -0.671 0.238 0.238 0.466   
## site.IDSTRT:year2021 -0.348 -0.710 0.252 0.252 0.252   
## site.IDVAUL:year2021 -0.348 -0.710 0.252 0.252 0.252   
## s.IDSTRT:2020 s.IDVAUL:2020 s.IDFRCH:2021 s.IDMOOS:2021  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021   
## site.IDFRCH:year2020   
## site.IDMOOS:year2020   
## site.IDPOKE:year2020   
## site.IDSTRT:year2020   
## site.IDVAUL:year2020 0.525   
## site.IDFRCH:year2021 0.251 0.251   
## site.IDMOOS:year2021 0.252 0.252 0.502   
## site.IDPOKE:year2021 0.238 0.238 0.474 0.477   
## site.IDSTRT:year2021 0.494 0.252 0.502 0.505   
## site.IDVAUL:year2021 0.252 0.493 0.502 0.505   
## s.IDPOKE:2021 s.IDSTRT:2021  
## site.IDFRCH   
## site.IDMOOS   
## site.IDPOKE   
## site.IDSTRT   
## site.IDVAUL   
## year2020   
## year2021   
## site.IDFRCH:year2020   
## site.IDMOOS:year2020   
## site.IDPOKE:year2020   
## site.IDSTRT:year2020   
## site.IDVAUL:year2020   
## site.IDFRCH:year2021   
## site.IDMOOS:year2021   
## site.IDPOKE:year2021   
## site.IDSTRT:year2021 0.476   
## site.IDVAUL:year2021 0.476 0.505   
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -2.1528695 -0.5688215 -0.1017070 0.4662106 3.5586151   
##   
## Residual standard error: 6.416662   
## Degrees of freedom: 1638 total; 1620 residual

# generalized linear hypotheses

site.ID.comp <- glht(no3.mod.arma.1.2.no.2018, linfct = mcp(site.ID = "Tukey"))

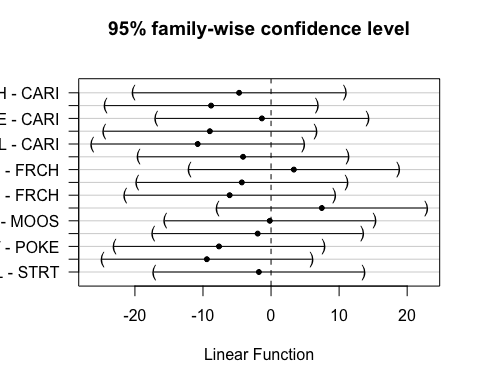
## Warning in mcp2matrix(model, linfct = linfct): covariate interactions found --  
## default contrast might be inappropriate

summary(site.ID.comp)

##   
## Simultaneous Tests for General Linear Hypotheses  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = dailyNO3 ~ site.ID \* year, data = mean\_daily\_no\_2018,   
## correlation = corARMA(form = ~julian | site.ID/year, p = 1,   
## q = 2), na.action = na.omit)  
##   
## Linear Hypotheses:  
## Estimate Std. Error z value Pr(>|z|)  
## FRCH - CARI == 0 -4.693 5.445 -0.862 0.955  
## MOOS - CARI == 0 -8.808 5.445 -1.617 0.587  
## POKE - CARI == 0 -1.345 5.445 -0.247 1.000  
## STRT - CARI == 0 -8.990 5.451 -1.649 0.566  
## VAUL - CARI == 0 -10.774 5.445 -1.979 0.354  
## MOOS - FRCH == 0 -4.114 5.387 -0.764 0.973  
## POKE - FRCH == 0 3.348 5.386 0.622 0.989  
## STRT - FRCH == 0 -4.296 5.392 -0.797 0.968  
## VAUL - FRCH == 0 -6.081 5.386 -1.129 0.869  
## POKE - MOOS == 0 7.463 5.386 1.386 0.736  
## STRT - MOOS == 0 -0.182 5.392 -0.034 1.000  
## VAUL - MOOS == 0 -1.967 5.386 -0.365 0.999  
## STRT - POKE == 0 -7.645 5.392 -1.418 0.716  
## VAUL - POKE == 0 -9.429 5.386 -1.751 0.498  
## VAUL - STRT == 0 -1.785 5.392 -0.331 0.999  
## (Adjusted p values reported -- single-step method)

plot(print(confint(site.ID.comp)))

##   
## Simultaneous Confidence Intervals  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = dailyNO3 ~ site.ID \* year, data = mean\_daily\_no\_2018,   
## correlation = corARMA(form = ~julian | site.ID/year, p = 1,   
## q = 2), na.action = na.omit)  
##   
## Quantile = 2.85  
## 95% family-wise confidence level  
##   
##   
## Linear Hypotheses:  
## Estimate lwr upr   
## FRCH - CARI == 0 -4.6931 -20.2124 10.8261  
## MOOS - CARI == 0 -8.8077 -24.3269 6.7115  
## POKE - CARI == 0 -1.3450 -16.8628 14.1728  
## STRT - CARI == 0 -8.9896 -24.5250 6.5457  
## VAUL - CARI == 0 -10.7743 -26.2922 4.7435  
## MOOS - FRCH == 0 -4.1145 -19.4662 11.2372  
## POKE - FRCH == 0 3.3481 -12.0022 18.6984  
## STRT - FRCH == 0 -4.2965 -19.6645 11.0715  
## VAUL - FRCH == 0 -6.0812 -21.4315 9.2691  
## POKE - MOOS == 0 7.4627 -7.8876 22.8129  
## STRT - MOOS == 0 -0.1820 -15.5500 15.1860  
## VAUL - MOOS == 0 -1.9667 -17.3170 13.3836  
## STRT - POKE == 0 -7.6446 -23.0112 7.7219  
## VAUL - POKE == 0 -9.4293 -24.7782 5.9195  
## VAUL - STRT == 0 -1.7847 -17.1513 13.5819

 Still showing no significance between the sites for 2019-2021 data

### BY YEAR

Due to the fact that we dont have data for 3 of the sites for 2018 the gls structure does not like that…so we will try this analysis by year and see if concentrations vary significantly from each other within years

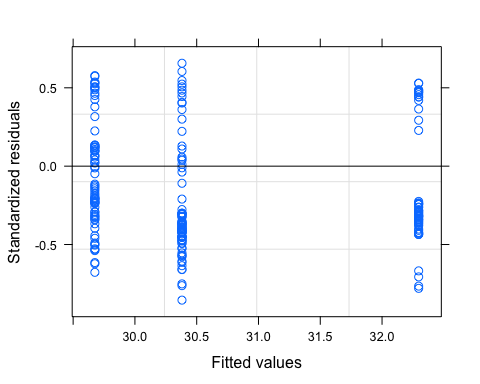
# NO3

# CorAR1 structure

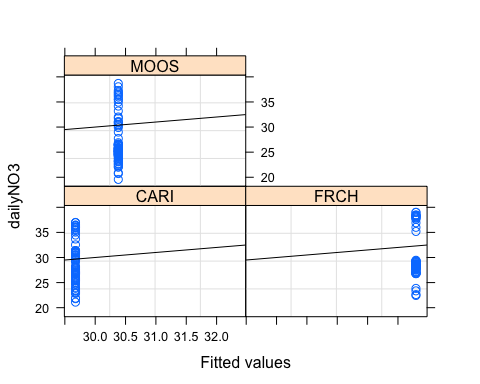
mean\_daily\_2018 <- subset(mean\_daily, year == "2018")  
  
no3.mod.gls.2018 <- gls(dailyNO3 ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

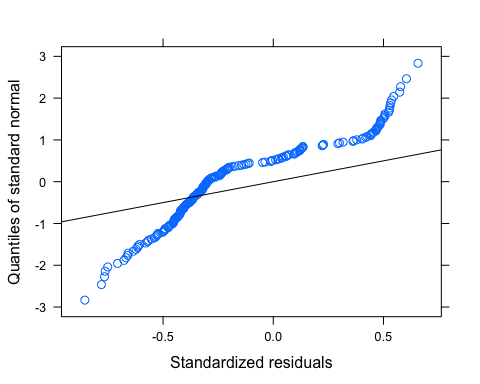
plot(no3.mod.gls.2018, resid(., type = "p") ~ fitted(.), abline = 0)



plot(no3.mod.gls.2018, dailyNO3 ~ fitted(.) | site.ID, abline = c(0,1))



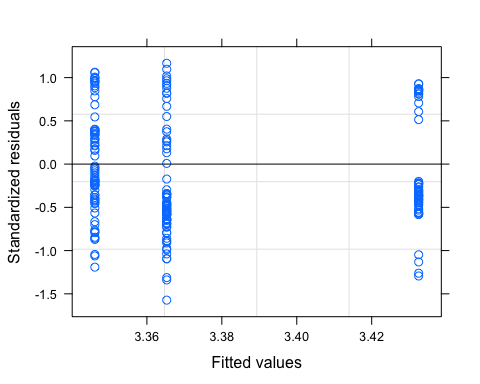
qqnorm(no3.mod.gls.2018, abline = c(0,1))

 Our normality is not great…lets log transform

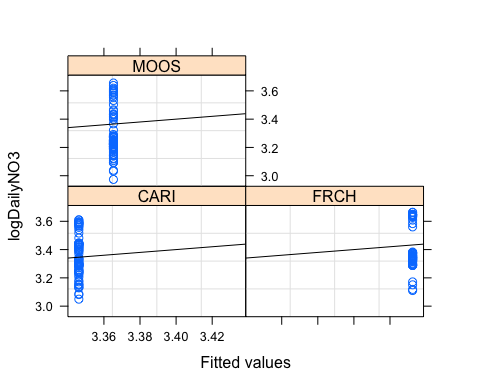
mean\_daily\_2018$logDailyNO3 <- log(abs(mean\_daily\_2018$dailyNO3))  
  
no3.mod.gls.2018.log <- gls(logDailyNO3 ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

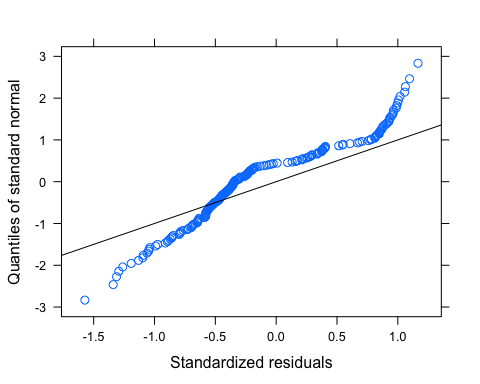
plot(no3.mod.gls.2018.log, resid(., type = "p") ~ fitted(.), abline = 0)



plot(no3.mod.gls.2018.log, logDailyNO3 ~ fitted(.) | site.ID, abline = c(0,1))



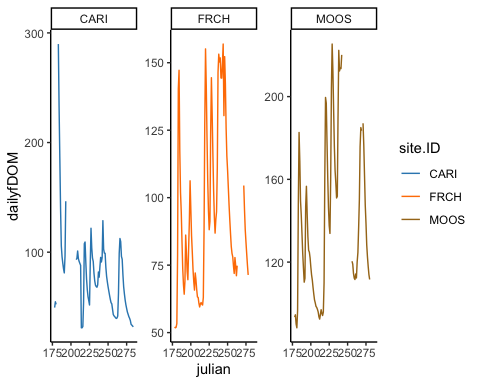
qqnorm(no3.mod.gls.2018.log, abline = c(0,1))

 Still not good for normality…we have to transform further before preceding

### fDOM

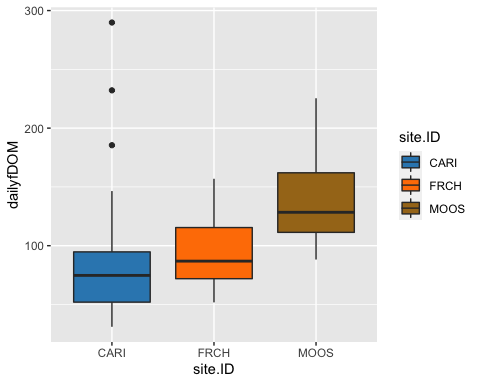
ggplot(mean\_daily\_2018, aes(x = julian, y = dailyfDOM, color = site.ID)) +  
 geom\_line() +  
 scale\_color\_manual(values=c("#3288BD","#FF7F00", "#A6761D")) +  
 facet\_wrap(~site.ID, scales = "free") +  
 theme\_classic()

## Warning: Removed 12 row(s) containing missing values (geom\_path).



ggplot(mean\_daily\_2018, aes(x = site.ID, y = dailyfDOM, fill = site.ID)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#3288BD","#FF7F00", "#A6761D"))

## Warning: Removed 38 rows containing non-finite values (stat\_boxplot).

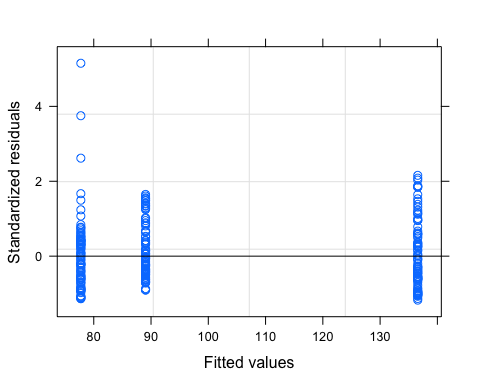


# corAR1 structure

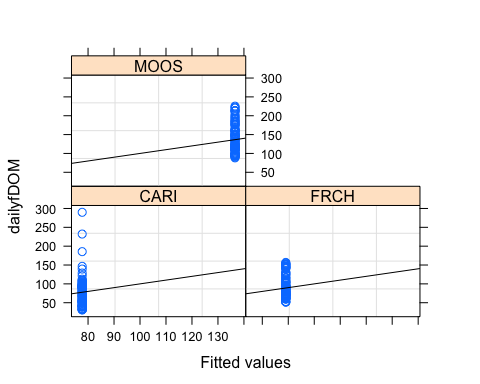
fDOM.mod.ar1.2018 <- gls(dailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

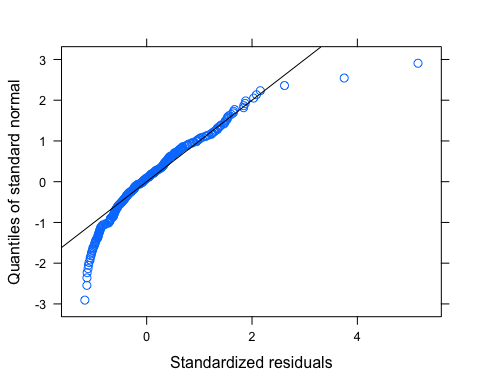
plot(fDOM.mod.ar1.2018, resid(., type = "p") ~ fitted(.), abline = 0)



plot(fDOM.mod.ar1.2018, dailyfDOM ~ fitted(.) | site.ID, abline = c(0,1))



qqnorm(fDOM.mod.ar1.2018, abline = c(0,1))



Let me log transform to see if its better

# log transformed

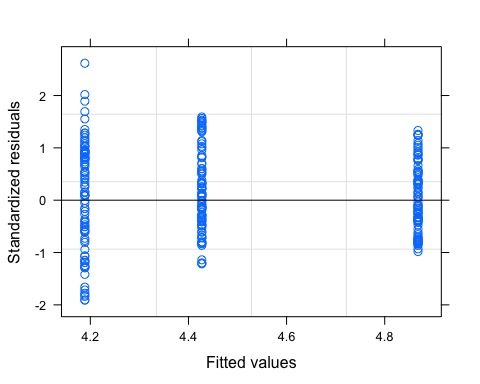
mean\_daily\_2018$logDailyfDOM <- log(mean\_daily\_2018$dailyfDOM)  
  
# removing clear outliers here   
which(mean\_daily\_2018$logDailyfDOM > 5.4)

## [1] 6 7 256 265

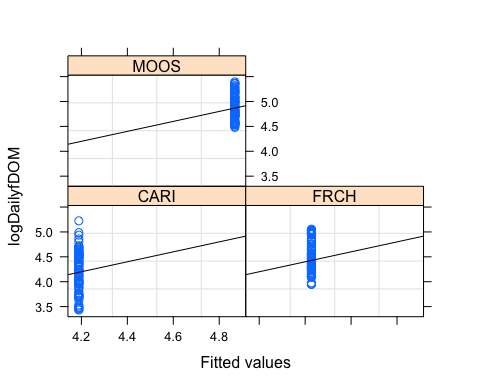
mean\_daily\_2018 <- mean\_daily\_2018 %>%  
 mutate(across(c(logDailyfDOM),   
 ~ifelse(logDailyfDOM > 5.4, NA, .)))  
  
fDOM.mod.ar1.log <- gls(logDailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

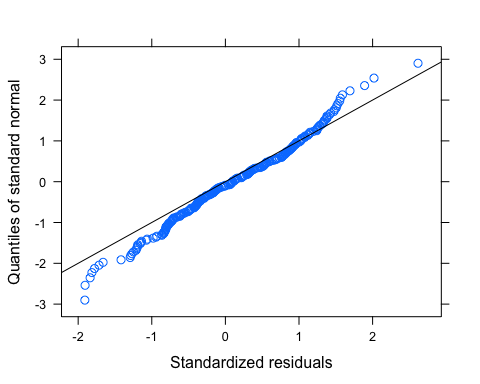
plot(fDOM.mod.ar1.log, resid(., type = "p") ~ fitted(.), abline = 0)



plot(fDOM.mod.ar1.log, logDailyfDOM ~ fitted(.) | site.ID, abline = c(0,1))

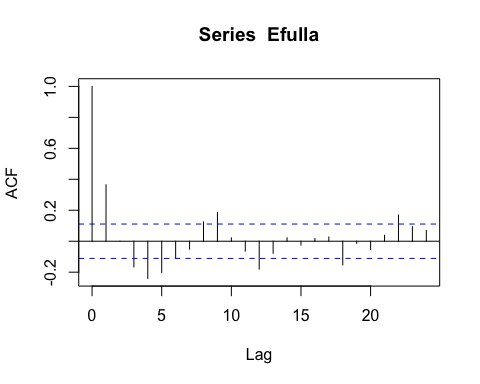


qqnorm(fDOM.mod.ar1.log, abline = c(0,1))

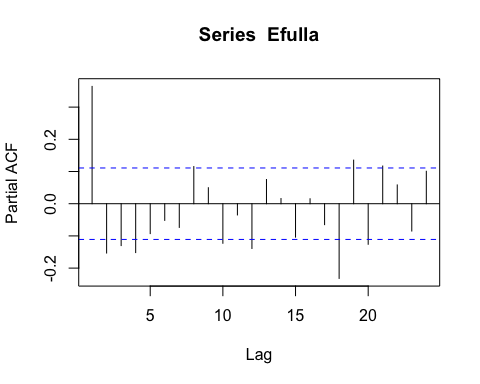


# ACF plot

Ear1<-residuals(fDOM.mod.ar1.log, type="normalized")  
I1<-!is.na(mean\_daily\_2018$logDailyfDOM)  
Efulla<-vector(length = length(mean\_daily\_2018$logDailyfDOM))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

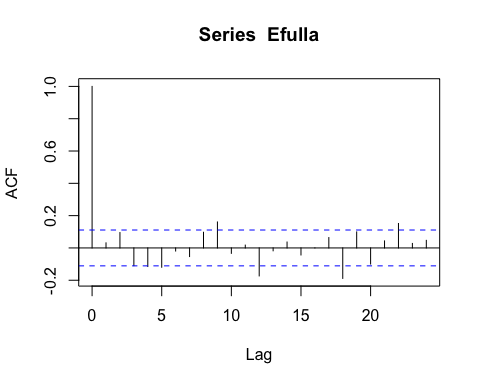
 AR1 is not adequately handling the temporal autocorrelation

# corARMA structure

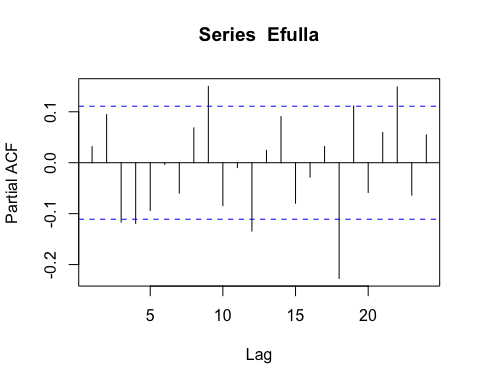
fDOM.mod.arma.1.1 <- gls(logDailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 1, q = 1))

# ACF plot

Ear1<-residuals(fDOM.mod.arma.1.1, type="normalized")  
I1<-!is.na(mean\_daily\_2018$logDailyfDOM)  
Efulla<-vector(length = length(mean\_daily\_2018$logDailyfDOM))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

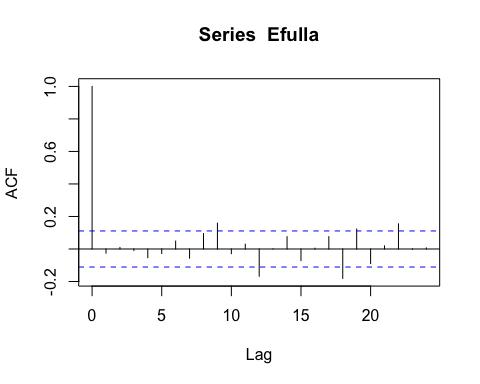
 Still autocorrelation

# corARMA structure

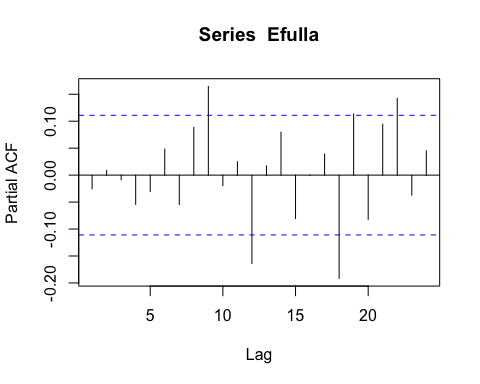
fDOM.mod.arma.2.1 <- gls(logDailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 2, q = 1))

# ACF plot

Ear1<-residuals(fDOM.mod.arma.2.1, type="normalized")  
I1<-!is.na(mean\_daily\_2018$logDailyfDOM)  
Efulla<-vector(length = length(mean\_daily\_2018$logDailyfDOM))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

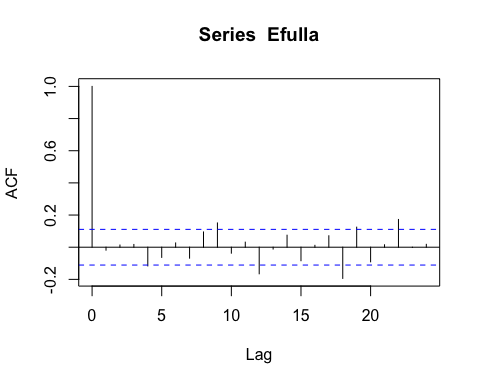
 gets worse in the later lags

# corARMA structure

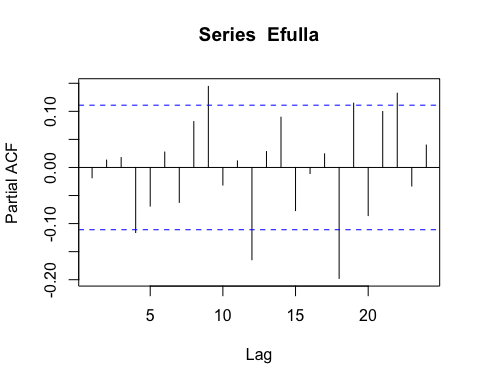
fDOM.mod.arma.1.2 <- gls(logDailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 1, q = 2))

# ACF plot

Ear1<-residuals(fDOM.mod.arma.1.2, type="normalized")  
I1<-!is.na(mean\_daily\_2018$logDailyfDOM)  
Efulla<-vector(length = length(mean\_daily\_2018$logDailyfDOM))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

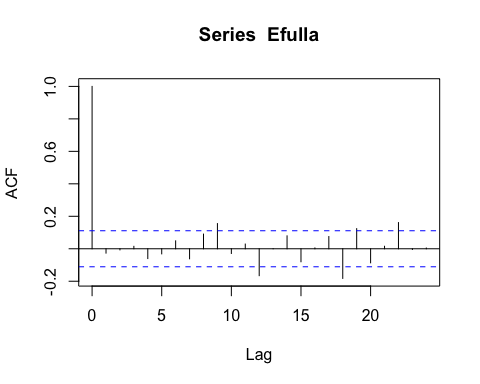
 Still autocorrelation in the later lags

# corARMA structure

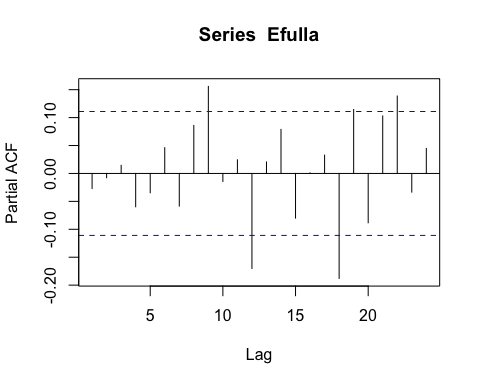
mean\_daily\_2018$site.ID <- as.factor(mean\_daily\_2018$site.ID)  
fDOM.mod.arma.2.2 <- gls(logDailyfDOM ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 2, q = 2))

# ACF plot

Ear1<-residuals(fDOM.mod.arma.2.2, type="normalized")  
I1<-!is.na(mean\_daily\_2018$logDailyfDOM)  
Efulla<-vector(length = length(mean\_daily\_2018$logDailyfDOM))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

 Still autocorrelation but this is the best one yet so lets interpret this model

# generalized linear hypotheses

site.ID.comp <- glht(fDOM.mod.arma.2.2, linfct = mcp(site.ID = "Tukey"))  
summary(site.ID.comp)

##   
## Simultaneous Tests for General Linear Hypotheses  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = logDailyfDOM ~ site.ID, data = mean\_daily\_2018, correlation = corARMA(form = ~julian |   
## site.ID, p = 2, q = 2), na.action = na.omit)  
##   
## Linear Hypotheses:  
## Estimate Std. Error z value Pr(>|z|)   
## FRCH - CARI == 0 0.2282 0.1422 1.605 0.244   
## MOOS - CARI == 0 0.6462 0.1430 4.520 <0.001 \*\*\*  
## MOOS - FRCH == 0 0.4180 0.1435 2.914 0.010 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## (Adjusted p values reported -- single-step method)

summary(fDOM.mod.arma.2.2)

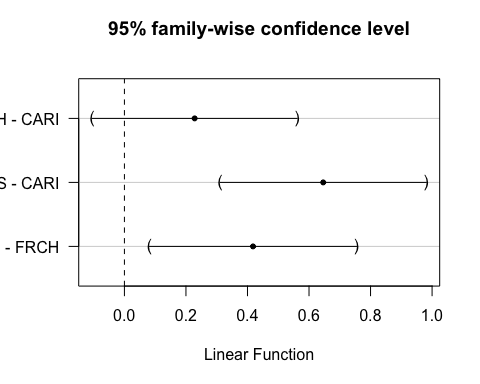
## Generalized least squares fit by REML  
## Model: logDailyfDOM ~ site.ID   
## Data: mean\_daily\_2018   
## AIC BIC logLik  
## -284.6658 -255.9678 150.3329  
##   
## Correlation Structure: ARMA(2,2)  
## Formula: ~julian | site.ID   
## Parameter estimate(s):  
## Phi1 Phi2 Theta1 Theta2   
## 1.1290767 -0.2991151 0.2559632 0.0758242   
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 4.251027 0.1001870 42.43093 0.0000  
## site.IDFRCH 0.228180 0.1421944 1.60470 0.1097  
## site.IDMOOS 0.646209 0.1429576 4.52029 0.0000  
##   
## Correlation:   
## (Intr) s.IDFR  
## site.IDFRCH -0.705   
## site.IDMOOS -0.701 0.494  
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -2.2633032 -0.6580045 -0.0231753 0.7443149 2.6930082   
##   
## Residual standard error: 0.360899   
## Degrees of freedom: 270 total; 267 residual

intervals(fDOM.mod.arma.2.2)

## Approximate 95% confidence intervals  
##   
## Coefficients:  
## lower est. upper  
## (Intercept) 4.05377000 4.2510270 4.4482840  
## site.IDFRCH -0.05178492 0.2281800 0.5081448  
## site.IDMOOS 0.36474181 0.6462095 0.9276772  
## attr(,"label")  
## [1] "Coefficients:"  
##   
## Correlation structure:  
## lower est. upper  
## Phi1 1.19767847 1.1290767 0.87493450  
## Phi2 -0.59100574 -0.2991151 0.06203559  
## Theta1 -0.08817157 0.2559632 0.67672874  
## Theta2 -0.13630247 0.0758242 0.28130267  
## attr(,"label")  
## [1] "Correlation structure:"  
##   
## Residual standard error:  
## lower est. upper   
## 0.2951396 0.3608990 0.4413101

plot(print(confint(site.ID.comp)))

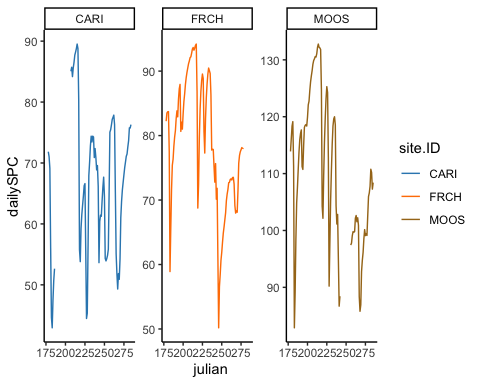
##   
## Simultaneous Confidence Intervals  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = logDailyfDOM ~ site.ID, data = mean\_daily\_2018, correlation = corARMA(form = ~julian |   
## site.ID, p = 2, q = 2), na.action = na.omit)  
##   
## Quantile = 2.3436  
## 95% family-wise confidence level  
##   
##   
## Linear Hypotheses:  
## Estimate lwr upr   
## FRCH - CARI == 0 0.22818 -0.10507 0.56143  
## MOOS - CARI == 0 0.64621 0.31117 0.98125  
## MOOS - FRCH == 0 0.41803 0.08181 0.75425

 This shows that FRCH and CARI are significantly different from CARI

### SPC

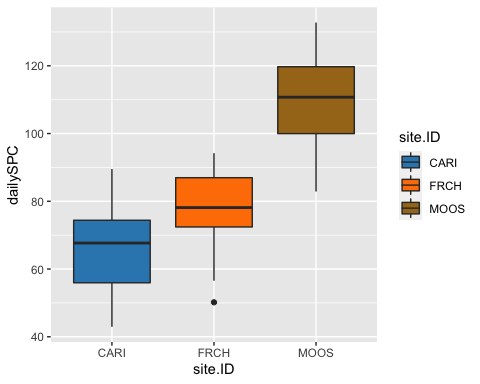
ggplot(mean\_daily\_2018, aes(x = julian, y = dailySPC, color = site.ID)) +  
 geom\_line() +  
 scale\_color\_manual(values=c("#3288BD","#FF7F00", "#A6761D")) +  
 facet\_wrap(~site.ID, scales = "free") +  
 theme\_classic()

## Warning: Removed 7 row(s) containing missing values (geom\_path).



ggplot(mean\_daily\_2018, aes(x = site.ID, y = dailySPC, fill = site.ID)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#3288BD","#FF7F00", "#A6761D"))

## Warning: Removed 29 rows containing non-finite values (stat\_boxplot).

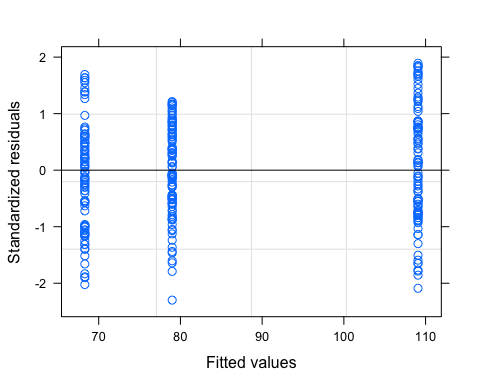


# corAR1 structure

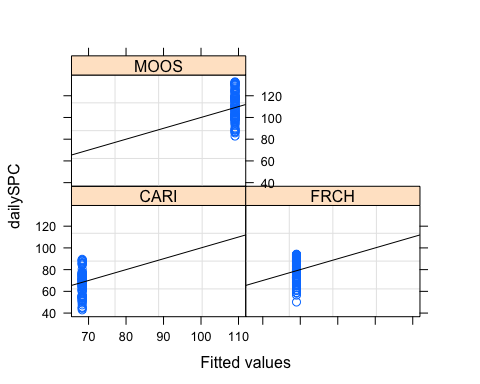
SPC.mod.ar1 <- gls(dailySPC ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

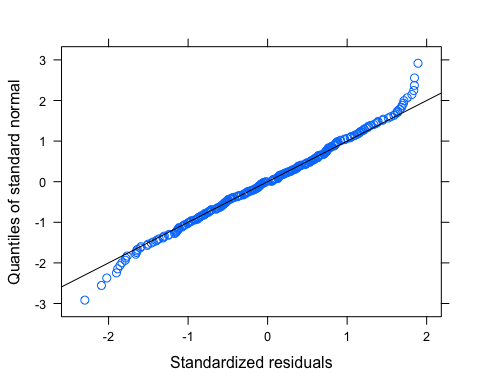
plot(SPC.mod.ar1, resid(., type = "p") ~ fitted(.), abline = 0)



plot(SPC.mod.ar1, dailySPC ~ fitted(.) | site.ID, abline = c(0,1))

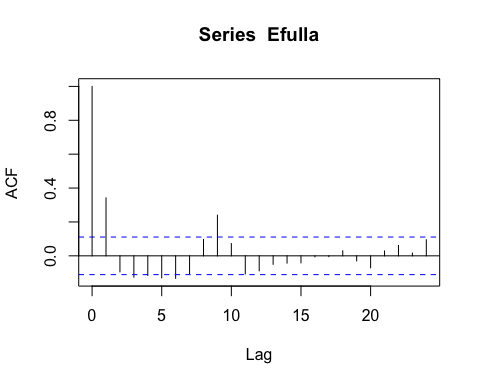


qqnorm(SPC.mod.ar1, abline = c(0,1))

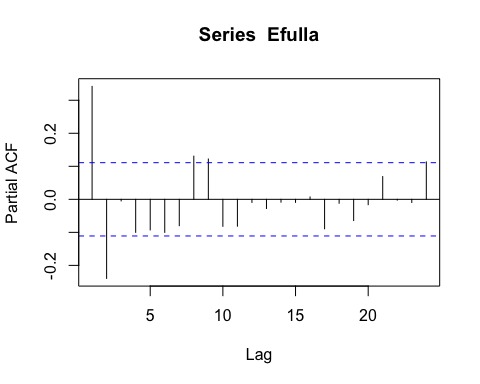


# ACF plot

Ear1<-residuals(SPC.mod.ar1, type="normalized")  
I1<-!is.na(mean\_daily\_2018$dailySPC)  
Efulla<-vector(length = length(mean\_daily\_2018$dailySPC))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

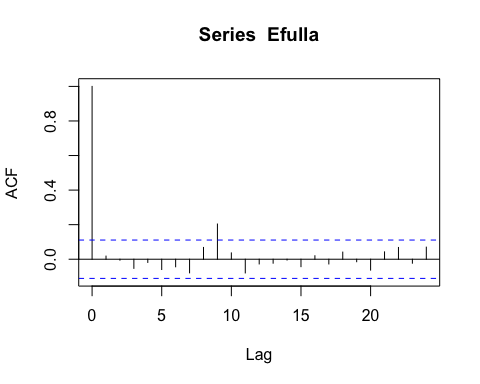
 AR1 is not adequately handling the temporal autocorrelation

# corARMA structure

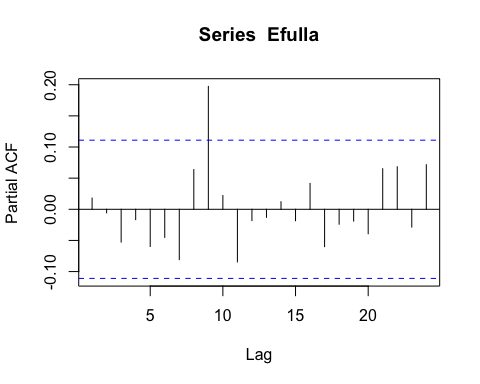
SPC.mod.arma.1.1 <- gls(dailySPC ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 1, q = 1))

# ACF plot

Ear1<-residuals(SPC.mod.arma.1.1, type="normalized")  
I1<-!is.na(mean\_daily\_2018$dailySPC)  
Efulla<-vector(length = length(mean\_daily\_2018$dailySPC))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

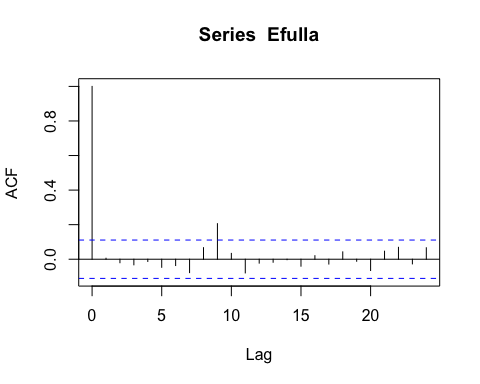
 Still autocorrelation at the 9th lag but that is it

# corARMA structure

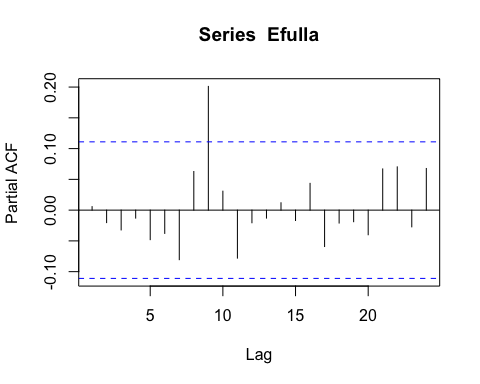
SPC.mod.arma.2.1 <- gls(dailySPC ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 2, q = 1))

# ACF plot

Ear1<-residuals(SPC.mod.arma.2.1, type="normalized")  
I1<-!is.na(mean\_daily\_2018$dailySPC)  
Efulla<-vector(length = length(mean\_daily\_2018$dailySPC))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

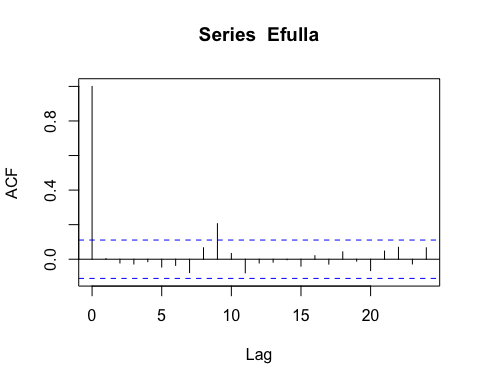
 Still that 9th lag

# corARMA structure

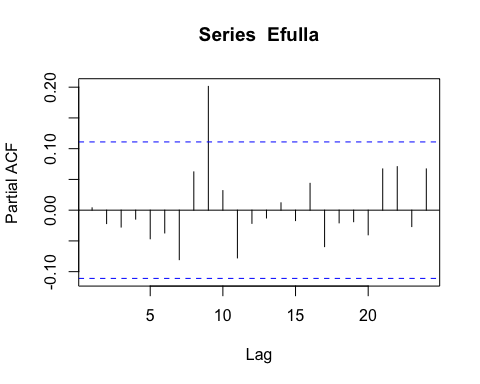
SPC.mod.arma.1.2 <- gls(dailySPC ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 1, q = 2))

# ACF plot

Ear1<-residuals(SPC.mod.arma.1.2, type="normalized")  
I1<-!is.na(mean\_daily\_2018$dailySPC)  
Efulla<-vector(length = length(mean\_daily\_2018$dailySPC))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

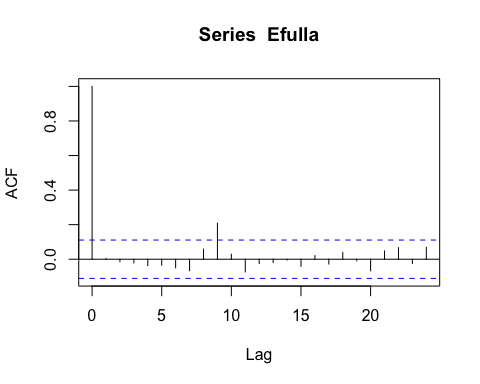
 Still autocorrelation in the later lags

# corARMA structure

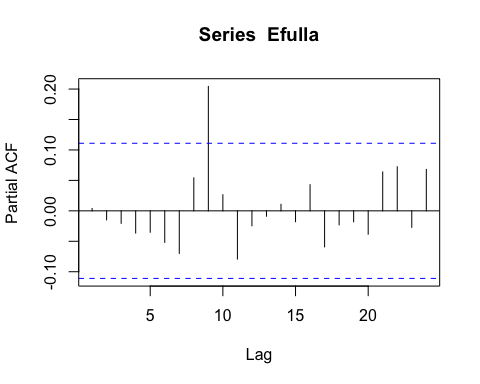
SPC.mod.arma.2.2 <- gls(dailySPC ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corARMA(form = ~ julian|site.ID, p = 2, q = 2))

# ACF plot

Ear1<-residuals(SPC.mod.arma.2.2, type="normalized")  
I1<-!is.na(mean\_daily\_2018$dailySPC)  
Efulla<-vector(length = length(mean\_daily\_2018$dailySPC))  
Efulla<-NA  
Efulla[I1]<-Ear1  
acf(Efulla, na.action=na.pass)



pacf(Efulla, na.action=na.pass)

 Still autocorrelation but this is the best one yet so lets interpret this model

# generalized linear hypotheses

site.ID.comp <- glht(SPC.mod.arma.1.1, linfct = mcp(site.ID = "Tukey"))  
summary(site.ID.comp)

##   
## Simultaneous Tests for General Linear Hypotheses  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = dailySPC ~ site.ID, data = mean\_daily\_2018, correlation = corARMA(form = ~julian |   
## site.ID, p = 1, q = 1), na.action = na.omit)  
##   
## Linear Hypotheses:  
## Estimate Std. Error z value Pr(>|z|)   
## FRCH - CARI == 0 11.072 5.812 1.905 0.137   
## MOOS - CARI == 0 41.498 5.771 7.190 <1e-04 \*\*\*  
## MOOS - FRCH == 0 30.426 5.729 5.311 <1e-04 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## (Adjusted p values reported -- single-step method)

summary(SPC.mod.arma.1.1)

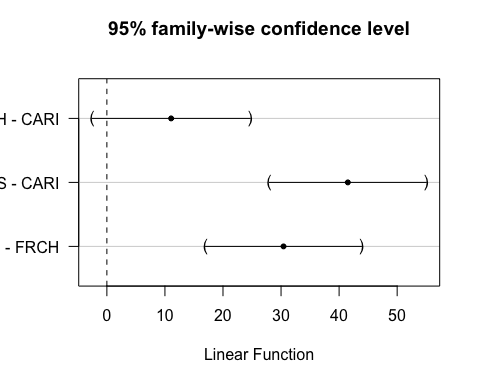
## Generalized least squares fit by REML  
## Model: dailySPC ~ site.ID   
## Data: mean\_daily\_2018   
## AIC BIC logLik  
## 1670.552 1692.36 -829.2758  
##   
## Correlation Structure: ARMA(1,1)  
## Formula: ~julian | site.ID   
## Parameter estimate(s):  
## Phi1 Theta1   
## 0.8448631 0.4651003   
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 67.80195 4.139394 16.379681 0.0000  
## site.IDFRCH 11.07197 5.812003 1.905019 0.0578  
## site.IDMOOS 41.49844 5.771485 7.190255 0.0000  
##   
## Correlation:   
## (Intr) s.IDFR  
## site.IDFRCH -0.712   
## site.IDMOOS -0.717 0.511  
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -2.36740465 -0.68326104 0.02168773 0.71463278 1.93689745   
##   
## Residual standard error: 12.12077   
## Degrees of freedom: 283 total; 280 residual

intervals(SPC.mod.arma.1.1)

## Approximate 95% confidence intervals  
##   
## Coefficients:  
## lower est. upper  
## (Intercept) 59.6536712 67.80195 75.95024  
## site.IDFRCH -0.3687935 11.07197 22.51274  
## site.IDMOOS 30.1374354 41.49844 52.85945  
## attr(,"label")  
## [1] "Coefficients:"  
##   
## Correlation structure:  
## lower est. upper  
## Phi1 0.7589000 0.8448631 0.9018861  
## Theta1 0.3534746 0.4651003 0.5636665  
## attr(,"label")  
## [1] "Correlation structure:"  
##   
## Residual standard error:  
## lower est. upper   
## 9.669261 12.120768 15.193820

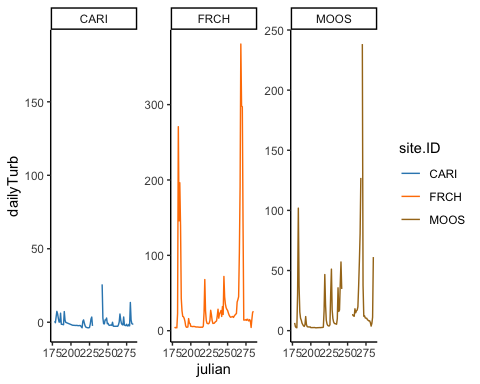
plot(print(confint(site.ID.comp)))

##   
## Simultaneous Confidence Intervals  
##   
## Multiple Comparisons of Means: Tukey Contrasts  
##   
##   
## Fit: gls(model = dailySPC ~ site.ID, data = mean\_daily\_2018, correlation = corARMA(form = ~julian |   
## site.ID, p = 1, q = 1), na.action = na.omit)  
##   
## Quantile = 2.343  
## 95% family-wise confidence level  
##   
##   
## Linear Hypotheses:  
## Estimate lwr upr   
## FRCH - CARI == 0 11.0720 -2.5456 24.6895  
## MOOS - CARI == 0 41.4984 27.9758 55.0211  
## MOOS - FRCH == 0 30.4265 17.0036 43.8493

 This shows that MOOS is significantly different than CARI and FRCH

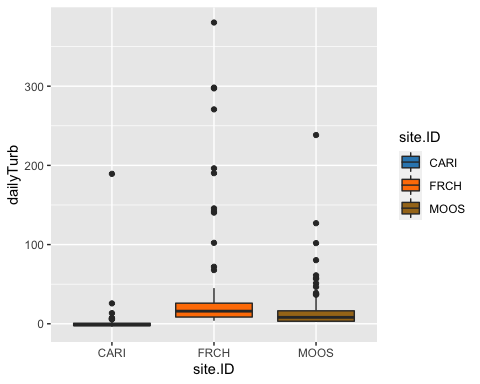
### Turb

ggplot(mean\_daily\_2018, aes(x = julian, y = dailyTurb, color = site.ID)) +  
 geom\_line() +  
 scale\_color\_manual(values=c("#3288BD","#FF7F00", "#A6761D")) +  
 facet\_wrap(~site.ID, scales = "free") +  
 theme\_classic()



ggplot(mean\_daily\_2018, aes(x = site.ID, y = dailyTurb, fill = site.ID)) +  
 geom\_boxplot() +  
 scale\_fill\_manual(values=c("#3288BD","#FF7F00", "#A6761D"))

## Warning: Removed 25 rows containing non-finite values (stat\_boxplot).

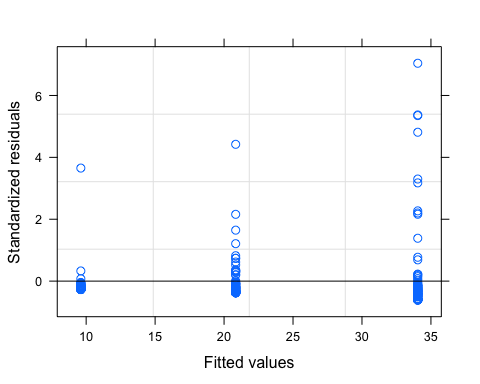


# corAR1 structure

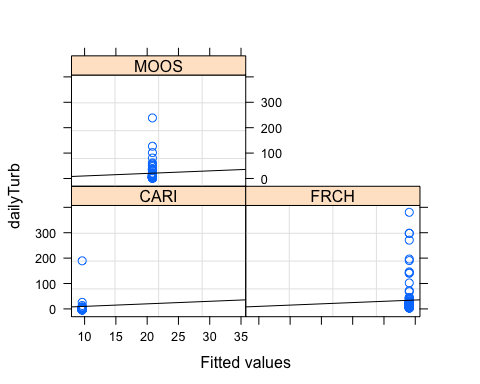
turb.mod.ar1 <- gls(dailyTurb ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

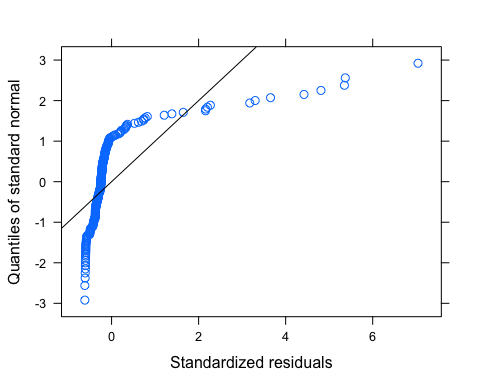
plot(turb.mod.ar1, resid(., type = "p") ~ fitted(.), abline = 0)



plot(turb.mod.ar1, dailyTurb ~ fitted(.) | site.ID, abline = c(0,1))



qqnorm(turb.mod.ar1, abline = c(0,1))

 Looks like we have lots of outliers here but our normality isnt good either so lets log transform first and then investigate outliers

# corAR1 structure

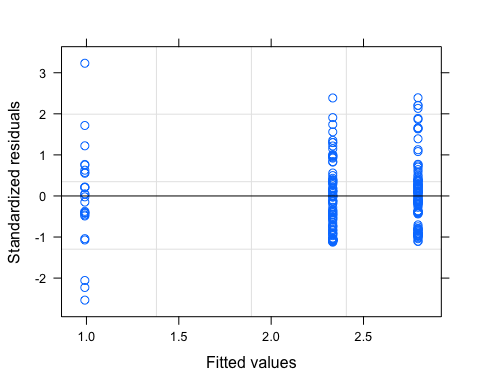
mean\_daily\_2018$logDailyTurb <- log(mean\_daily\_2018$dailyTurb)

## Warning in log(mean\_daily\_2018$dailyTurb): NaNs produced

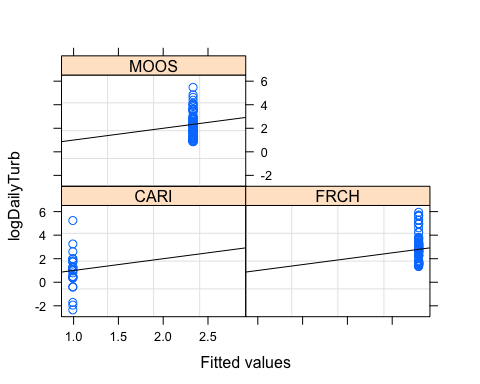
turb.mod.ar1.log <- gls(logDailyTurb ~ site.ID,   
 na.action = na.omit,   
 data = mean\_daily\_2018,   
 correlation = corAR1(form = ~ julian|site.ID))

# diagnostic plots

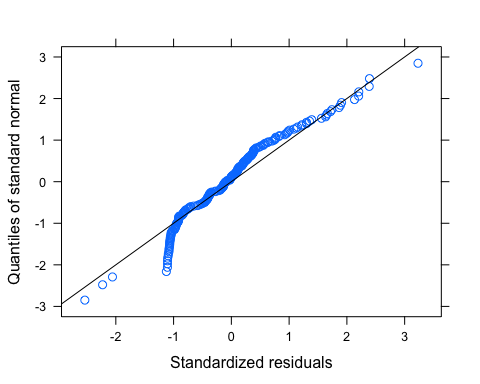
plot(turb.mod.ar1.log, resid(., type = "p") ~ fitted(.), abline = 0)



plot(turb.mod.ar1.log, logDailyTurb ~ fitted(.) | site.ID, abline = c(0,1))



qqnorm(turb.mod.ar1.log, abline = c(0,1))

 It looks like we may need to further transform Turbidity