Workshop 3: Timeseries Models

Kristy Sullivan

1/31/2020

Objectives

The primary objective of this analysis is to create an Autoregressive Integrated Moving Average (ARIMA) model of daily net ecosystem exchange (NEE) rates for a mangrove scrub forest along the coast of South Florida. The model will be optimized by identifying the best-fitting predictive parameters (i.e. PAR, salinity, temperature, etc), identifying seasonality components, testing for stationarity (Dickey-Fuller), and applying the appropriate transformations to optimize the fit of the model.

Methods

Site Information

The data used in this exercise were collected from Florida Coastal Everglades Long-term Ecological Research (FCE LTER) site TS/Ph-7 (*Figure 1*). This is a mangrove dominated site along the Taylor River slough within the Everglades National Park. Data were obtained from an eddy covariance tower at the site.

Figure 1: Map of the FCE LTER sites, note site TS/Ph-7.

Statistical Analysis

Timeseries data for NEE were first visualized and then corrected for outliers using the tsclean function. The corrected data were then deconstructed into observed, trend, seasonal, and random plots to better understand the behavior of the data. A Dickey-Fuller test for stationarity was performed to determine if an ARIMA model would be an appropriate fit. The data were then tested for autocorrelation. The ARIMA model was fitted using the auto.arima function to generate optimal p (lag order), d (degree of differencing), and q (moving average) values. The fit of the auto-generated ARIMA values were tested by determining if the residuals were normally distributed (ACF & PACF plots). Adjustments were made to parameter values if needed. Potential models were compared to each other via comparison of AIC values. THe model with the lowest AIC was selected as the best fitting model. This model was then used to create a 30 day forecast of mangrove ecosystem NEE.

Results (minimum of 1 plot and one table)

Five potential ARIMA models for NEE were created using different predictive parameters (*Table 1*). Of the five potential models, arima.nee5 which was modeled based on PAR was the best predictor for mangrove NEE (AIC = 683.68) (*Figure 2*). No transformations were needed to improve the fit of this model as there was no lag time (*Figure 3*). Other models with different parameters or adjusted p, d, and q values did not provide a good fit for observed NEE (AIC(s) ~ 700).

Table 1: Potential ARIMA models and respective model parameters, N value, and AIC value.

Model	Parameter	N	AIC
arima.nee1	NEE	9	703.09
arima.nee2	NEE w/ corrected values	18	704.77
arima.nee3	Max salinity	12	706.19
arima.nee4	Extreme salinity (>25 ppt)	9	700.57
arima.nee5	PAR	9	683.68

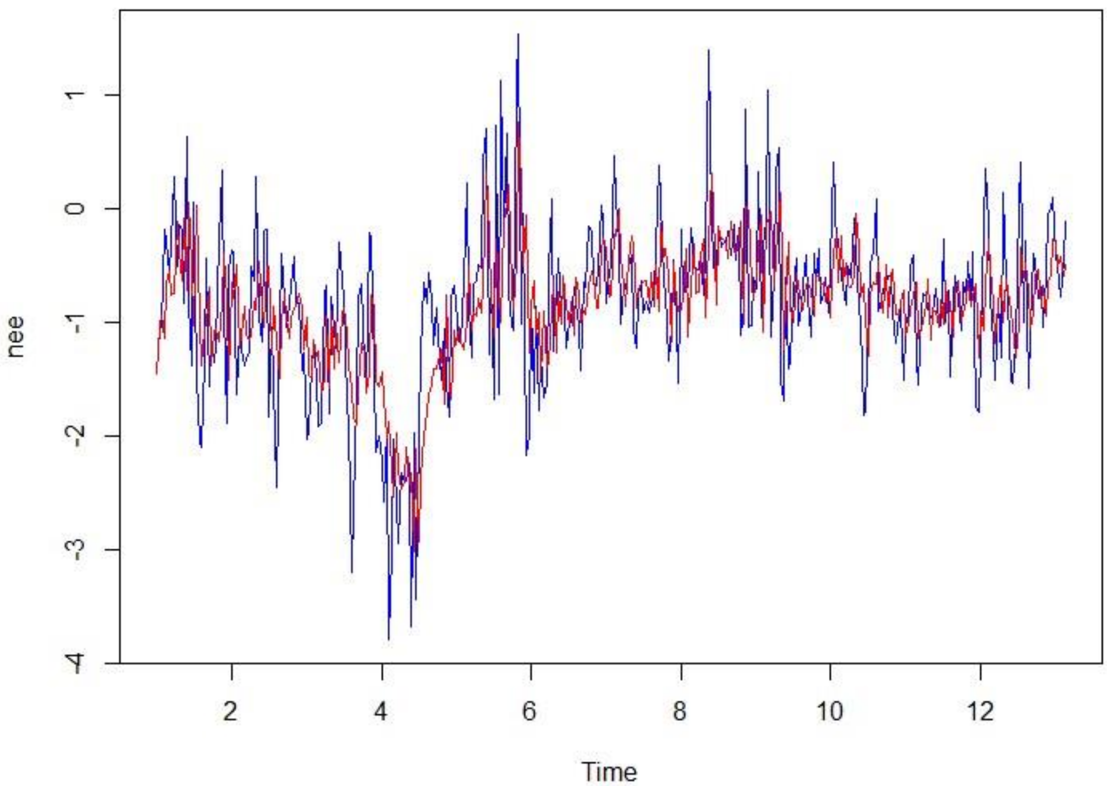

Figure 2: NEE timeseries with outliers removed (blue) and fitted arima.nee5 model using PAR as the predictor (red). The X-axis represents NEE in g C m^2 day^−1, Y axis represents time in months (1 = Jan, 12 = Dec.

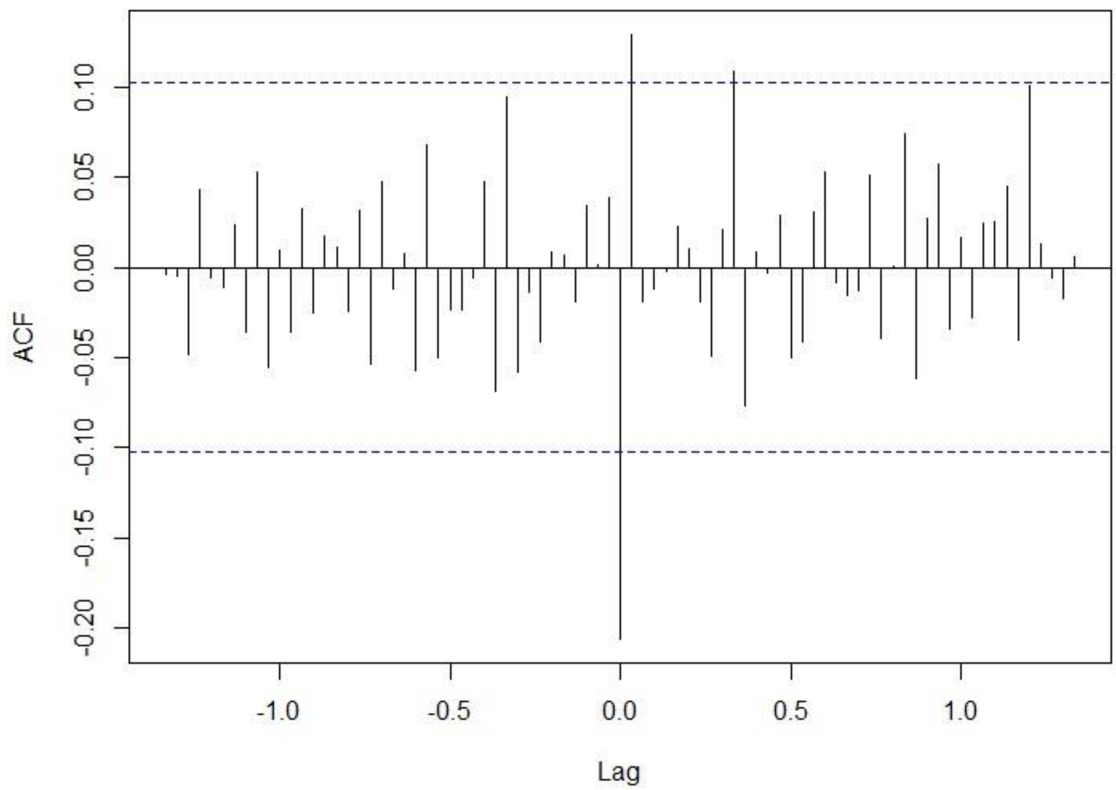

Figure 3: ACF plot indicating no significant lags in the modeled data.

Discussion

Via comparison of the AIC values for five different ARIMA models, PAR was determined to be the best-fitting predictive parameter of daily mangrove ecosystem NEE. Photosynthesis is greatly influenced by light availability and intensity, therefore, PAR is a suitable predictor of net ecosystem exchange of carbon. Other parameters such as salinity or temperature do not perform as well in this model. In order to test the fit of this model across multiple years, a longer dataset of mangrove ecosystem NEE will be needed. If the AIC holds or improves, we will then be able to make more accurate forecasts of NEE for the South Florida mangrove ecosystem. In years of high storm activity and mangrove destruction, PAR may no longer be a suitable predictor. Instead, parameters such as percent canopy cover may provide a basis for the best fitting ARIMA model.