Exploring ENSO Dynamics via Time Series Analysis of Monthly Niño 3.4 Values from 1950 to 2023

Lanz Railey A. Fermin, Francis Ira S. Velaque

Abstract

El Niño–Southern Oscillation (ENSO) is a climate phenomenon involving fluctuating wind and sea surface temperatures in the Pacific Ocean. Linear time series models have been integral components of ENSO forecasting. The paper aimed to employ such method on monthly Niño 3.4 values, a sea surface temperature metric, from 1950 to 2023 in an attempt analyze trends and potentially forecast future ENSO events. Ljung-Box and ADF tests were used to check for autocorrelation and stationarity, respectively. Seasonality was apparent and second-order differencing was performed to fix this issue. Using the ACF and PACF plots, eleven possible models were fitted. The best-fit model was an ARMA$(3, 3)$ with the lowest AIC score 2323.52 and is given by $W_t = -1.0458W_{t-1} - 0.2286W_{t-2} - 0.1819W_{t-3} + Z_t - 0.9833Z_{t-1} - 0.9925Z_{t-2} + 0.9774Z_{t-3}$. An analysis of the model revealed the dependence of Niño 3.4 values with previous observations. Its close relation with another ENSO metric was also discovered, as explained by the lags of the ARMA model. Despite this, the innate seasonality of the data calls for the utilization of more sophisticated models to yield better potential in predicting ENSO events.

1 Introduction

The El Niño–Southern Oscillation (ENSO) is a coupled ocean-atmosphere climatological disturbance that originated in the Pacific Ocean. It is described as the most energetic interannual variation of climate system on Earth, given its characterization as an irregular, yet periodic fluctuation of wind and sea surface temperatures that occur every three to seven years [1]. Alternating warm and cool events of ENSO, commonly known as El Niño and La Niña, respectively, frequently lead to extensive climate anomalies, trigger intensification of tropical cyclone genesis, and affect several facets of human activity such as agriculture, energy production, and water resources [2].

The strength and phase of a given ENSO event are usually defined using an index. However, numerous such indices exist and there is no consensus as to which is the best quantitative measure to accurately define its onset and classify its phase, given the innate complexity of ocean-atmosphere dynamics [3]. Some of the usual and known ENSO indicators include air pressure metrics like Southern Oscillation Index, wind-based markers like 200-mb zonal winds, and more importantly, sea surface temperature (SST) indices like Niño 1+2, Niño 3, Niño 4, and Niño 3.4 [4]. In modern climatology, the use of SST indices and their anomalies—defined as the magnitude of departure from the average SST—gained increased viability mainly due to their consistency of available data at fixed time intervals.

Poor prediction of ENSO events can lead to disastrous consequences that affect not only the normal flow of land and aquatic ecosystems, but also anthropogenic functions such as the world economy. For example, conventional models were not able to foresee the magnitude of the El Niño event in 2015–2016 [5]. This became one of the strongest El Niño on record and caused widespread climate impacts that matched peak historical patterns like temperature and precipitation [6]. Hence, the need for continuous analysis and forecasting of ENSO events remains essential.

Submission for MATH 62.2 Time Series and Forecasting, AY 2022-2023 Second Semester

Linear models have long been part of routine predictions done by climate centers worldwide. Despite their limitations regarding the dynamic nature of climatology and randomness of some variables, such statistical models could still depict much of the important ocean-atmosphere interactions using observed data, including ENSO. For instance, Serykh & Sonechkin (2021) explored the relationship between ENSO and the Global Atmospheric Oscillation, and used the latter's index values to predict potential ENSO events with a minimum 12-month lead time [7].

Niño 3.4, along with other Niño indices, measures SST averaged within different points within or near the equatorial Pacific Ocean. It has been utilized to develop homogeneous, observational-based SST climate records over long periods of time, allowing for a statistical framework to be used in analyzing ENSO events [8]. Further, it has been touted as a good metric for forecasting ENSO in the tropical Pacific due to its reliability in representing SST variations associated with ENSO phases. While it is not the lone indicator, its consistency and reliability has made it a standard metric.

This paper attempts to analyze trends of ENSO fluctuations by employing a linear time series model in the observed Niño 3.4 index values. By identifying a sufficiently-fit model to the observed index values, we can derive insights that capture dynamics of potential ENSO events in relation to observed sea surface temperatures. Understanding the temporal patterns present in Niño 3.4 is crucial in identifying its predictive power in forecasting future ENSO events.

2 Methods

2.1 Data

The dataset used in this paper came from the Niño 3.4 index values, as provided by the US National Weather Service - Climate Prediction Center of the National Oceanic and Atmospheric Administration (NOAA) [9]. The values are derived from the Extended Reconstructed Sea Surface Temperature version 5, which are raw data of SSTs collected by NOAA from various measuring structures such as ships, buoys, and satellites in the Niño 3.4 area, spanning from coordinates 5N-5S, 170W-120W.

The Niño 3.4 dataset covers SST observations from January 1950 to December 2023, with a total of 888 data points. A plot of the dataset can be seen in Figure 1.

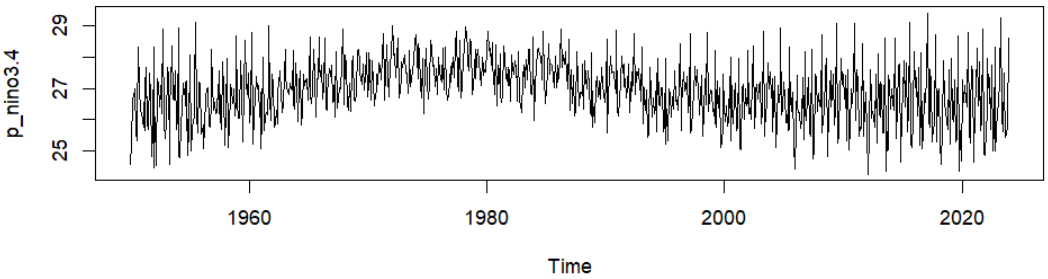

Figure 1: Plot of $\{X_t\}$.

It is important to note that previous research has verified some important properties of the Niño 3.4 dataset values, such as its high correlation with other climatological variables and ENSO indicators like sea level pressure and precipitation in the tropical Pacific [10]; its near-Gaussian distribution [11]; and its characteristic as the most skillful region in predicting ENSO events [12].

2.2 Theorems

2.2.1 Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller (ADF) Test is commonly used to test the stationarity of a time series. The null hypothesis states that the data is not stationary. In order to perform an analysis of the time series data, we aim to reject the null hypothesis. That is, the time series data should be stationary.

2.2.2 Ljung-Box Test

The autocorrelation of time series data can be checked using the Ljung-Box Test. The null hypothesis claims that the data is not autocorrelated. For us to fit a model in the time-series data, we aim to reject the null hypothesis. That is, we want to ensure that there is a correlation between data points in the time series at different time intervals.

2.2.3 Autoregressive process of order p

By definition, a time series $\{X_t\}$ is an autoregressive process of order p, denoted by AR(p), if

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t$$

where $\{Z_t\} \sim WN(0, \sigma^2)$ and Z_t is uncorrelated with X_s for each $s < t$. To identify the order p of an AR time series, we use the sample partial autocorrelation function (PACF). For an AR(p) series, the sample PACF "cuts off" at lag p.

2.2.4 Moving average process of order q

A time series $\{X_t\}$ is a moving-average process of order q, denoted by MA(q) if

$$X_t = \mu + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $Z_t \sim WN(0, \sigma^2)$. We identify the order q for an MA process using the autocorrelation function (ACF) where the sample ACF "cuts off" at lag q.

2.2.5 Autoregressive moving average process of orders p and q

A time series $\{X_t\}$ is ARMA(p, q) if it is stationary and

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $\phi_p \neq, \theta_q \neq 0$, and $\sigma^2 > 0$. The parameters p and q are called the autoregressive and moving average orders, respectively.

If X_t has a nonzero mean μ, we set $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$ and write the model as

$$X_t = \alpha + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $Z_t \sim WN(0, \sigma^2)$.

2.2.6 Akaike Criterion Information (AIC)

Given that multiple models are generated from the previous processes, AIC is commonly utilized to select the best-fit model. Here, lower AIC scores indicate a better fit for the time series data.

2.3 Assumptions

Let the Niño 3.4 be a time series data denoted by $\{X_t\}$. We fitted a model with a small number of coefficients since these models are easier to understand and explain [13]. For this specific paper, we restricted p and q to be both at most 3. For the ADF test and the Ljung-Box Test, we used a 95% level of confidence. Furthermore, we used differencing to remove the trend or seasonal components from the data by repeated differencing at one or more lags in order to generate a noise sequence [14].

2.4 Data Exploration

We used RStudio (v2024.04.1) and the libraries `TSA`, `forecast`, and `tseries` for the data exploration and modeling. To model the time series data, we need to check for stationarity and autocorrelation.

We first use the ADF test for stationarity. Due to RStudio limitations, there is no exact output for the p-value. However, we are given a prompt that the p-value is smaller than the printed p-value which is 0.01. Since p-value $< 0.01 < 0.05$, it follows that we reject the null hypothesis. That is, Niño 3.4 time series is stationary.

We then use the Ljung-Box test to check for autocorrelation. The p-value is $3.446 \times 10^{-6} < 0.05$ which leads us to reject the null hypothesis. Hence, we can conclude that the time series data is autocorrelated. With this, we can now proceed with order determination.

2.5 Order determination

In this section, we determined possible AR(p), MA(q), and ARMA(p, q) models where both p and q are at most 3. Figure 2 shows the plots of the PACF and ACF of the Niño 3.4 time series.

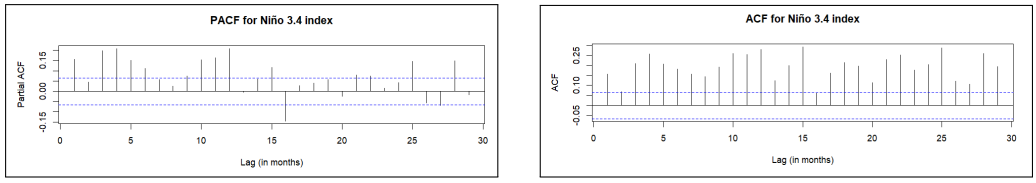

Figure 2: PACF and ACF plot of $\{X_t\}$.

Based on the ACF plots, there are signs of seasonality. This is one of the instances where the ADF test fails to account for seasonality [15]. We tried log transformation but it did not solve the problem of seasonality. Hence, we utilized differencing with `d=1`. Let $Y_t = (1 - B)X_t$. The plots for its PACF and ACF are seen in Figure 3.

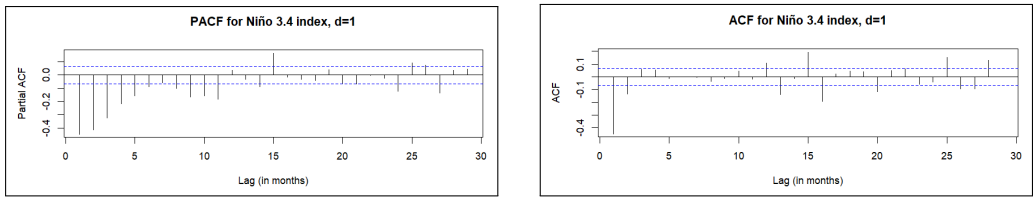

Figure 3: PACF and ACF plot of $\{Y_t\}$.

With `d=1`, the presence of seasonality is still apparent in the PACF plot. Hence, we continued with `d=2`. Let $W_t = (1 - B)^2 X_t = (1 - B)Y_t$. The new PACF plot is seen in Figure 4 below.

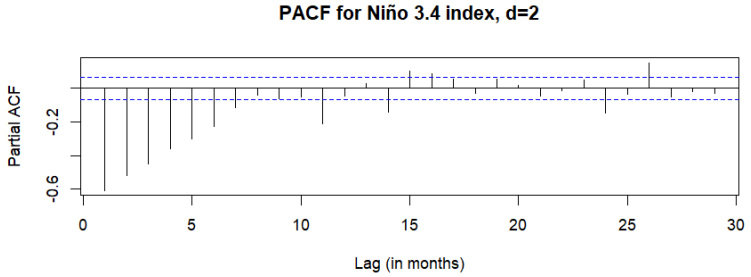

Figure 4: PACF plot of $\{W_t\}$.

Notice that the PACF plot cuts off at lags $p = 1, 2, 3$. Figure 5 shows the new ACF plot. Observe that the ACF plot cuts off at lags $q = 1, 2$.

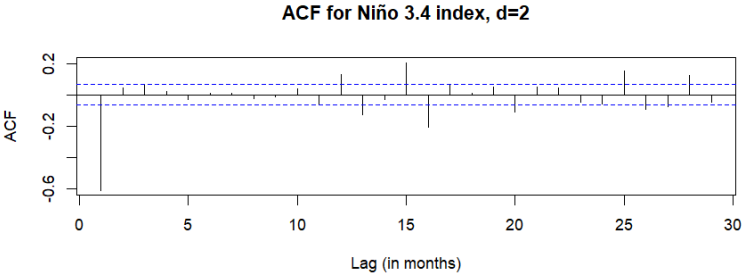

Figure 5: ACF plot of $\{W_t\}$.

Since second-order differencing was done, we need to do the ADF test and Ljung-Box test again to ensure stationarity and autocorrelation. For the ADF test, the p-value is smaller than the printed p-value at 0.01 < 0.05. Hence, we reject the null hypothesis. That is, W_t is stationary. The Ljung-Box test, on the other hand, returns a p-value of $2.2 \times 10^{-16} < 0.05$ which leads us to reject the null hypothesis. We can then conclude that W_t is also autocorrelated.

Given that W_t is both stationary and autocorrelated, we can now proceed fitting the models. The possible AR(p), MA(q), and ARMA(p, q) models consist of the combination of the p and q values derived from the corresponding PACF and ACF plots.

3 Results and Discussion

3.1 Model Fitting

In this section, we determine the best-fit model for our Niño 3.4 dataset. Recall that we have 11 candidate AR, MA, and ARMA models, with varying lag orders.

Using the `Arima()` function of the library `forecast` in `R`, we obtain the coefficients for the equation of each model, their respective information criteria like the AIC, Akaike Information Corrected Criterion, Bayesian Information Criterion, and other metrics such as the log-likelihood and variance of residuals.

Table 1 summarizes the AIC score for each candidate model. Note that we primarily focus on comparing AIC as the criterion for identifying the best-fit model.

Table 1: Summary of AIC scores.

Model	AIC
MA(1)	2965.027
MA(3)	2344.487
AR(1)	3488.556
AR(2)	3211.895
AR(3)	3004.528
ARMA$(1, 1)$	2769.756
ARMA$(2, 1)$	2600.537
ARMA$(3, 1)$	2498.902
ARMA$(1, 3)$	2341.707
ARMA$(2, 3)$	2329.205
ARMA$(3, 3)$	2323.515

We see that the model with the lowest AIC is ARMA$(3, 3)$. And so, it can be said that our Niño 3.4 time series data, W_t, is an ARMA$(3, 3)$ process. We proceed by writing the model in functional form. Table 2 lists the coefficients of the model, obtained from the same `Arima()` function called earlier.

Table 2: Coefficients of the ARMA$(3, 3)$ model.

ar1	ar2	ar3	ma1	ma2	ma3	mean
-1.0458	-0.2286	-0.1819	-0.9833	-0.9925	0.9774	0

And so, the equation ARMA$(3, 3)$ model is:

$$W_t = -1.0458W_{t-1} - 0.2286W_{t-2} - 0.1819W_{t-3} + Z_t - 0.9833Z_{t-1} - 0.9925Z_{t-2} + 0.9774Z_{t-3}$$

3.2 Model Interpretation

The model equation we have obtained for the best-fit model in the previous subsection simply demonstrates that the observed values of W_t is somehow dependent not only on the past three observations, but also on the error terms of those past three months.

Given that the model is ARMA$(3, 3)$, it might be of interest to examine the appropriateness of having lags of the AR and MA component both set to 3 by defining a metric closely related to the Niño 3.4 SSTs: the Oceanic Niño Index (ONI). It is the most commonly used index to formally designate an ENSO event at a given point in time. The ONI is computed as the three-month running mean of the SST anomalies in the Niño 3.4 region. An El Niño or La Nina phase is declared if the observed ONI values exceed the $+0.5^{\circ}$C or -0.5°C threshold for at least five consecutive months [16].

From this definition, we can draw a potential justification behind the optimality of having lags $(3, 3)$ for the ARMA model. As Niño 3.4 data is used, in itself, as an index to forecast potential ENSO events, and since the ONI is inherently dependent on the past three observations, it can be said that the best-fit ARMA$(3, 3)$ model is aligned with the underlying dynamics between Niño 3.4 and ONI, and their respective use as indices to forecast ENSO events.

The nature of the best-fit model as an ARMA$(3, 3)$ may also be indicative of the inherent complexities embedded behind ENSO forecasting processes. Previous studies have demonstrated that forecasting models are only able to predict the occurrence of intense and regular El Niño events; yet the transition to Neutral phases and sudden drops to La Nina events remained as inaccurate predictions [17, 18]. Moreover, an inverse relationship between the length of lead time and accuracy of the forecasting model was also discovered [19].

One of the most documented issues with regard to ENSO forecasting is the spring predictability barrier, in which the prediction skill of a given model significantly decreases during the boreal spring (months of March-April-May) [20]. This further demonstrates the innate seasonality of the Niño 3.4 data – an issue that was only addressed in the current study using differencing.

To conclude, we recommend exploring more advanced models to fit the existing Niño 3.4 data such as `SARIMA` and `SARIMAX`. These models can be used to better account for the seasonality of the data, potentially improving the reliability of ENSO forecasting using time series methods.

References

[1] C. F. Ropelewski and M. S. Halpert. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Monthly weather review*, 115(8):1606–1626, 1987.

[2] M. L'Heureux, A. Levine, M. Newman, C. Ganter, J. Luo, M. Tippett, and T. Stockdale. El niño southern oscillation in a changing climate. *ENSO Prediction*, pages 227–246, 2020.

[3] D. E. Hanley, M. A. Bourassa, J. J. O'Brien, S. R. Smith, and E. R. Spade. A quantitative evaluation of ENSO indices. *Journal of Climate*, 16(8):1249–1258, 2003.

[4] S. Mir and M. A. Arbab. ENSO dataset & comparison of deep learning models for ENSO forecasting. *Earth Science Informatics*, pages 1–6, 2024.

[5] M. L. L'Heureux, K. Takahashi, A. B. Watkins, A. G. Barnston, E. J. Becker, T. E. Di Liberto, and A. T. Wittenberg. Observing and predicting the 2015/16 El Niño. *Bulletin of the American Meteorological Society*, 98(7):1363–1382, 2017.

[6] A. Santoso, M. J. McPhaden, and W. Cai. The defining characteristics of ENSO extremes and the strong 2015/2016 El Niño. *Reviews of Geophysics*, 55(4):1079–1129, 2017.

[7] I. V. Serykh and D. M. Sonechkin. El Niño forecasting based on the global atmospheric oscillation. *International Journal of Climatology*, 41(7):3781–3792, 2021.

[8] M. J. McPhaden, A. Santoso, and W. Cai. Introduction to El Niño Southern Oscillation in a changing climate. *El Niño Southern Oscillation in a changing climate*, pages 1–19, 2020.

[9] Monthly Atmospheric and SST Indices. `https://cpc.ncep.noaa.gov/data/indices/ersst5.nino.mth.91-20.ascii`, 2024.

[10] A. G. Barnston, M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt. Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bulletin of the American Meteorological Society*, 93(5):631–651, 2012.

[11] E. K. Jin, J. L. Kinter, B. Wang, C. K. Park, I. S. Kang, B. P. Kirtman, ..., and T. Yamagata. Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dynamics*, 31:647–664, 2008.

[12] M. Newman and P. D. Sardeshmukh. Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophysical Research Letters*, 44(16):8520–8529, 2017.

[13] J. Ledolter and B. Abraham. Parsimony and Its Importance in Time Series Forecasting. *Technometrics*, 1981.

[14] P. Brockwell and R. Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

[15] A. A. Bawdekar, B. R. Prusty, and K. Bingi. Sensitivity analysis of stationarity tests' outcome to time Series facets and test parameters. *Mathematical Problems in Engineering*, 2022.

[16] M. L'Heureux. An overview of the 2007-08 La Niña and boreal wintertime variability. *Climate Prediction S&T Digest*, 2, 2008.

[17] D. Chen, M. A. Cane, A. Kaplan, S. E. Zebiak, and D. Huang. Predictability of El Niño over the past 148 years. *Nature*, 428(6984):733–736, 2004.

[18] M. H. Glantz. Shades of chaos: lessons learned about lessons learned about forecasting El Niño and its impacts. *International Journal of Disaster Risk Science*, 6:94–103, 2015.

[19] M. Latif, T. P. Barnett, M. A. Cane, M. Flügel, N. E. Graham, H. Von Storch, and S. E. Zebiak. A review of ENSO prediction studies. *Climate Dynamics*, 9:167–179, 1994.

[20] W. Duan and C. Wei. The 'spring predictability barrier' for ENSO predictions and its possible mechanism: results from a fully coupled model. *International Journal of Climatology*, 33(5):1280–1292, 2013.

A Appendix

A.1 Codes in R

```r
library('tseries')
library('TSA')
library('forecast')
library('MASS')

# Read CSV
enso = read.csv('C:/Users/Admin/Downloads/test3.1.csv', header=TRUE)
enso

# The dataset consists of two cols. We store each into a unique variable.
p_nino3.4 = ts(enso$nino3.4_sst, start=c(1950,1), frequency=12)

plot(p_nino3.4)

# Test for Stationarity
adf.test(p_nino3.4)

# Test for Autocorrelation
Box.test(p_nino3.4, type='Ljung')

# Remark. nino3.4 are stationary and autocorrelated.

#####
# Correlogram generation
# Note that the lags generated are in terms of years. However, our data is monthly.
# So we need to find equivalences, to find a lag h (in month) which will be the
    AR/MA potential order.

# For p_nino3.4,
acf_nino = acf(p_nino3.4, plot=FALSE)
acf_nino$lag = acf_nino$lag * 12
plot(acf_nino, main="ACF for Nio 3.4 index", xlab="Lag (in months)")

pacf_nino = pacf(p_nino3.4, plot=FALSE)
pacf_nino$lag = pacf_nino$lag * 12
plot(pacf_nino, main="PACF for Nio 3.4 index", xlab="Lag (in months)")

# Notes
# ACF --> q = 1, 2, 3
# PACF --> p = 1, 3

stl_nino = decompose(p_nino3.4)
plot(stl_nino)

# Seasonality can be observed. So it might be worthwhile to do differencing.
# For d = 1
nino_d = diff(p_nino3.4, differences=1)
adf.test(nino_d)
Box.test(nino_d, type='Ljung')

acf_nino_d = acf(nino_d, plot=FALSE)
acf_nino_d$lag = acf_nino_d$lag * 12
plot(acf_nino_d, main="ACF for Nio 3.4 index, d=1", xlab="Lag (in months)")

pacf_nino_d = pacf(nino_d, plot=FALSE)
pacf_nino_d$lag = pacf_nino_d$lag * 12
plot(pacf_nino_d, main="ACF for Nio 3.4 index, d=1", xlab="Lag (in months)")

## Seasonality is still observed so we proceed with d=2
# For d = 2
```

```r
nino_d = diff(p_nino3.4, differences=2)
adf.test(nino_d)
Box.test(nino_d, type='Ljung')

acf_nino_d = acf(nino_d, plot=FALSE)
acf_nino_d$lag = acf_nino_d$lag * 12
plot(acf_nino_d, main="ACF for Nio 3.4 index, d=2", xlab="Lag (in months)")

pacf_nino_d = pacf(nino_d, plot=FALSE)
pacf_nino_d$lag = pacf_nino_d$lag * 12
plot(pacf_nino_d, main="ACF for Nio 3.4 index, d=2", xlab="Lag (in months)")

# ACF --> q = 1, 3
# PACF --> p = 1, 2, 3
# Why cut off until 3? Parsimony

# For nino_diff
nino_ma1 = Arima(nino_d, order=c(0,0,1))
nino_ma3 = Arima(nino_d, order=c(0,0,3))
nino_ar1 = Arima(nino_d, order=c(1,0,0))
nino_ar2 = Arima(nino_d, order=c(2,0,0))
nino_ar3 = Arima(nino_d, order=c(3,0,0))
nino_arma11 = Arima(nino_d, order=c(1,0,1))
nino_arma21 = Arima(nino_d, order=c(2,0,1))
nino_arma31 = Arima(nino_d, order=c(3,0,1))
nino_arma13 = Arima(nino_d, order=c(1,0,3))
nino_arma23 = Arima(nino_d, order=c(2,0,3))
nino_arma33 = Arima(nino_d, order=c(3,0,3))

cbind(nino_ma1$aic, nino_ma3$aic, nino_ar1$aic, nino_ar2$aic, nino_ar3$aic,
      nino_arma11$aic, nino_arma21$aic, nino_arma31$aic,
      nino_arma13$aic, nino_arma23$aic, nino_arma33$aic)

nino_arma33


nino_arma33.a = arima(nino_d, order=c(3,0,3))
nino_arma33.a
# Best model is ARMA(3,3)
```