ARIMA_Workshop

Kristin Briggs

1/24/2020

Objectives

Evaluate temporal patterns of daily mangrove NEE to determine the influence of seasonality and other trends, as well as stochastic events, and the magnitude of their influence on the long-term patterns. Also, improve the model performance using explanatory variables.

Methods

Examine and plot a time series of daily mangrove NEE, then remove outliers and decompose the time series by extracting the seasonal, trend and irregular behavior in the data. Tests for assumptions of stationarity (constant variance) in the dataset are conducted (ADF). Autocorrelation function is also used to test for time lag correlation. Independence of time series data was tested using Ljunk-Box test. Reduced autocorrelation by widening the moving average window, and select best fit models using AIC. Compare the influence of salinity and PAR on model fit.

Site Information

An LTER mangrove site in the Everglades along the Taylor Slough in the short/dwarf mangrove area, toward the lower end of the mangrove productivity gradient. https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/TSPh7info.jpg (https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/TSPh7info.jpg) TS/Ph-7b Site

Statistical Analysis

Use ARIMA package to adjust and select best fit models. Compare the influence of salinity and PAR on model fit. Minimize AIC to improve model fit.

Results

Testing Assumptions: ADF confirms stationary data (p-value=0.01), Ljung-Box test confirmed independence (p-value =0.23). The comparison of the ARIMA generated model (nee1), to the adjusted ARIMA model (nee2) resulted in AIC values of 706.09 and 704.77, respectively.
Fitting independent variables to explain time series patterns: Salinity: Including salinity in the model (nee3) did not improve it's fit, nee3 AIC = 705.41, compared to our previous best model (nee2) the model. And yet, including an extreme salinity (over 25n ppt) indicator did improve the model fit (AIC of 400.77). https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/timeSal.jpg

(https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/timeSal.jpg) https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/ARIMAtable1.jpg (https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/ARIMAtable1.jpg)

PAR: Using the photosynthetically active radiation as a, in the model improved it's fit to NEE patters (PAR driven model AIC = 683.68). https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/timePAR.jpg (https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/timePAR.jpg) https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/ARIMAtable2.jpg (https://github.com/KLBriggs/Workshop-Biology-Spring2020/blob/master/images/ARIMAtable2.jpg)

Discussion

Since mangroves are coastal species with tolerance for a range of salinities, its likely they have stable growth despite fluctuations in salinity. This also helps explain why extreme salinity was a better predictor of mangrove NEE, since it may reflect saline conditions outside mangrove tolerance range and reduce productivity. Photosynthetically active radiation (PAR) was a better predictor of mangrove ecosystem exchange than extreme salinity, which makes sense considering the relationship between PAR and productivity, but this may be influenced by some collinearity between these variables. Also PAR only indirectly reflects respiration rates which would be better evaluated looking at the nighttime NEE data. Further analysis using night data and comparing which models (explanatory variables) fit the daytime and nighttime data the best may help parcel how productivity/respiration are influenced in variables differently, ultimately determining the systems carbon balance and NEE.

library(zoo) library(tseries) library(forecast) library(xts) nee -> ts(mangroves$nee, start = 1, frequency = 30) nee < -ts(mangroves nee, start= 1, frequency=30) par(mfrow=c(1,1), mai=c(0.25,0.8,0.1, 0.1)) plot(nee, typ="l", ylab= "NEE", xlab="") plot(nee) lines(tsclean(nee), col="red") nee <- tsclean(nee) nee.d <- decompose(nee, 'multiplicative') plot(nee.d) adf.test(nee) acf(nee, lag.max=45)

arima.nee1 <-auto.arima(nee, trace=TRUE) tsdisplay(residuals(arima.nee1), lag.max=45) arima.nee2 <-arima(nee , order=c(10,1,3), seasonal= list(order=c(2,0,2)))

tsdisplay(residuals(arima.nee2), lag.max= 30)

AIC(arima.nee1, arima.nee2)

par(mfrow=c(1,1)) plot(nee , typ="l"); lines(fitted(arima.nee2),col="red") checkresiduals(arima.nee2, lag=36)

par(mfrow=c(1,1)) plot(nee , typ="l"); lines(fitted(arima.nee2),col="red")

plot(forecast(arima.nee2, h=30))

sal <- ts(mangroves$salinity.max, start= 1, frequency=30)

par(mfrow=c(1,1), mai=c(0.25,0.8,0.1, 0.1)) plot(sal , typ="l", ylab= "Salinity", xlab="")

plot(sal , typ="l", ylab= "Salinity", xlab="") lines(tsclean(sal) , col="red")

sal <- tsclean(sal)

```
sal.d <- decompose(sal, ‘multiplicative’) plot(sal.d)

adf.test(sal) adf.test(diff(sal))

ccf( diff(sal),nee, na.action = na.pass, lag.max=40, plot=TRUE)

arima.nee3 <-auto.arima(nee, xreg=c(diff(sal),0), trace=TRUE)

AIC(arima.nee2, arima.nee3 )

sal.i <- sal sal.i[sal.i < 25 ]<- 0 sal.i[sal.i >= 25 ]<- 1 plot(sal.i)

arima.nee4 <-auto.arima(nee, xreg=sal.i, trace=TRUE)

AIC(arima.nee2,arima.nee4 )

checkresiduals(arima.nee4, lag=36)

par(mfrow=c(1,1)) plot(nee , typ=“l”); lines(fitted(arima.nee4),col=“red”)
```

use PAR as explanatory variable, challenge for better model

```
photo <- ts(mangroves$par, start= 1, frequency=30)

par(mfrow=c(1,1), mai=c(0.25,0.8,0.1, 0.1)) plot(photo , typ=“l”, ylab= “PAR”, xlab="")

plot(photo, typ=“l”, ylab= “PAR”, xlab="“) lines(tsclean(photo) , col=”red")

photo <- tsclean(photo)

photo.d <- decompose(photo, ‘multiplicative’) plot(photo.d)

adf.test(photo)

adf.test(diff(photo))

ccf( diff(photo),nee, na.action = na.pass, lag.max=40, plot=TRUE)

arima.nee.p1 <-auto.arima(nee, xreg=c(diff(photo),0), trace=TRUE)

AIC(arima.nee4,arima.nee.p1) checkresiduals(arima.nee.p1, lag=36)
```