# Homework 4

Philip Situmorang

## STAT 535 - Homework 4

The file ChinaImports2.txt gives monthly imports by the U.S. from China, in millions of dollars, for the period January 1989 through December 2019.

1. Graph both the monthly imports and the logged monthly imports vs. time. Comment in detail on what you can determine from the plots. Mark economic downturns in the plots.

**Discussion:**

**TREND** - Upward trend is visible. In the imports plot (blue line), imports seem to increase exponentially or by a constant multiplicative factor until the 2009 recession period (highlighted in orange) where it begins to flatten. In the log Imports plot (red line) log Imports seem to increase by a constant additive amount before flattening after the 2009 recession.

**SEASONALITY** - The plots indicate strong seasonality in imports. Imports tend to reach its highest peak around October annually. After dropping in November and December, Imports tend to increase in January, making a small peak before dropping again in February. From February to October, imports tend to increase at a steady rate.

**VOLATILITY** - There is evidence of increasing volatility in the imports plot. For example, it is evident that imports tend to rise and drop more sharply after the 2009 recession period. This suggest a multiplicative model is preferred than an additive one.

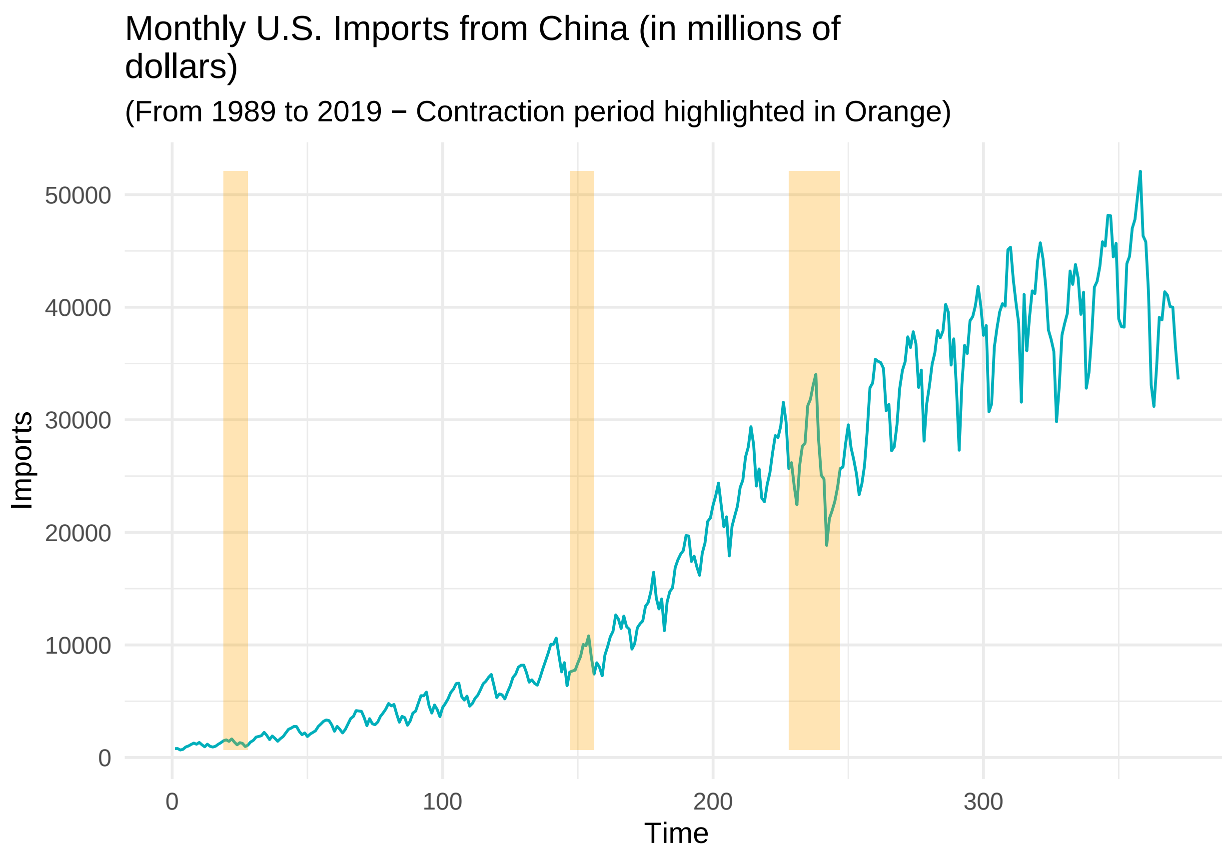
**IMPACT OF ECONOMIC DOWNTURN** - Economic downturn tend to cause imports to drop, as is evident in the 2001 and 2009 contractions (the drop is less prominent in 1990 where imports from China is still very low compared to the following periods).

**OUTLIER** - The drop in imports in February of 2009 is likely an outlier.

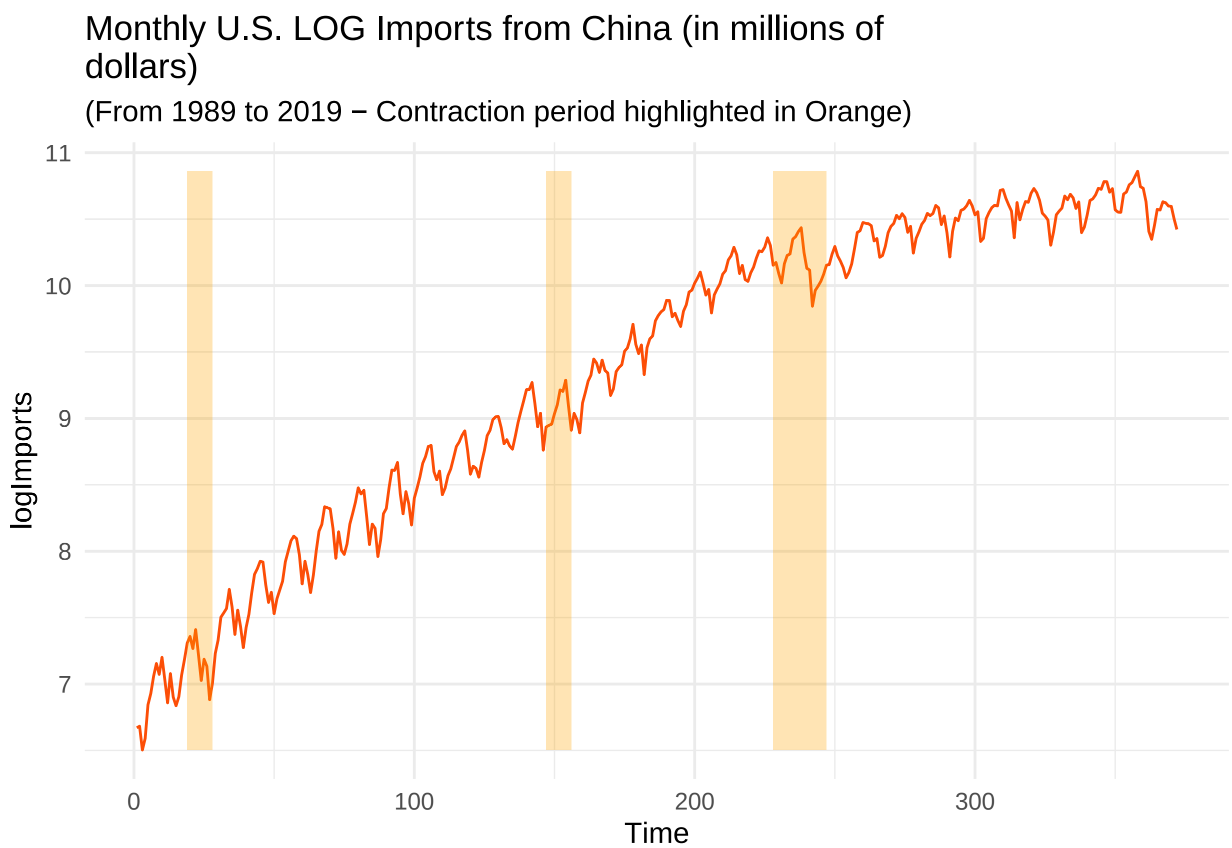
```
f <- file.choose("ChinaImports2.txt")
imports <- read.csv(f)
```

```
imports <- imports %>% mutate (year_month = paste(Year,Month,sep="-"))
```

```
# modified
cycle <- data.frame(from = c('1945-2','1948-11','1953-7','1957-8','1960-4','1969-12','1973-11','1980-1'
                    to = c('1945-10','1949-10','1954-5','1958-4','1961-2','1970-11','1975-3','1980-7','
```




```
ggplot(data=dat, aes(Time, logImports)) +
theme_minimal() +
geom_line(color = "#FC4E07", size = .5) +
geom_rect(data=rects, inherit.aes=FALSE, aes(xmin=start, xmax=end, ymin=min(dat$logImports),
ymax=max(dat$logImports), group=group), color="transparent", fill="orange",
alpha=0.3) +
labs(title = "Monthly U.S. LOG Imports from China (in millions of
dollars)", subtitle = "(From 1989 to 2019 - Contraction period highlighted in Orange)")
```




2. Fit a regression model with seasonal and significant calendar components to the differences of the logged monthly imports.

```
# add fMonth variables
imports <- imports %>% mutate(fMonth = as.factor(Month))
```

```
# add differences column
logImports_lst <- imports$logImports
logDiff <- diff(logImports_lst,lag=1)

# insert NA to the beginning of the differences
logDiff <- c(NA, logDiff)

# combine to dataframe
imports$logDiff <- logDiff
```

```
model1<-lm(logDiff~fMonth+c348+s348+c432+s432, data = imports);
```

```
model2 <- lm(logDiff~fMonth+c348+s348, data = imports)
anova(model2, model1)
```

```
## Analysis of Variance Table
##
## Model 1: logDiff ~ fMonth + c348 + s348
## Model 2: logDiff ~ fMonth + c348 + s348 + c432 + s432
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1    357 1.5883
## 2    355 1.5667  2  0.021616 2.4491 0.08783 .
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(a) Describe your fitted model.

**Discussion:** The fitted model is a regression model with the differenced log imports of order 1 (logDiff) as y variable and fMonth dummy variables and calendar pairs 348 and 432 as X variables. The differenced log monthly imports column was created by using the diff() function on the logImports and adding 'NA' to the very first period in the data frame (Jan 1989) which does not have a corresponding differenced value.

The partial F-test shows that calendar pairs 432 has marginal significance of 0.087 but we include it in the regression model nevertheless.

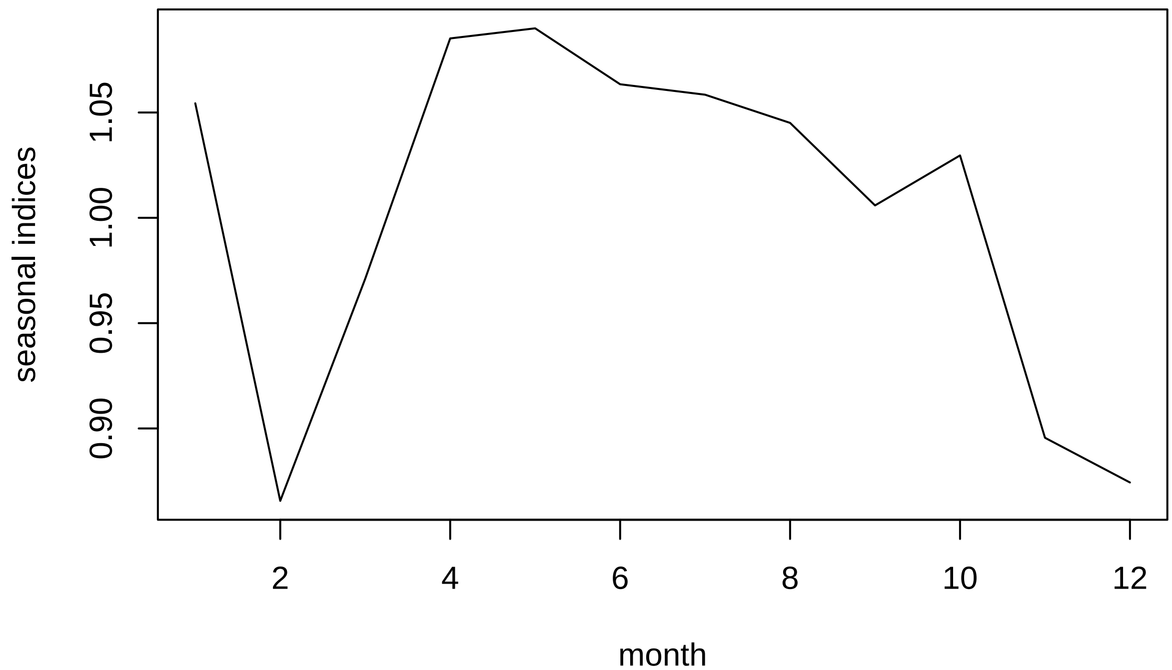
(b) Tabulate and plot the estimated seasonal indices for the imports series. Provide careful interpretation of the estimates.

**Discussion:** Looking at the plot of the indices below, we see that on average imports drop in the month of February before steadily increasing until the month of October before dropping again in the month of November and December. Imports then jumps in January before dropping again in February.

More specifically, looking at final indices, since February has seasonal index of .87 we can expect sales to fall roughly 13 percent below the level of the trend. In March, imports on average increase to 3 percent below the level of trend (March has a seasonal index of .97). From April to October, imports increase and on average go above the level of trend as indicated by indices which exceed 1.00. In November and December they fall to roughly 10 and 13 percent below level of trend. In January imports rise to roughly 5 percent above trend level.

```r
b1<-coef(model1)[1]
b2<-coef(model1)[2:12]+b1
b3<-c(b1,b2)
seas<-b3-mean(b3)

seas.ts<-ts(exp(seas))
plot(seas.ts,ylab="seasonal indices",xlab="month")
```




```
month <- seq(12)
seas_indices <- exp(seas)
seas_df <- data.frame(month, seas, seas_indices)
print.data.frame(tbl_df(seas_df))
```

```
## Warning: `tbl_df()` was deprecated in dplyr 1.0.0.
## Please use `tibble::as_tibble()` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.
```

```
##    month         seas seas_indices
## 1      1  0.052948420    1.0543753
## 2      2 -0.144305347    0.8656234
## 3      3 -0.029387793    0.9710398
## 4      4  0.081739214    1.0851728
## 5      5  0.086141128    1.0899601
## 6      6  0.061475925    1.0634049
## 7      7  0.056809263    1.0584539
## 8      8  0.044079249    1.0450652
## 9      9  0.005870053    1.0058873
## 10    10  0.029194210    1.0296245
## 11    11 -0.110288413    0.8955758
## 12    12 -0.134275909    0.8743488
```

(c) Perform a residual analysis for your model, examining the plot of the residuals vs. time, a residual normal quantile plot, the residual acf and pacf plots, and the residual spectral density (along with Bartlett's test). Discuss all.

**Discussion:**

**Residuals vs. Time Plot -** In the residuals vs. time plot, the flatness of the residuals indicate that trend has largely been captured. However, we do see that the model has not been reduced to white noise as we see bursts in certain periods.

**QQPLOT -** Most of the observations fall on or close to the normal line, though we do see that many are relatively further away from the line at the two tails. Not enough to conclude that the residuals are not normally distributed.

**ACF plot -** The ACF plot shows nonsmooth pattern, which indicate that trend has largely been captured. However we do see significant autocorrelation structure at lags 1, 2, 3, and 11, as well as at lags 12, 24, and 36, indicating that there remains uncaptured autocorrelation structure as well as seasonality structure. The autocorrelation at lags 12, 24, and 36 indicate that there exists dynamic seasonality within the time series, and the current model does not adequately capture it (we use one fMonth variables which only allows capturing one seasonality structure). The lags

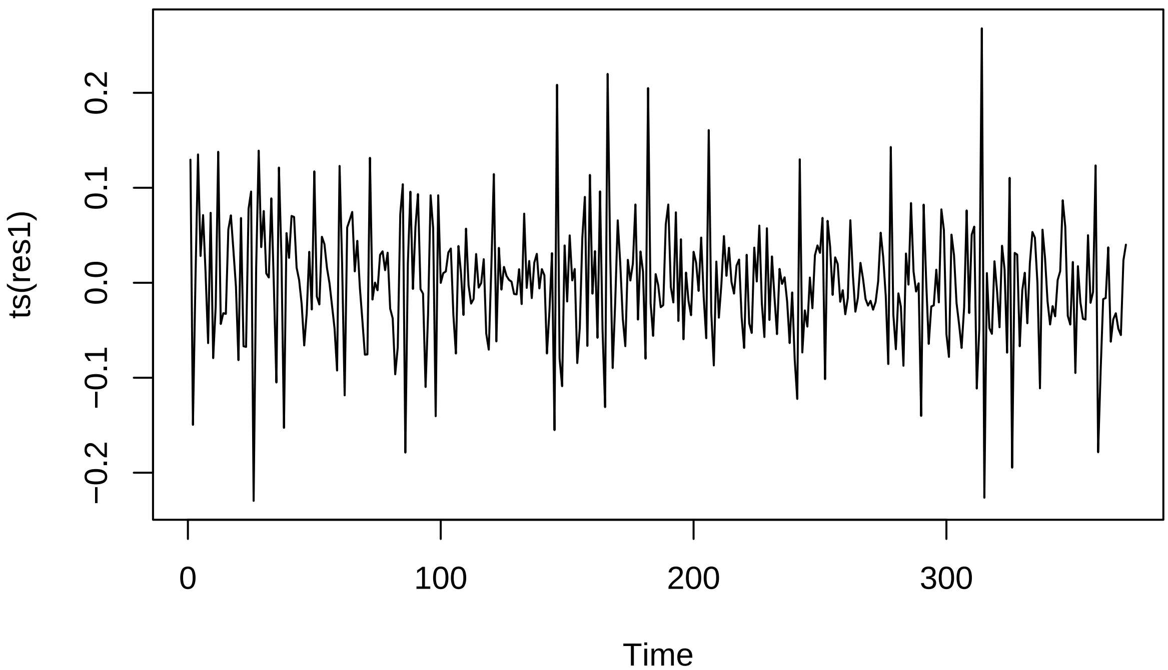
**PACF plot -** There are significant spikes in lags 1, 2, 10, and 25, which indicates uncaptured moving average structure in the data. We also see significant spikes at lags 12 and 36, which indicate that higher order moving average structure has not been captured.

**Spectral Density plot -** The spectral density plot shows uncaptured seasonality structure as indicated by the peaks at the seasonal frequencies (red lines). The plot also indicates that the model has not been reduced to white noise as the distance between the highest and lowest point is twice the upper half of the blue measurement line. The plot indicates, however, that trend has been captured as indicated by the low peak at the beginning of the plot.

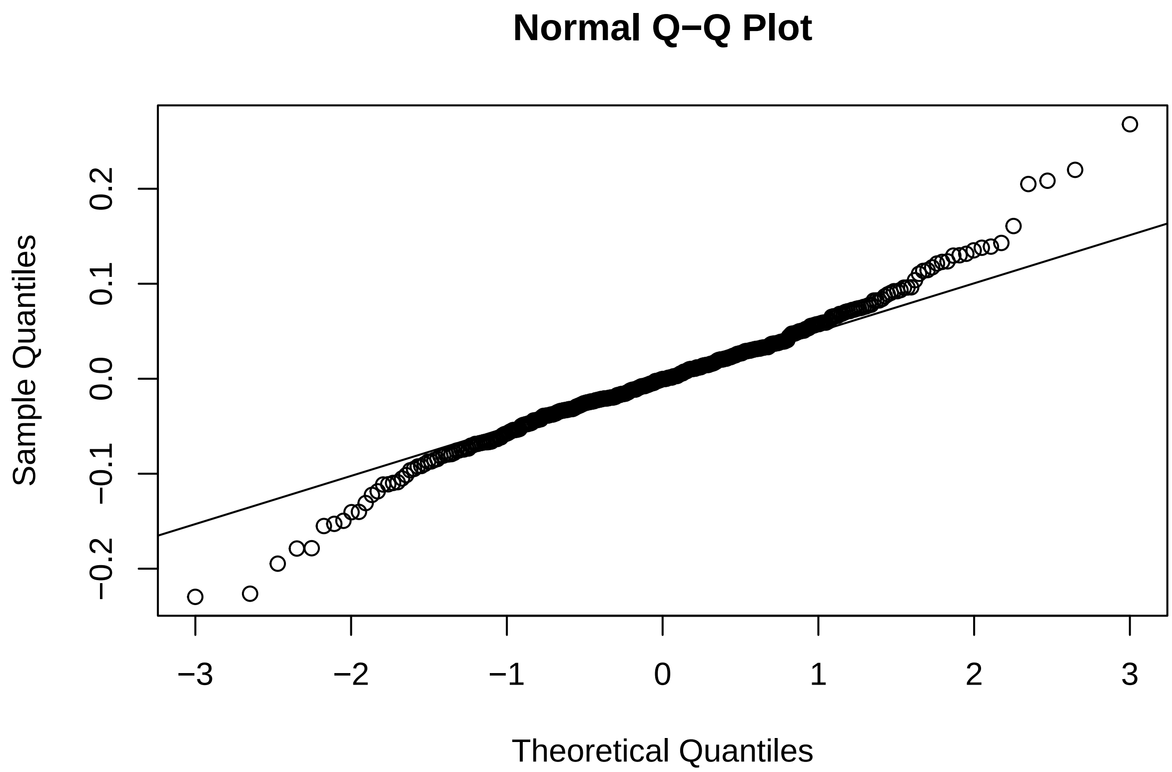
**Bartlett's test -** Bartlett's test yields a p-value close to 0, which indicates that we should reject the null hypothesis that the model has been reduced to white noise. This is consistent with our analysis of

```
res1 <- resid(model1)
```

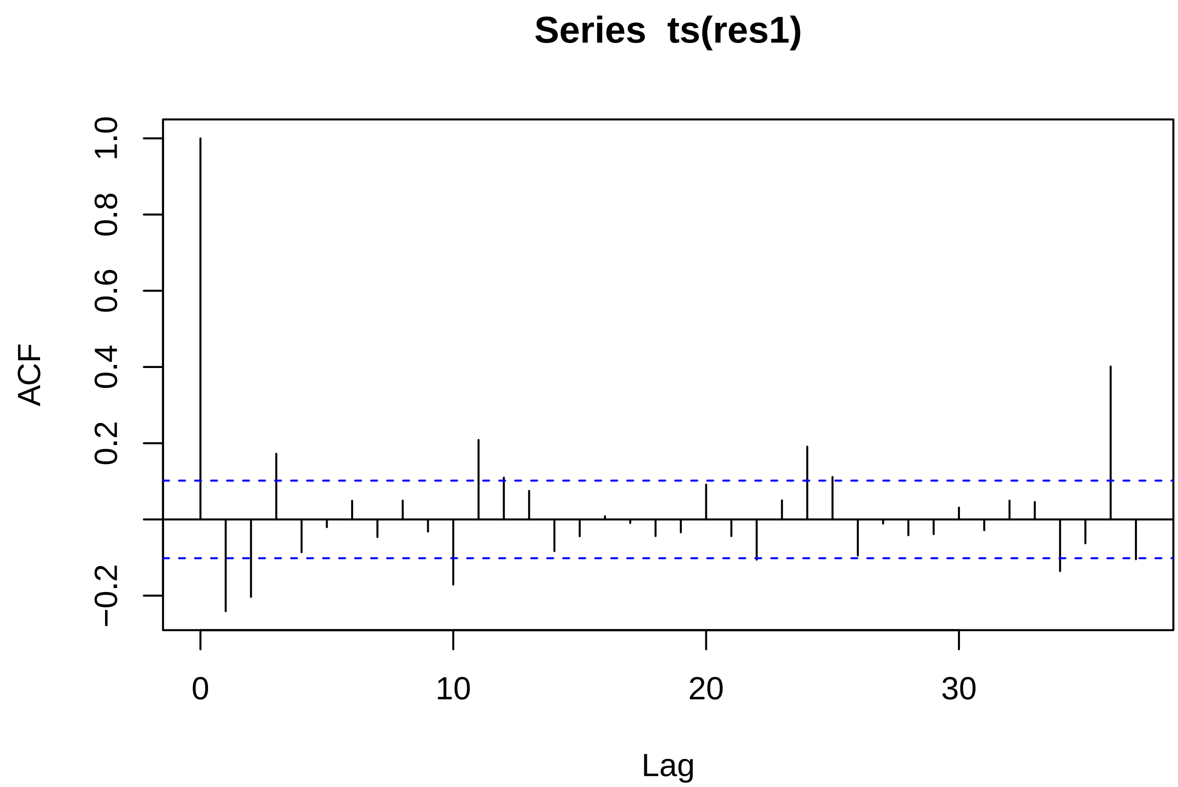
```
plot(ts(res1))
```



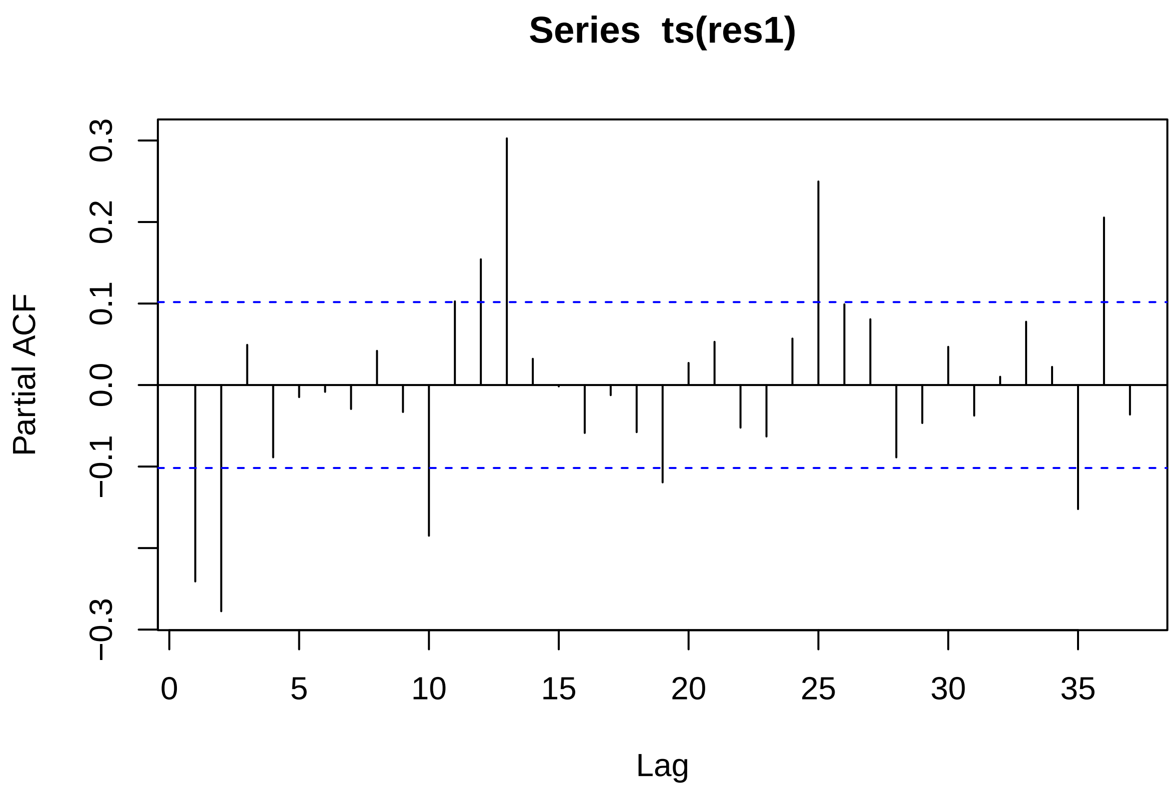

```
qqnorm(res1)
qqline(res1)
```



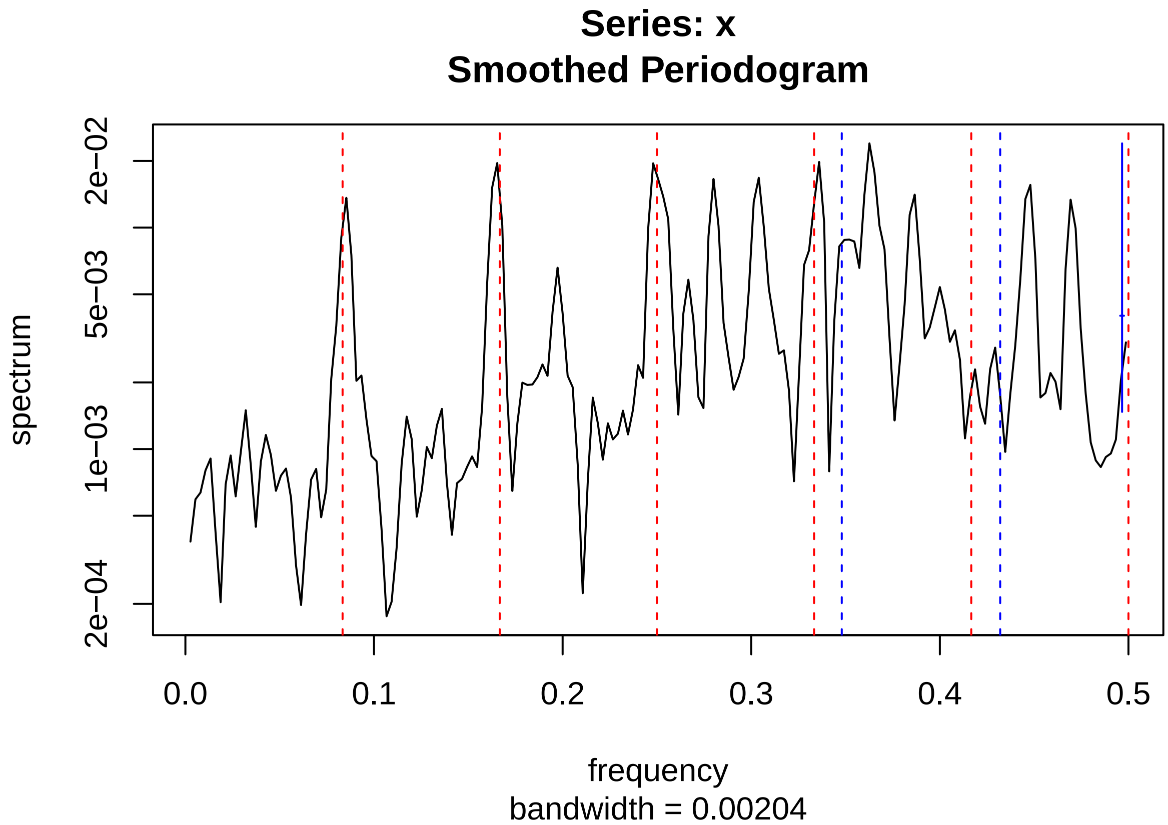

```
acf(ts(res1), 37)
```




```
pacf(ts(res1), 37)
```




```
spectrum(res1, span=3)
abline(v=c(1/12,2/12,3/12,4/12,5/12,6/12),col="red",lty=2)
abline(v=c(0.348,0.432),col="blue",lty=2)
```



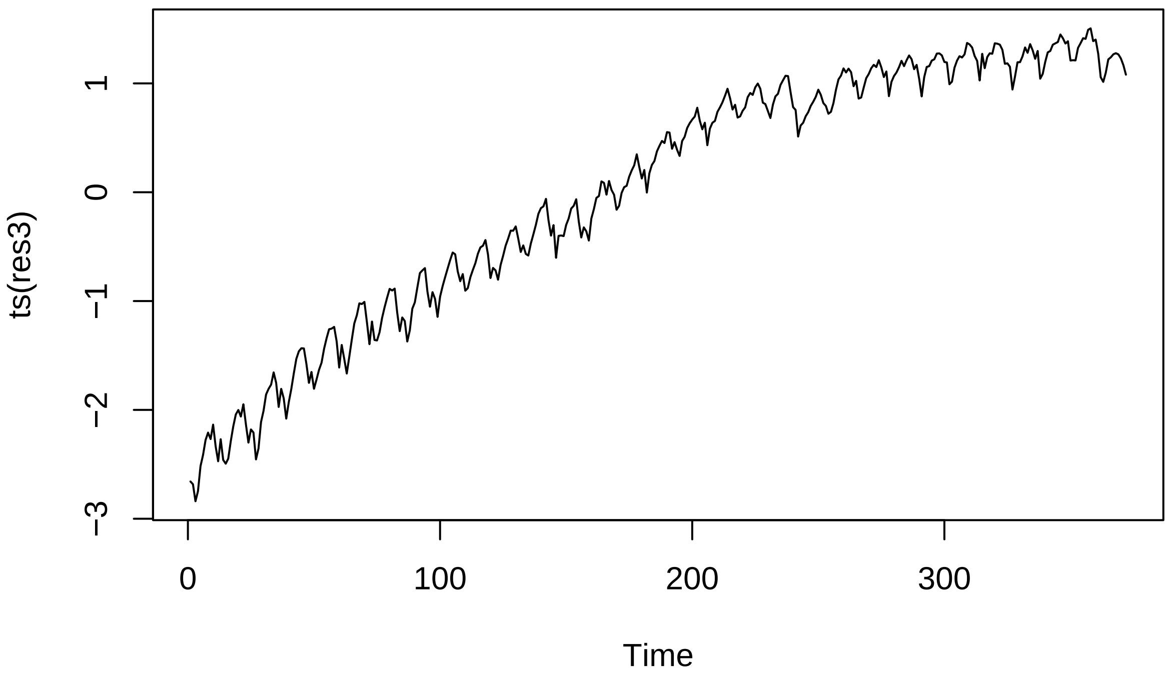

```
bartlettB.test(ts(res1))
```

```
##
##  Bartlett B Test for white noise
##
## data:
## = 3.1802, p-value = 3.285e-09
```

3. (a) Fit a regression model with just the calendar variables to the logged monthly imports series. [The calendar pairs will not be significant in this model. That's okay— do so and proceed as below.]

```
model3 <- lm(logImports~c348+s348+c432+s432, data = imports)
```

```
res3 <- resid(model3)
plot(ts(res3))
```



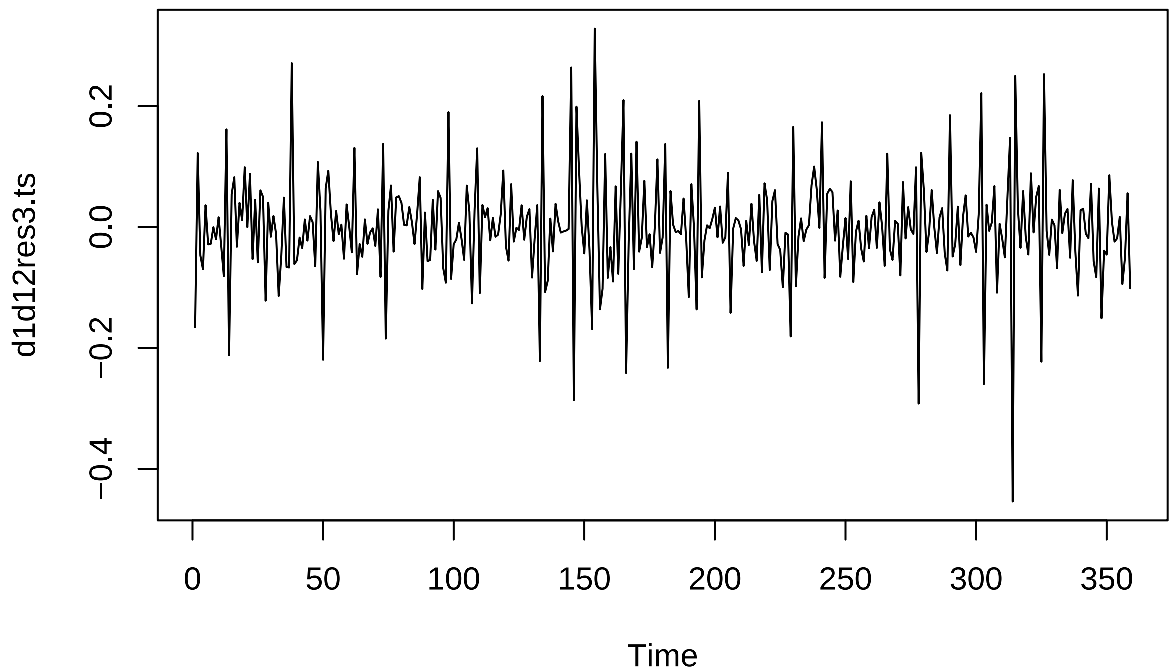

(b) Fit a seasonal ARIMA model to the residuals from the model in part (a).

**Discussion:** The first set of R-chunks below are for the initial analysis of where to start in terms of ARIMA structure. The following set of codes are fitting of different ARIMA models to see which one is best. The discussion on how we arrive at the final model follows in the next section. In this section we fit three ARIMA models and attempt to see which one is best based on AIC and residual analysis. The three ARIMA models are:

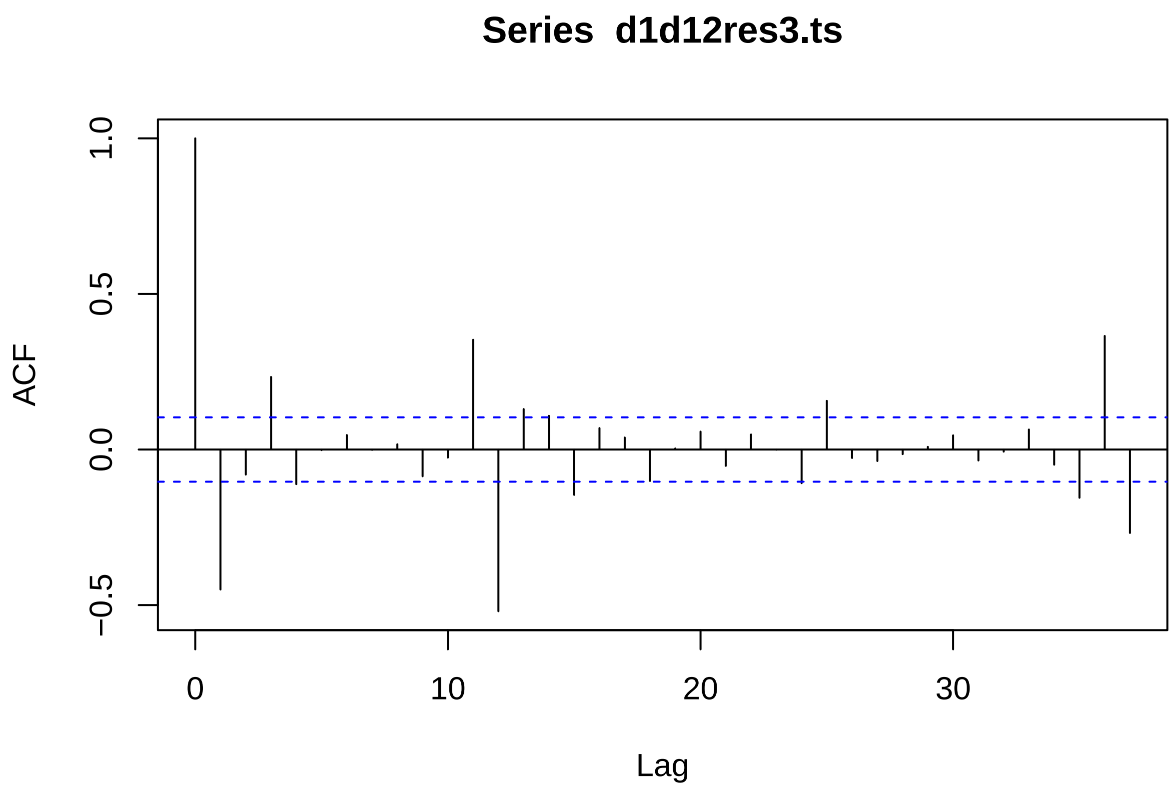
ARIMA(11,1,11)(1,1,1)12 ARIMA(0,1,11)(1,1,1)12 ARIMA(11,1,0)(1,1,1)12

```
d1d12res3.ts<-ts(diff(diff(res3),12))
```

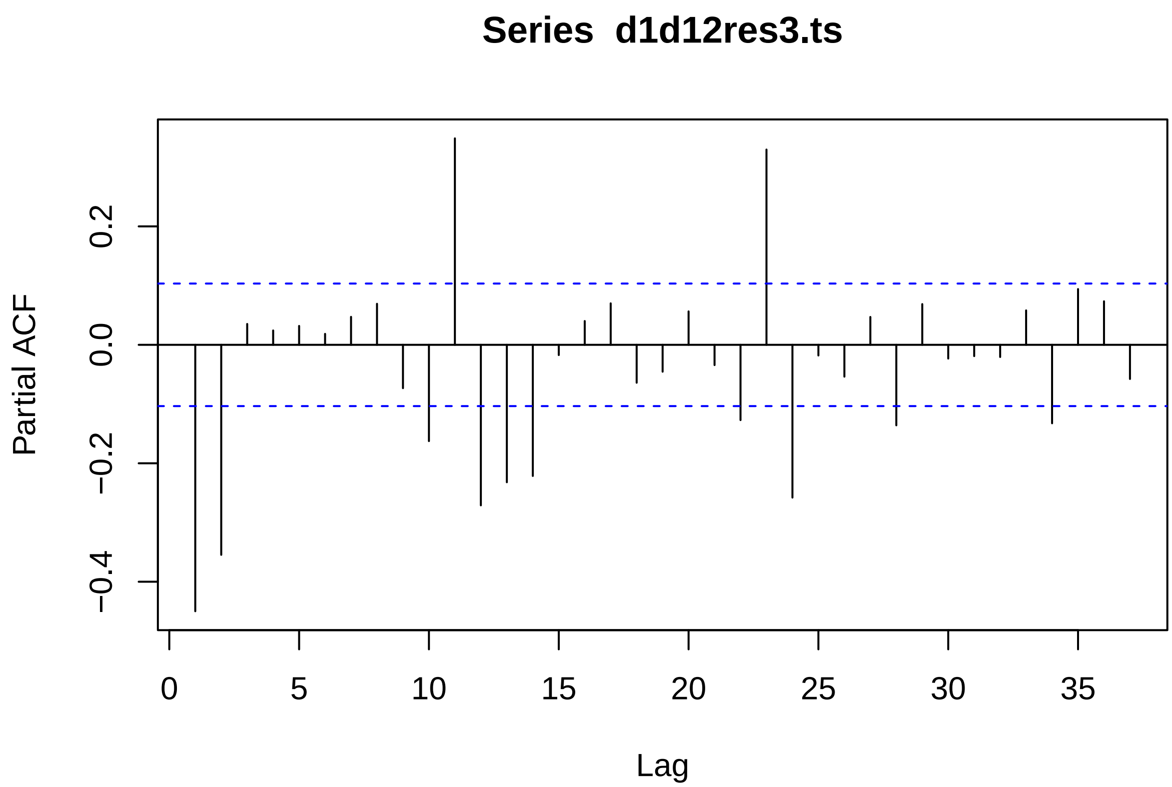
```
plot(d1d12res3.ts)
```




```
acf(d1d12res3.ts, 37)
```




```
pacf(d1d12res3.ts, 37)
```



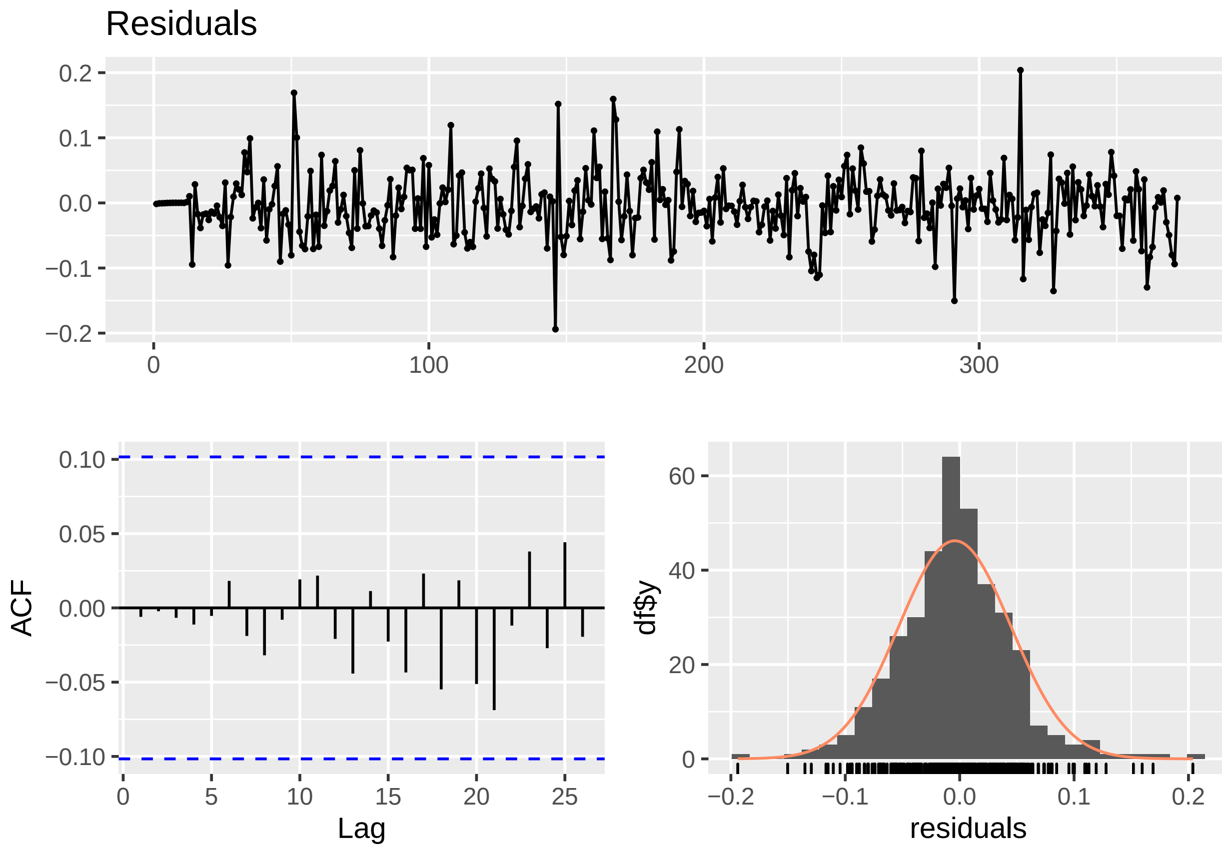

**Fit and analysis of ARIMA(11,1,11)(1,1,1)12**

```
res3.ts <- ts(res3)
model4<-arima(res3.ts,order=c(11,1,11),seasonal=list(order=c(1,1,1),period=12))
model4
```

```
##
## Call:
## arima(x = res3.ts, order = c(11, 1, 11), seasonal = list(order = c(1, 1, 1),
##     period = 12))
##
## Coefficients:
##           ar1      ar2      ar3      ar4      ar5      ar6      ar7      ar8
##       -0.9483  -0.6312  -0.3173  -0.3800  -0.2958  -0.4220  -0.6317  -0.9505
## s.e.   0.3841   0.5377   0.4502   0.2498   0.1980   0.1779   0.2437   0.3563
##           ar9     ar10    ar11     ma1     ma2     ma3     ma4     ma5     ma6
##       -0.5054  -0.0392  0.0314  0.5082  0.1596  0.1849  0.2661  0.1732  0.2757
## s.e.   0.5330   0.4139  0.1645  0.3797  0.3695  0.2114  0.1040  0.0921  0.0914
##          ma7     ma8     ma9     ma10    ma11     sar1     sma1
##       0.5786  0.7704  0.0919  -0.2666  0.3214  -0.1423  -0.5993
## s.e.  0.1556  0.2890  0.4204   0.2021  0.1364   0.0867   0.0706
##
## sigma^2 estimated as 0.002513:  log likelihood = 552.96,  aic = -1055.92
```

```
checkresiduals(ts(resid(model4)))
```

```
## Warning in modeldf.default(object): Could not find appropriate degrees of
## freedom for this model.
```

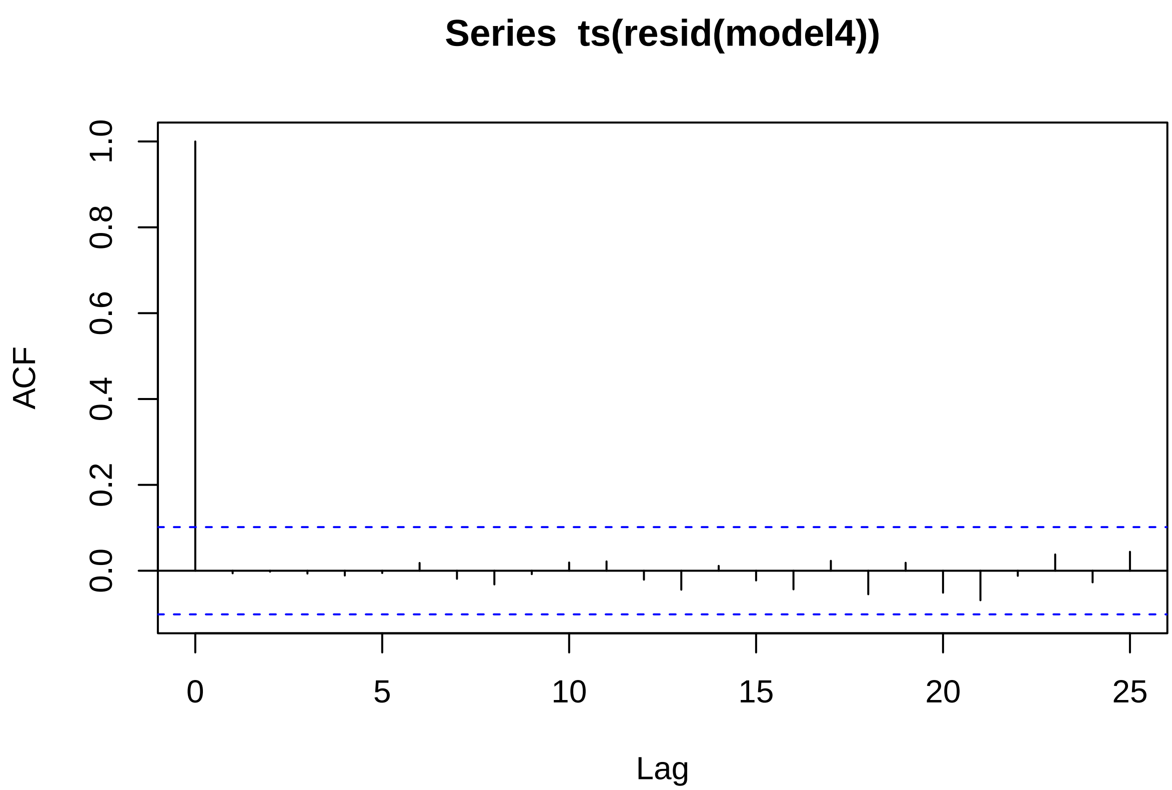



```
coeftest(model4)
```

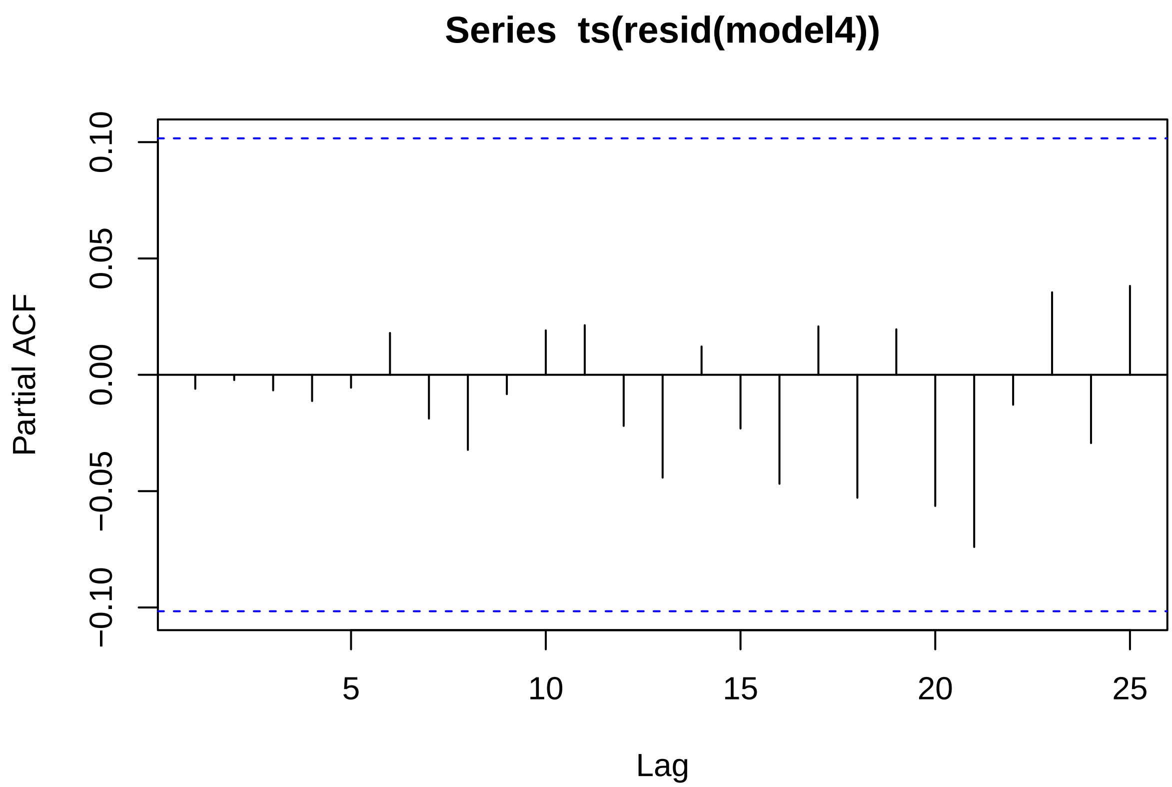
```
##
## z test of coefficients:
##
##        Estimate Std. Error z value  Pr(>|z|)
## ar1  -0.948261   0.384123 -2.4686 0.0135629 *
## ar2  -0.631218   0.537712 -1.1739 0.2404369
## ar3  -0.317277   0.450175 -0.7048 0.4809435
## ar4  -0.379950   0.249798 -1.5210 0.1282519
## ar5  -0.295846   0.198011 -1.4941 0.1351532
## ar6  -0.421956   0.177896 -2.3719 0.0176955 *
## ar7  -0.631699   0.243741 -2.5917 0.0095507 **
## ar8  -0.950521   0.356252 -2.6681 0.0076278 **
## ar9  -0.505360   0.533025 -0.9481 0.3430796
## ar10 -0.039220   0.413871 -0.0948 0.9245015
## ar11  0.031365   0.164511  0.1907 0.8487950
## ma1   0.508240   0.379729  1.3384 0.1807568
## ma2   0.159627   0.369521  0.4320 0.6657529
## ma3   0.184944   0.211388  0.8749 0.3816260
## ma4   0.266112   0.103996  2.5589 0.0105013 *
## ma5   0.173227   0.092080  1.8813 0.0599357 .
## ma6   0.275718   0.091376  3.0174 0.0025495 **
## ma7   0.578628   0.155611  3.7184 0.0002005 ***
## ma8   0.770436   0.288991  2.6659 0.0076771 **
## ma9   0.091886   0.420392  0.2186 0.8269824
```

```
## ma10 -0.266574   0.202100 -1.3190 0.1871612
## ma11  0.321359   0.136431  2.3555 0.0184992 *
## sar1 -0.142297   0.086708 -1.6411 0.1007739
## sma1 -0.599297   0.070610 -8.4875 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

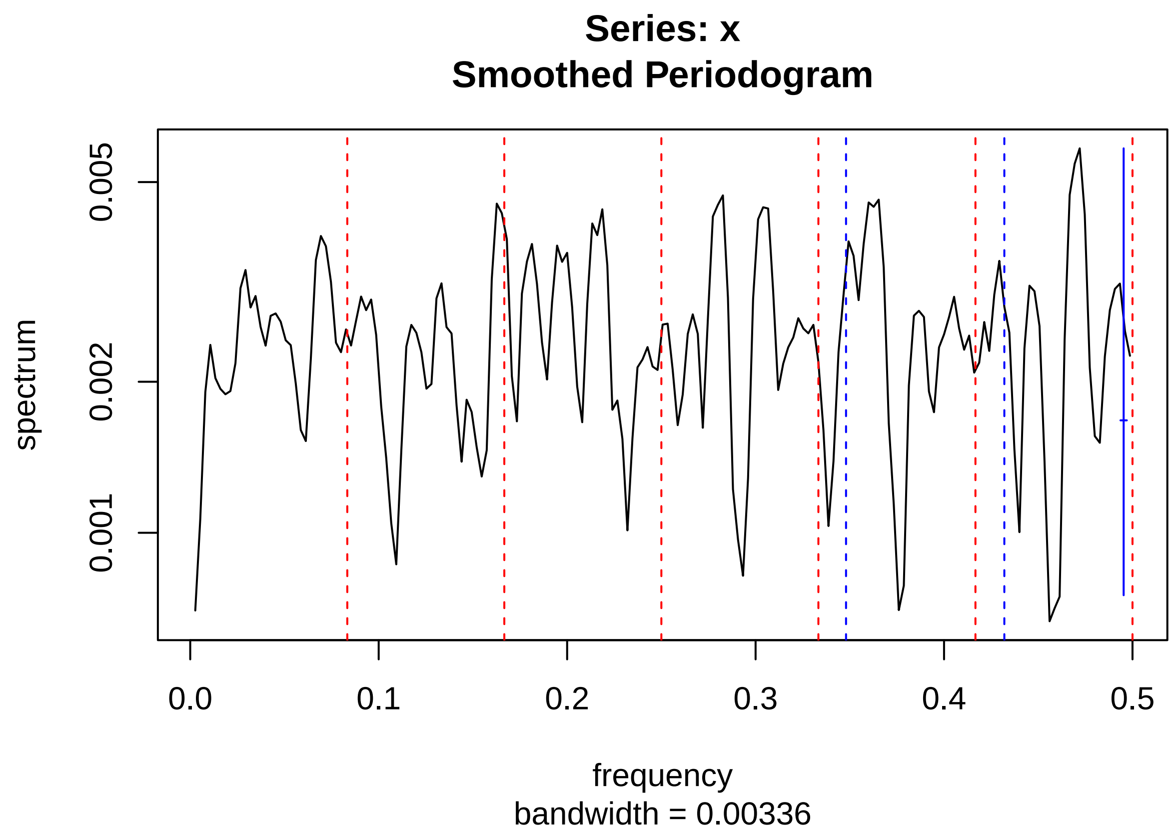
```
acf(ts(resid(model4)))
```




```
pacf(ts(resid(model4)))
```




```r
spectrum(resid(model4), span=5)
abline(v=c(1/12,2/12,3/12,4/12,5/12,6/12),col="red",lty=2)
abline(v=c(0.348,0.432),col="blue",lty=2)
```




```
bartlettB.test(ts(resid(model4)))
```

```
##
##  Bartlett B Test for white noise
##
## data:
## = 0.30128, p-value = 1
```

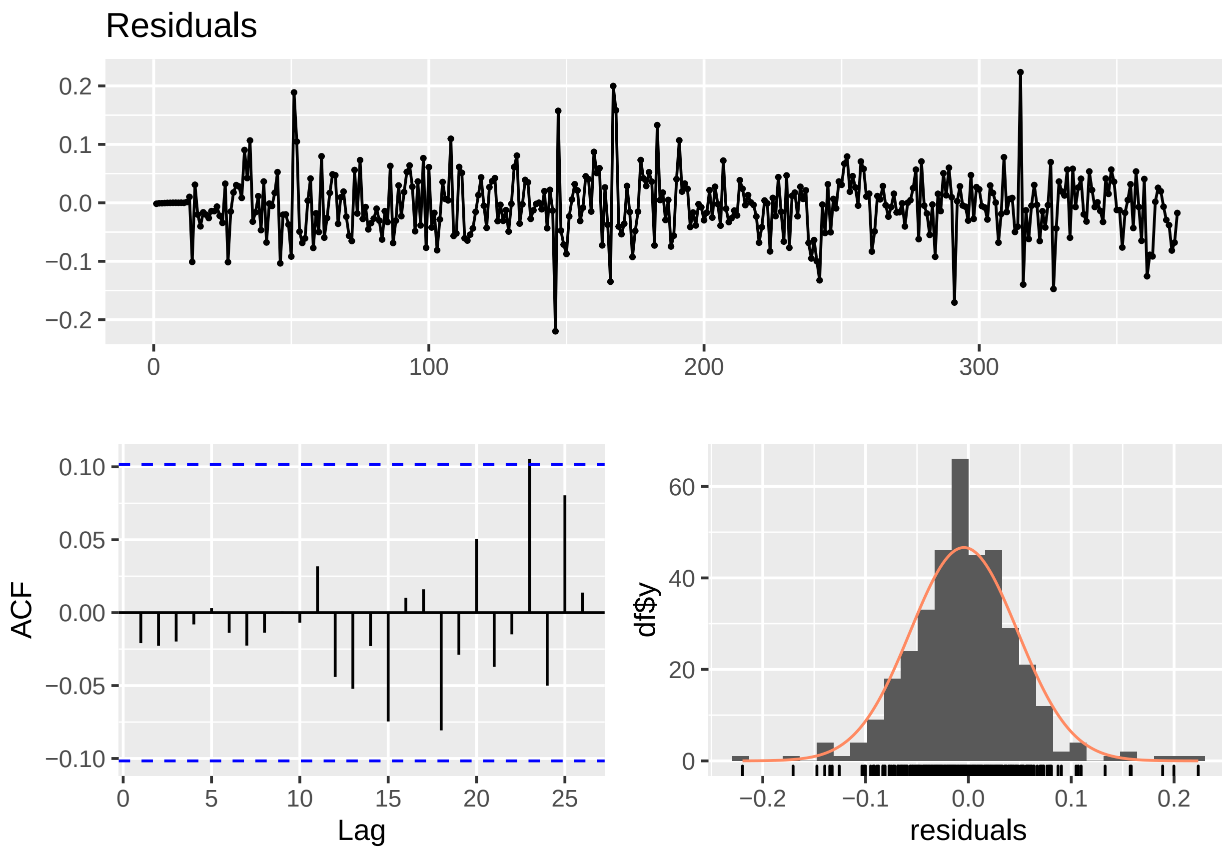
**Fit and analysis of ARIMA(0,1,11)(1,1,1)12**

```
model5<-arima(res3.ts,order=c(0,1,11),seasonal=list(order=c(1,1,1),period=12))
model5
```

```
##
## Call:
## arima(x = res3.ts, order = c(0, 1, 11), seasonal = list(order = c(1, 1, 1),
##     period = 12))
##
## Coefficients:
##           ma1      ma2     ma3      ma4     ma5     ma6     ma7      ma8
##       -0.4461  -0.0608  0.2397  -0.1616  0.0665  0.0150  0.0725  -0.0736
## s.e.   0.0536   0.0572  0.0570   0.0594  0.0603  0.0659  0.0609   0.0538
##           ma9     ma10    ma11     sar1     sma1
##       -0.0013  -0.0085  0.2324  -0.2815  -0.6704
## s.e.   0.0651   0.0585  0.0564   0.0672   0.0456
##
## sigma^2 estimated as 0.002839:  log likelihood = 537.38,  aic = -1046.76
```

```
checkresiduals(ts(resid(model5)))
```

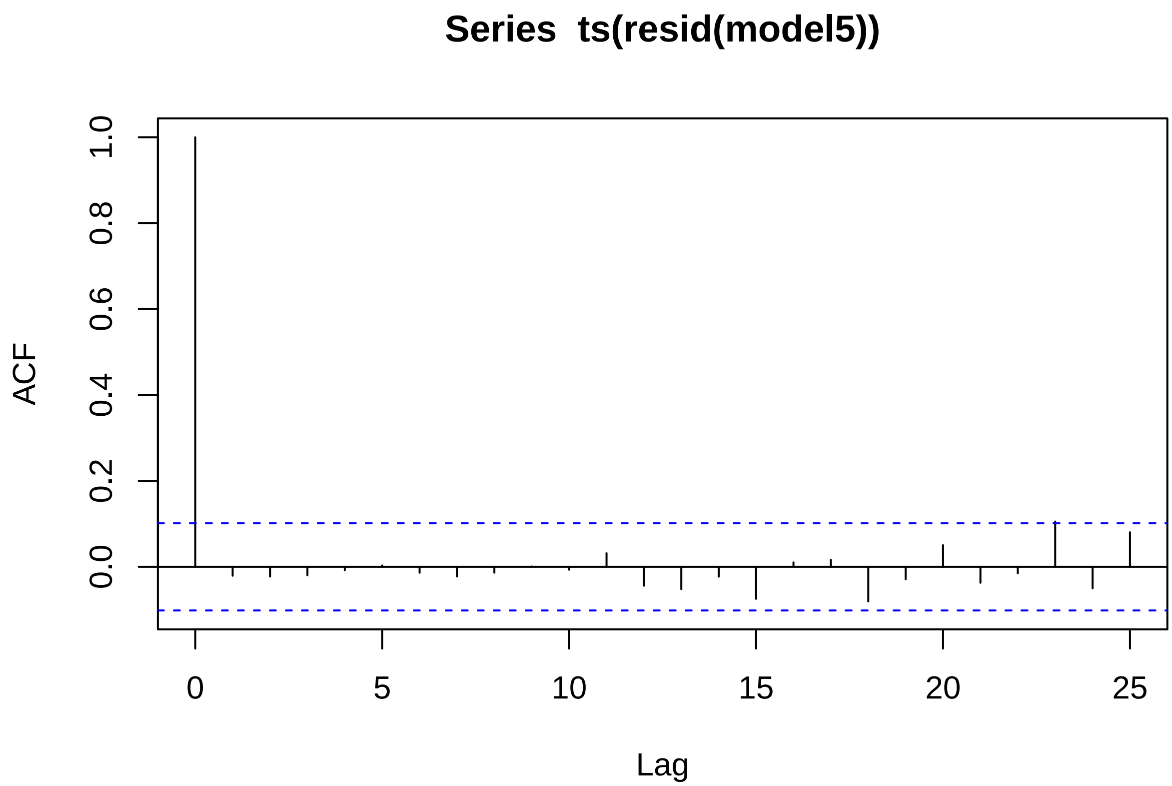
```
## Warning in modeldf.default(object): Could not find appropriate degrees of
## freedom for this model.
```



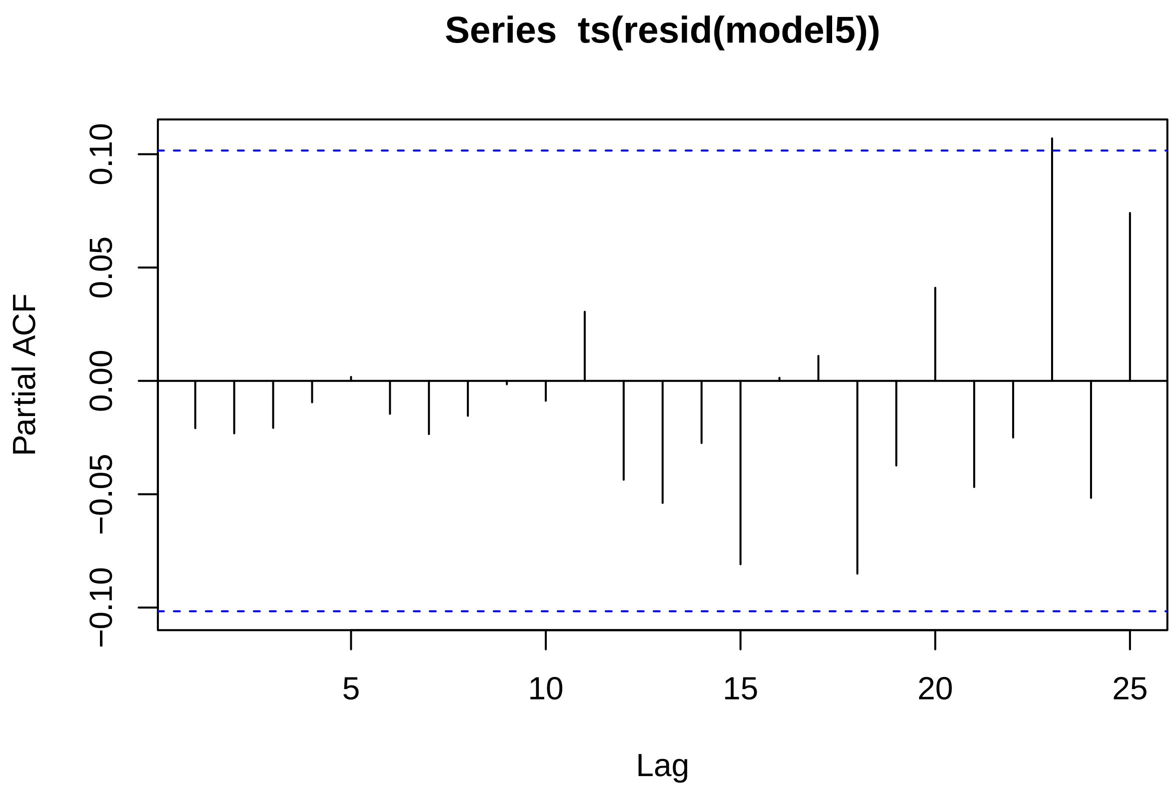

```
coeftest(model5)
```

```
##
## z test of coefficients:
##
##        Estimate Std. Error  z value  Pr(>|z|)
## ma1  -0.4461310  0.0535763  -8.3270 < 2.2e-16 ***
## ma2  -0.0607736  0.0572027  -1.0624  0.288043
## ma3   0.2396881  0.0570122   4.2042 2.621e-05 ***
## ma4  -0.1616336  0.0593534  -2.7232  0.006464 **
## ma5   0.0665166  0.0602881   1.1033  0.269892
## ma6   0.0150491  0.0659029   0.2284  0.819372
## ma7   0.0724682  0.0609098   1.1898  0.234139
## ma8  -0.0736183  0.0538176  -1.3679  0.171337
## ma9  -0.0012769  0.0650554  -0.0196  0.984340
## ma10 -0.0084839  0.0584674  -0.1451  0.884628
## ma11  0.2323652  0.0564483   4.1164 3.848e-05 ***
## sar1 -0.2815049  0.0671707  -4.1909 2.779e-05 ***
## sma1 -0.6704438  0.0455845 -14.7077 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
acf(ts(resid(model5)))
```




```
pacf(ts(resid(model5)))
```




```
bartlettB.test(ts(resid(model5)))
```

```
##
##  Bartlett B Test for white noise
##
## data:
## = 0.48505, p-value = 0.9727
```

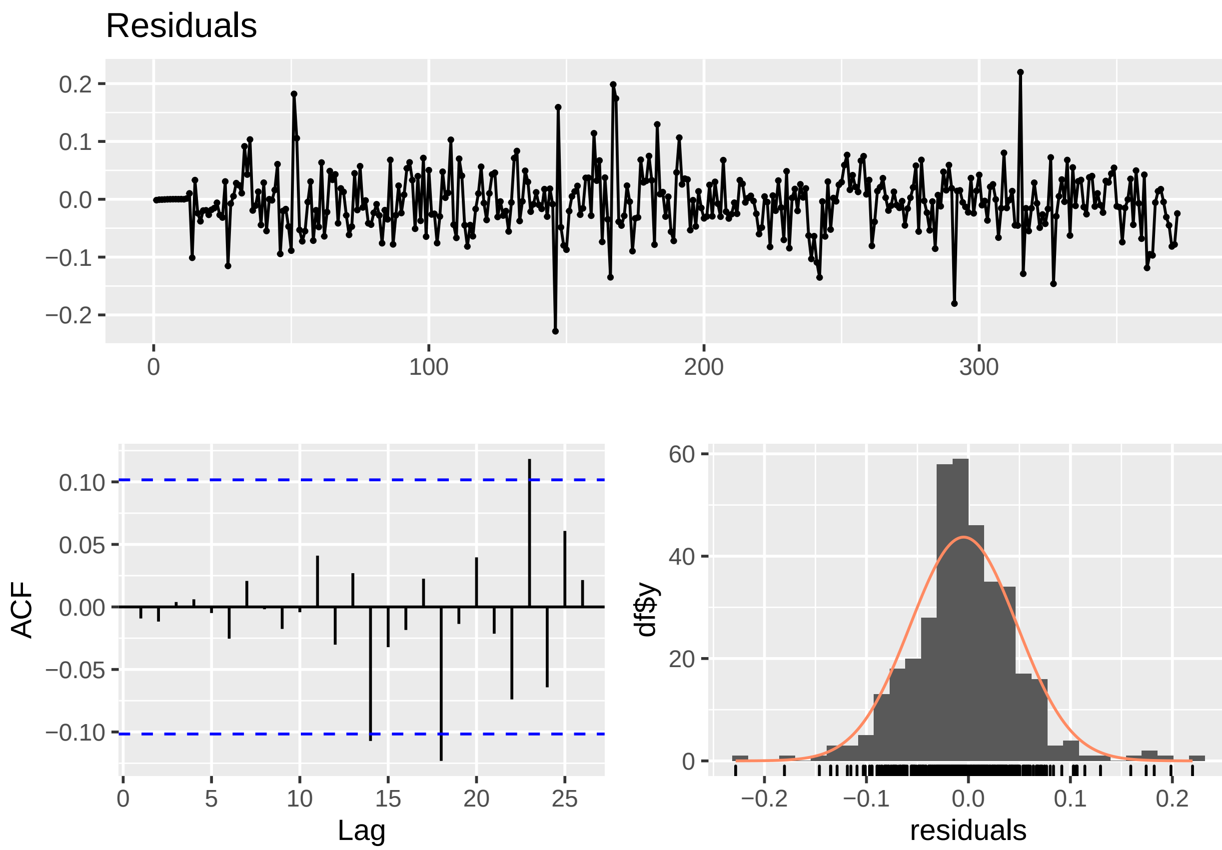
**Fit and analysis of ARIMA(11,1,0)(1,1,1)12**

```
model6<-arima(res3.ts,order=c(11,1,0),seasonal=list(order=c(1,1,1),period=12))
model6
```

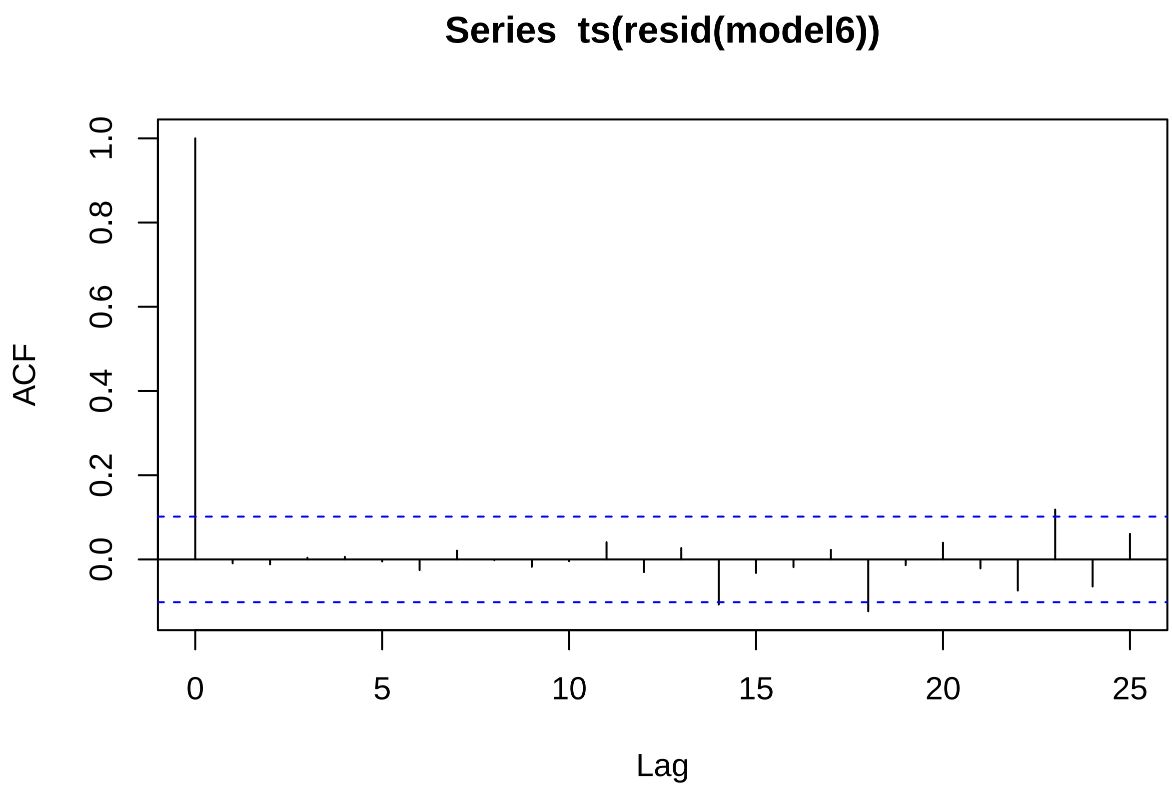
```
##
## Call:
## arima(x = res3.ts, order = c(11, 1, 0), seasonal = list(order = c(1, 1, 1),
##     period = 12))
##
## Coefficients:
##           ar1      ar2     ar3      ar4     ar5     ar6     ar7      ar8
##       -0.4750  -0.2770  0.0467  -0.0405  0.0383  0.0129  0.0687  -0.0270
## s.e.   0.0533   0.0577  0.0601   0.0608  0.0603  0.0610  0.0606   0.0602
##           ar9     ar10   ar11     sar1     sma1
##       -0.0108  -0.0388  0.200  -0.1482  -0.6319
## s.e.   0.0601   0.0597  0.055   0.0725   0.0528
##
## sigma^2 estimated as 0.002864:  log likelihood = 536.09,  aic = -1044.17
```

```
checkresiduals(ts(resid(model6)))
```

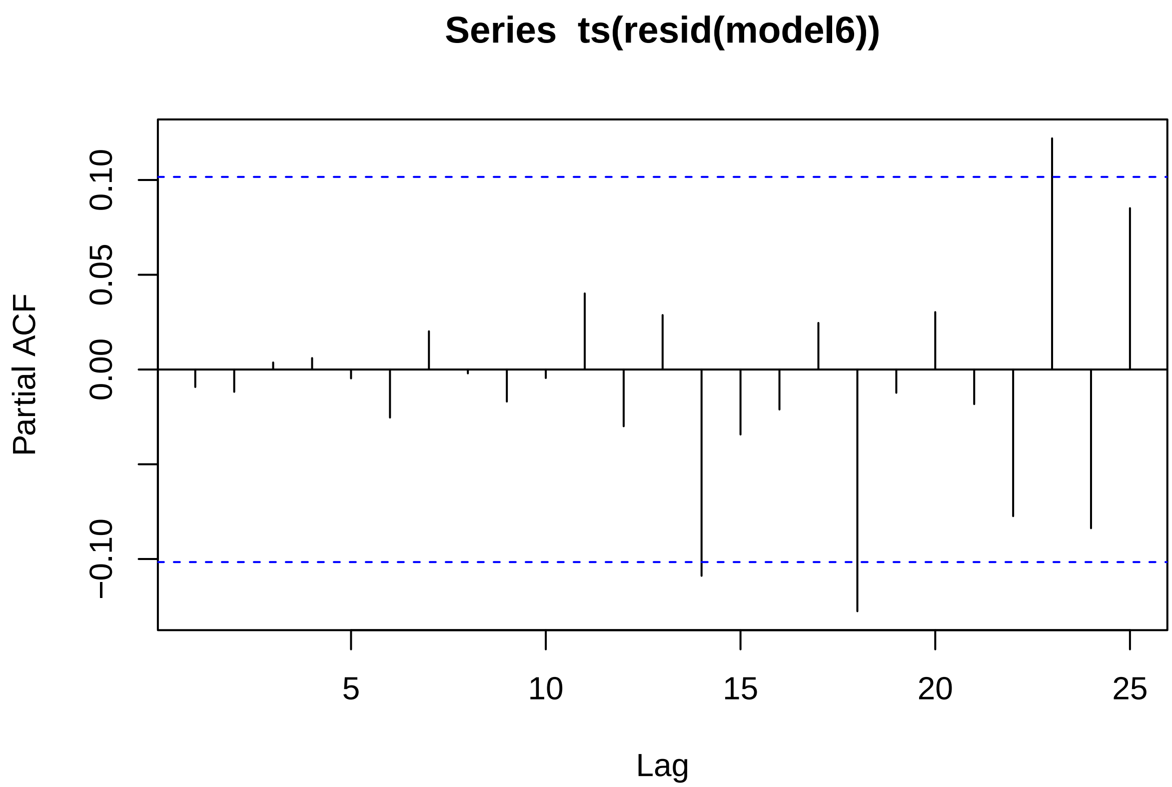
```
## Warning in modeldf.default(object): Could not find appropriate degrees of
## freedom for this model.
```




```
acf(ts(resid(model6)))
```




```
pacf(ts(resid(model6)))
```




```
bartlettB.test(ts(resid(model6)))
```

```
##
##  Bartlett B Test for white noise
##
## data:
## = 0.40032, p-value = 0.9972
```

(i) Describe your fitted ARIMA model.

**Discussion:** As we see both seasonality and trend in the plot of the residuals from part(a), we difference the residuals using orders 1 and 12. We store the differenced residuals in d1d12res3.ts, and use the time series data to create ACF and PACF plots. We analyze the two plots to determine a good place to start in fitting our model.

In our ACF plot we see significant spikes at lags 1, 3 and 11, as well as lags 12, 24, 36 suggesting that an ARIMA(0,1,11)(0,1,1)12 may be a good place to start.

Looking at the PACF plot, we see significant spikes at lags 1, 2 and 11, as well as lags 12 and 24, suggesting that ARIMA(11,1,0)(1,1,0)12 may also be a good candidate.

ARIMA(0,1,11)(0,1,1)12 model has AIC of -1046.76 and ARIMA(11,1,0)(1,1,0)12 has an AIC of -1044.17, so the former is slightly better in terms of AIC. The ARIMA(0,1,11)(0,1,1)12 model also has better ACF/PACF plot results, it has only one significant autocorrelation structure in lag 23 in both ACF and PACF plots.

We test fitting a third model with structure of ARIMA(11,1,11)(1,1,1)12 and see that the model has even better residuals results with no autocorrelation structure remaining visible in the ACF/PACF plot. There is a concern for overfitting, but the model yields an AIC score of -1055.92, which suggests that it is not an overfit model despite the increased number of variables.

We select the ARIMA(11,1,11)(1,1,1)12 as our final model. The model has an AIC of -1055.92, the best from the three models which we originally fit. It has 11 AR and 11 MA variables in the trend component and 1 AR and 1 MA variable in the seasonal component of the ARIMA model. The order of differencing is 1 for trend and 1 for seasonality.

(ii) Tabulate and plot estimated static seasonal indices. Compare the static estimates to those obtained via regression in part 2(b), with a table and a plot. Present the dynamic seasonal index estimates graphically as with the examples in the notes. Discuss the dynamic estimates. What information do they provide about changes over time of seasonal structure?

**Discussion:**

**STATIC ESTIMATES**

In the static seasonal index estimates we see that the regression and ARIMA estimates differ in several of the months. The table and plot of the indices are shown below.

**DYNAMIC ESTIMATES**

In the dynamic seasonal index estimates we do see presence of dynamic seasonality. Perhaps most prominent is the uniform increase of seasonal index in the months of March and November from 1990s until the end of the time series. We see in the graph that indices for both months cross their respective static seasonal indices level around the year 2001.

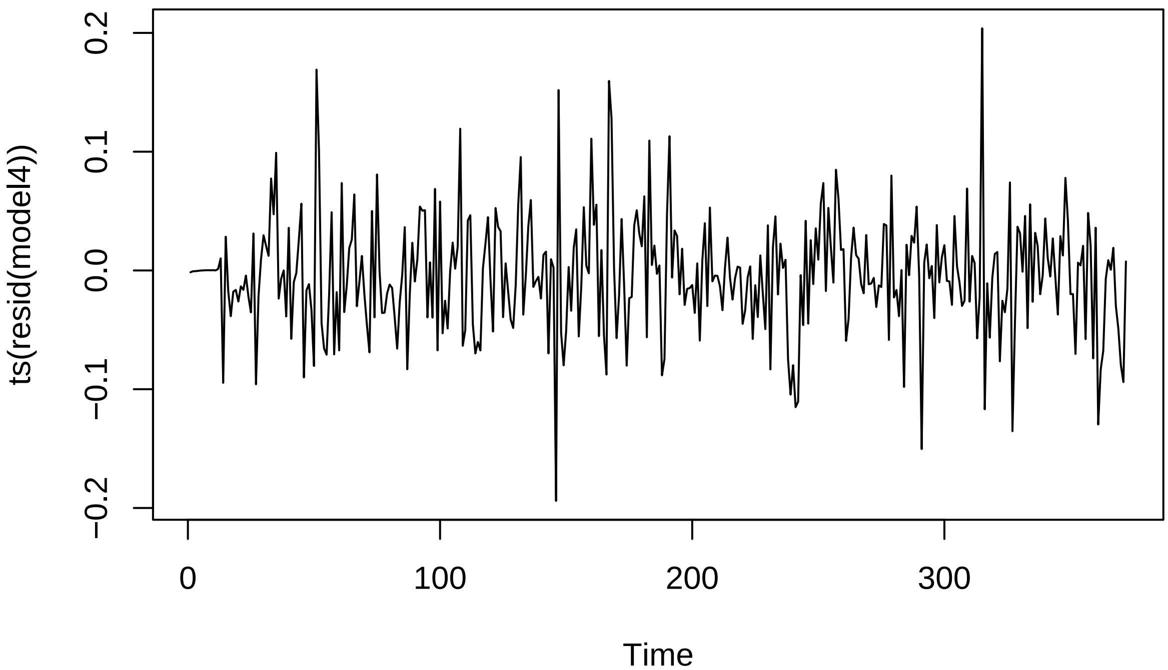
Otherwise there is much fluctuation in each month's seasonal indices over the span of three decades. Below are observations of each month's indices.

**January**: Indices are generally higher in 1990-2000 than in 2000-2019 **February**: Indices in february show slight decreasing trend since 1990s, although throughout indices fluctuate consistently. **March**: Strong evidence that march indices have generally increased over the years **April**: There is no apparent trend in increase or decrease over the years - however indices do fluctuate consistently around the April seasonal index **May**: A slight decreasing trend until 2010 and a slight increasing trend thereafter. **June**: June indices are consistently above 1.00, meaning imports have always been above the level of trend in the month of June. There is fluctuation in indices with perhaps a slight downward trend. **July**: Apparent fluctuation of indices over three decades with perhaps a slight downward trend **August**: August has strong volatility in terms of fluctuation of indexes over the years, perhaps one of the strongest. Indices are always above 1.00 in the month of august. **September**: Apparent fluctuation of indices over three decades with perhaps a slight upward trend **October**: Relatively stable fluctuation around the static seasonal index over three decades **November**: Strong evidence that november indices have generally increased over the years **December**: Indices indicate that imports consistently fall below the level of trend in the month of December. Indices fluctuate around the static seasonal index and are generally higher in the year 2000 and after.

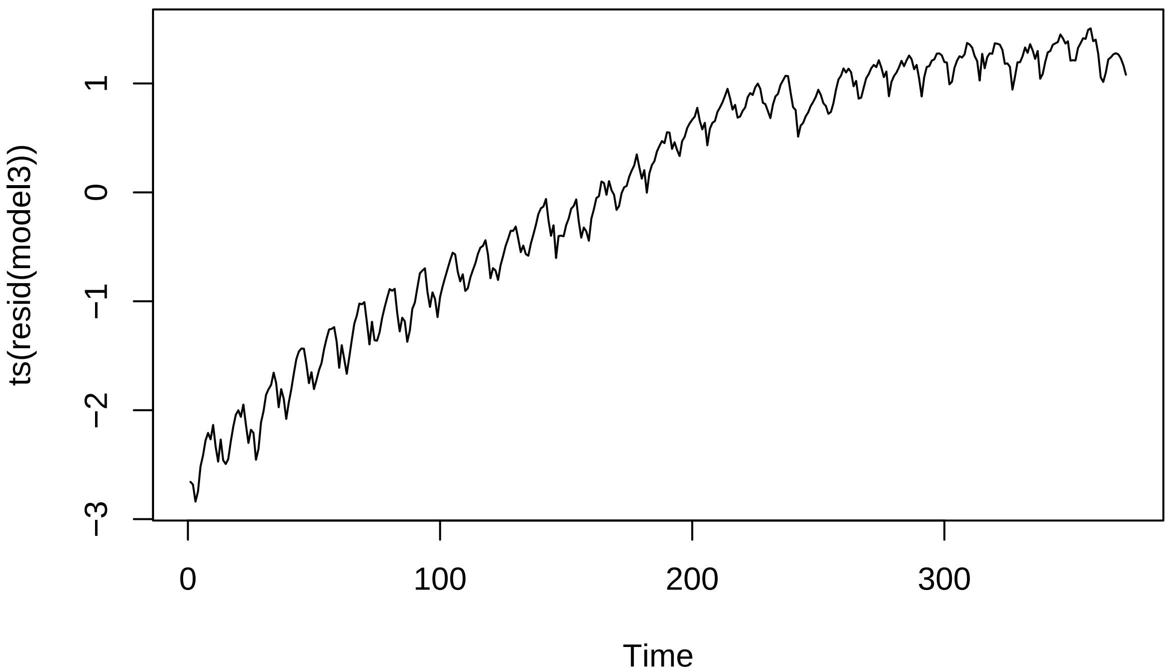
**STATIC ESTIMATES TABLE AND PLOT**

```
sel<-1:12
arimapred<-resid(model3)[-sel]-resid(model4)[-sel]
```

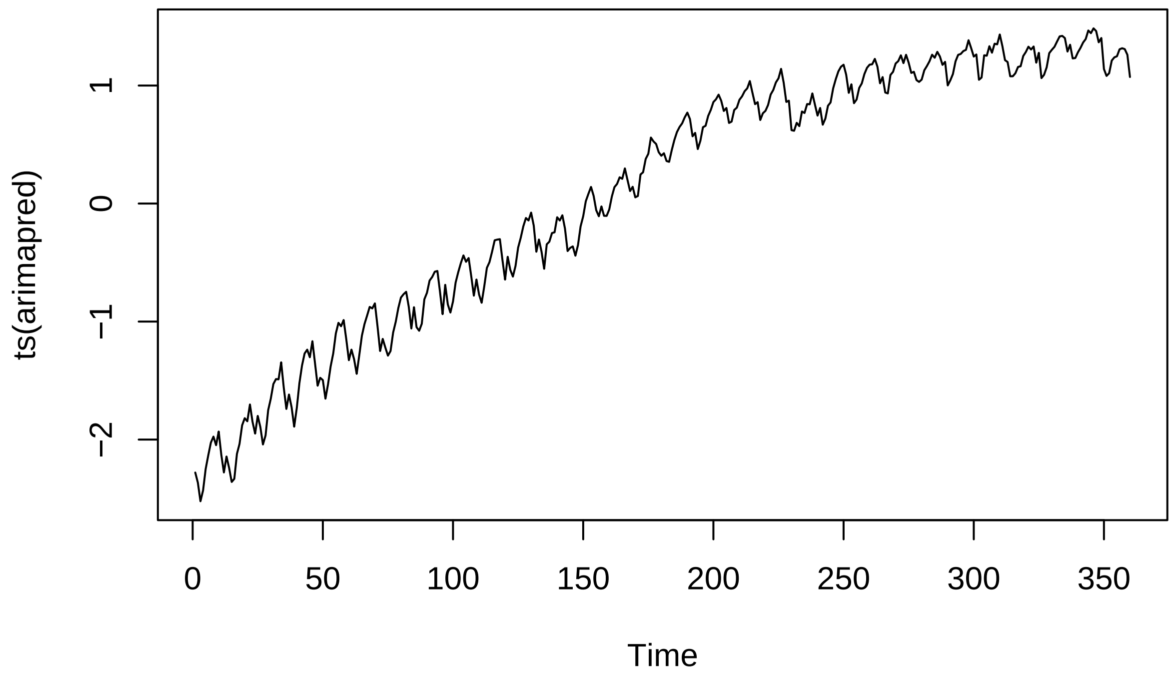
```
plot(ts(resid(model4)))
```



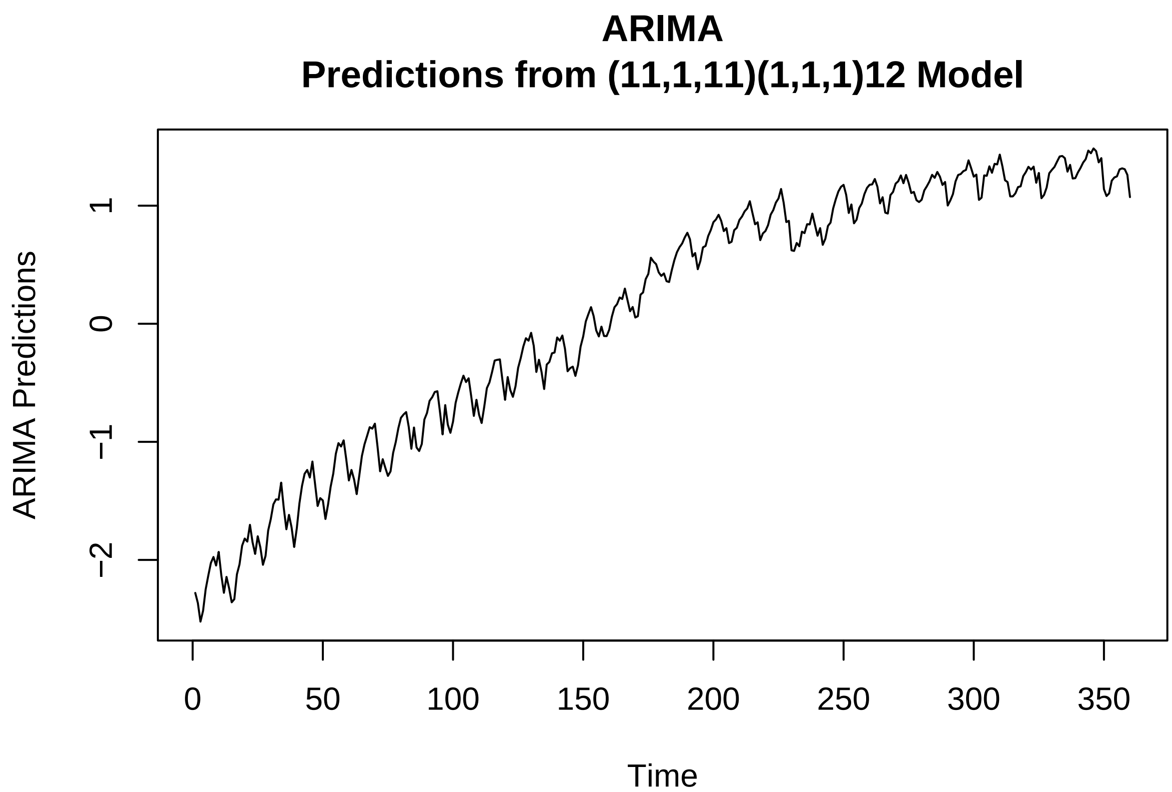

```
plot(ts(resid(model3)))
```




```
plot(ts(arimapred))
```

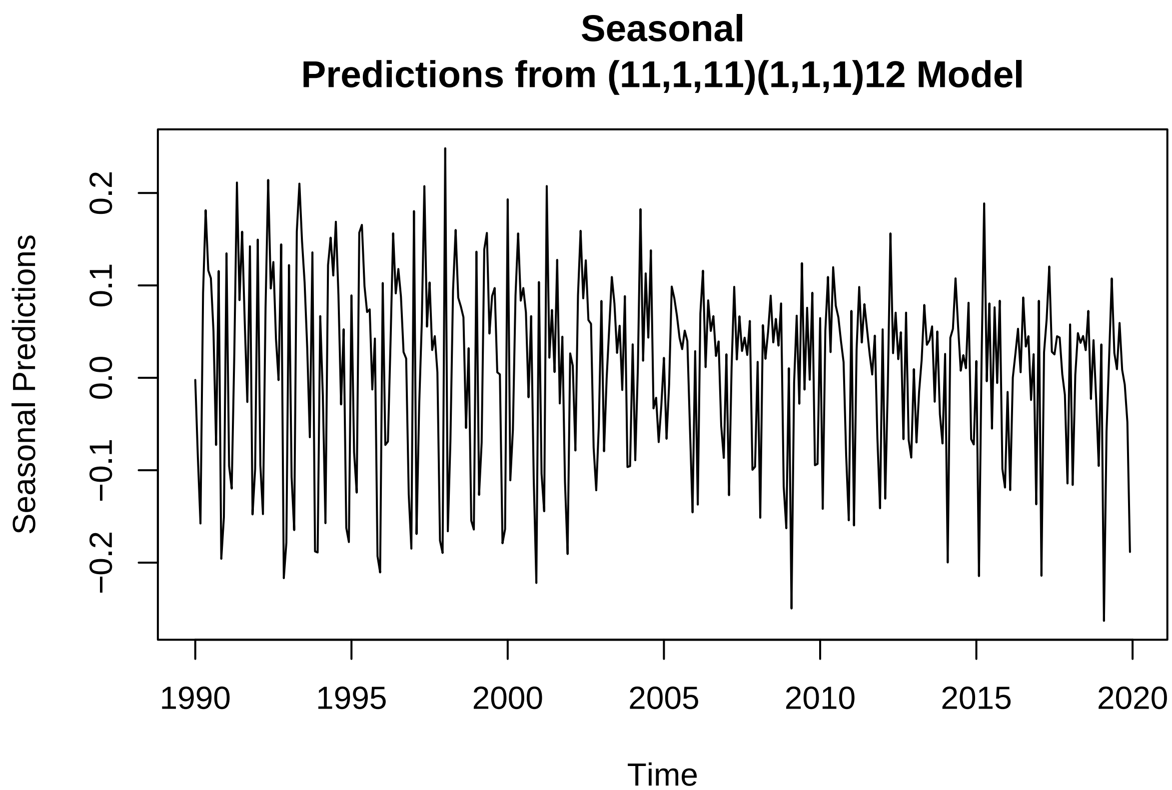



```
arimapred.ts<-ts(arimapred)
plot(arimapred.ts,xlab="Time",ylab="ARIMA Predictions",main="ARIMA
Predictions from (11,1,11)(1,1,1)12 Model")
```




```
# we do see trend remaining so we attempt to remove trend using differencing
model7 <- arima(arimapred.ts,order=c(0,1,0),seasonal=list(order=c(0,0,0),period=12))
arimapred2<- resid(model7)
```

```
arimapred2.ts<-ts(arimapred2,start=c(1990,1),freq=12)
plot(arimapred2.ts,xlab="Time",ylab="Seasonal Predictions",main="Seasonal
Predictions from (11,1,11)(1,1,1)12 Model")
```




```
monmeans<-tapply(arimapred2,imports$Month[-sel],mean)
seas2<-monmeans-mean(monmeans)
seas2_indices <- exp(seas2)
```

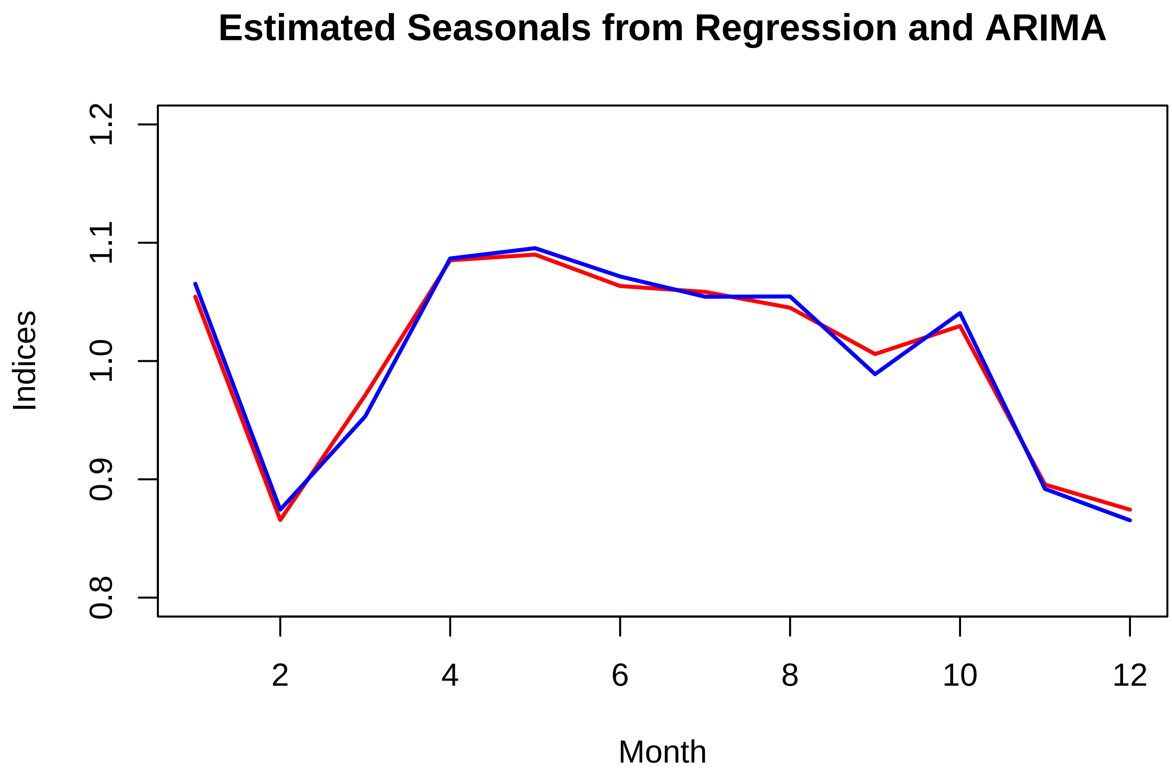
### TABLE OF SEASONAL INDICES FROM REGRESSION AND ARIMA

```
cbind(1:12,seas_indices,seas2_indices)
```

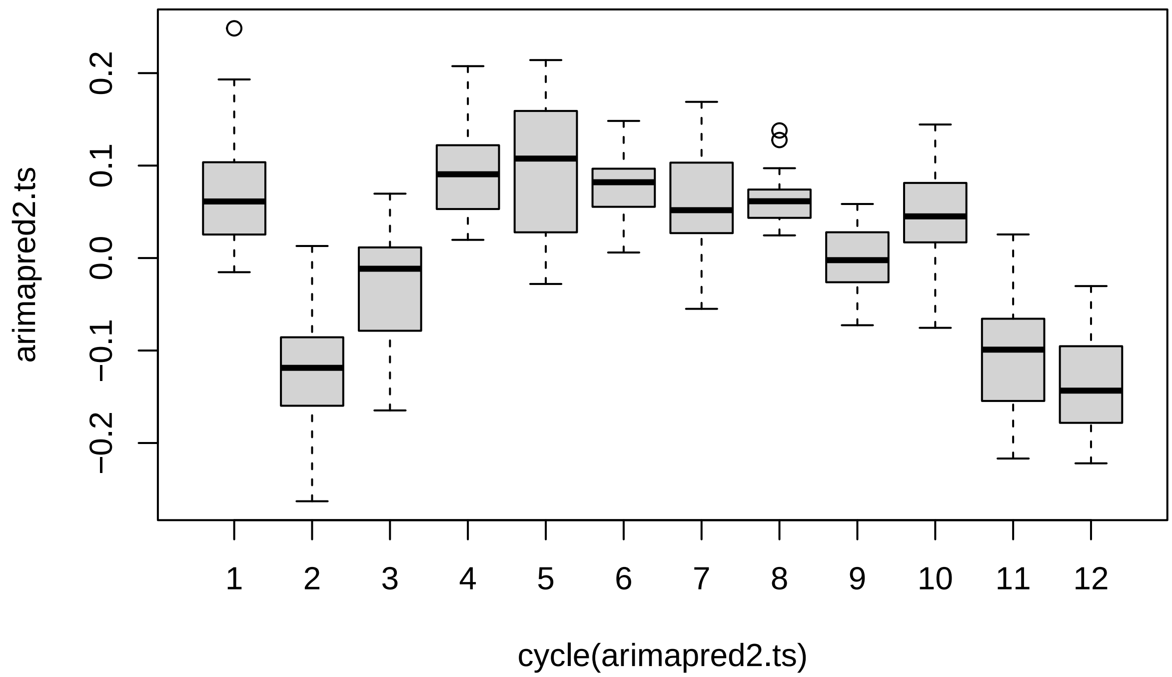
```
##                  seas_indices seas2_indices
## (Intercept)  1    1.0543753     1.0653650
## fMonth2      2    0.8656234     0.8743382
## fMonth3      3    0.9710398     0.9532041
## fMonth4      4    1.0851728     1.0866749
## fMonth5      5    1.0899601     1.0954021
## fMonth6      6    1.0634049     1.0714824
## fMonth7      7    1.0584539     1.0543535
## fMonth8      8    1.0450652     1.0545830
## fMonth9      9    1.0058873     0.9888734
## fMonth10    10    1.0296245     1.0405858
## fMonth11    11    0.8955758     0.8918844
## fMonth12    12    0.8743488     0.8653340
```

### PLOT OF SEASONAL INDICES FROM REGRESSION AND ARIMA

```
plot(ts(seas_indices),xlab="Month",ylab="Estimated Seasonal
Indices",main="Estimated Seasonals from Regression and ARIMA",ylim=c(0.8,1.2),lty=1,lwd=2,col="red")
lines(ts(seas2_indices),lty=1,lwd=2,col="blue")
```



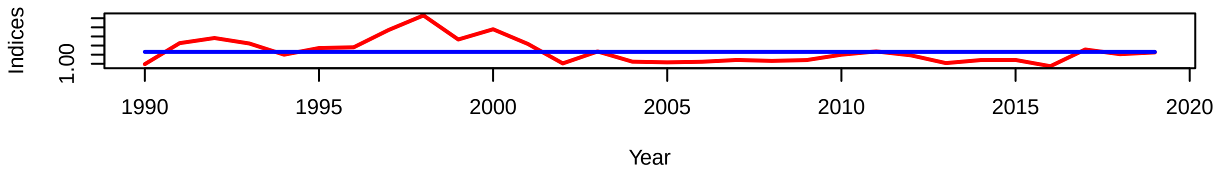

```
arimapred2.ts<-ts(arimapred2,start=c(1990,1),freq=12)
boxplot(arimapred2.ts~cycle(arimapred2.ts))
```



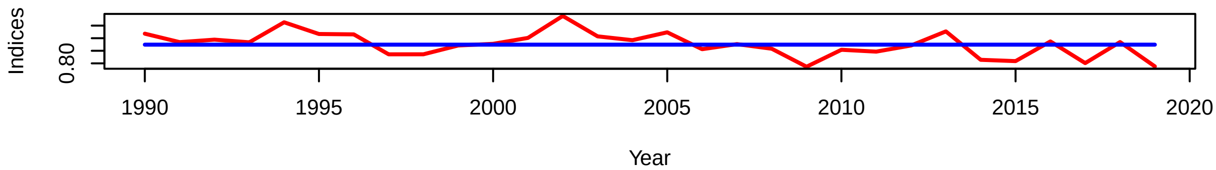

**DYNAMIC ESTIMATES**

```
y<-arimapred2
seasm<-matrix(rep(0,360),ncol=30)
j<--11
for(i in 1:30){
j<-j+12;j1<-j+11
seasm[,i]<-exp(y[j:j1]-mean(y[j:j1]))
}
year<-seq(1990,2019)
seas2m<-matrix(rep(seas2_indices,30),ncol=30)
name<-
c("January","February","March","April","May","June","July","August","Se
ptember","October","November","December")
par(mfrow=c(3,1))
for(i in 1:12){
plot(year,seasm[i,],xlab="Year",ylab="Indices",main=name[i],type="l",lwd=2,col="red")
lines(year,seas2m[i,],lty=1,lwd=2,col="blue")
}
```

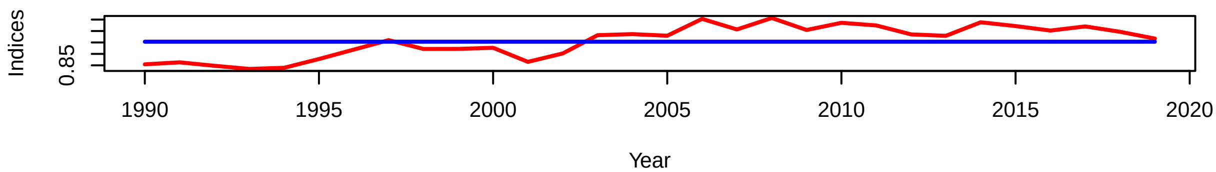
**January**

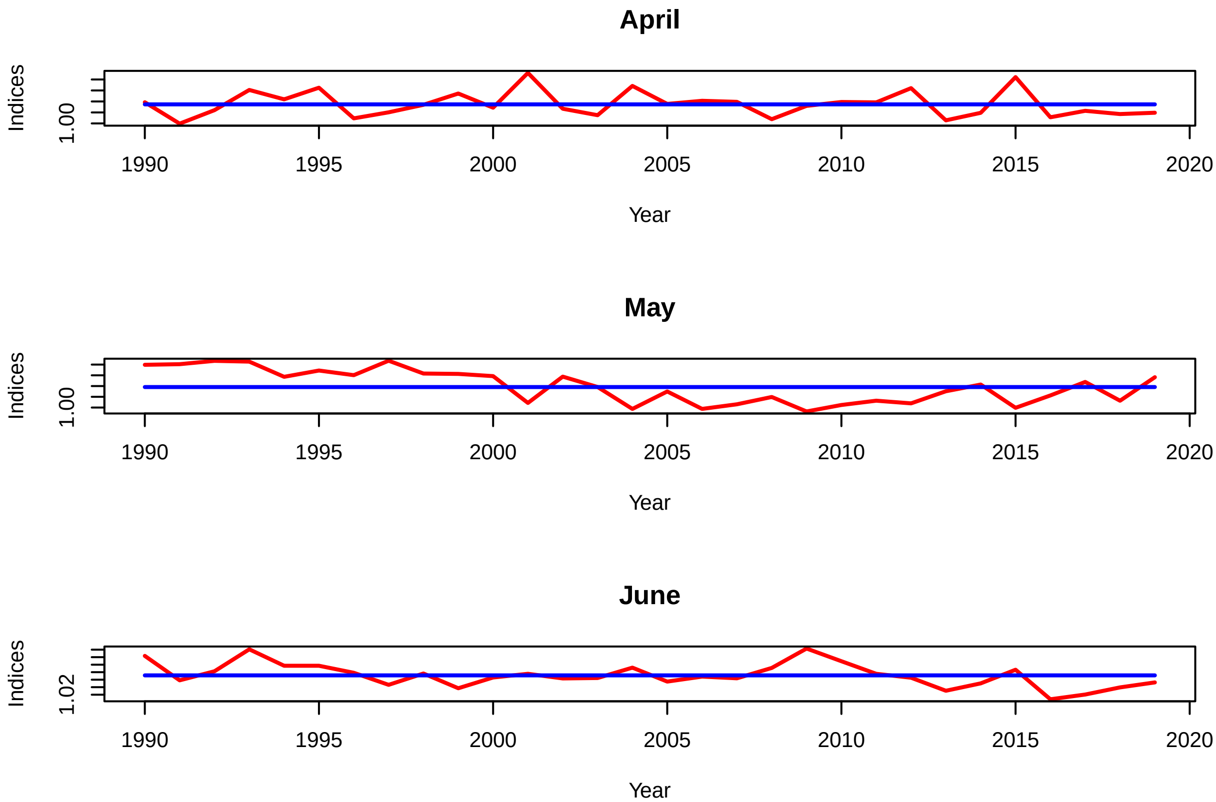


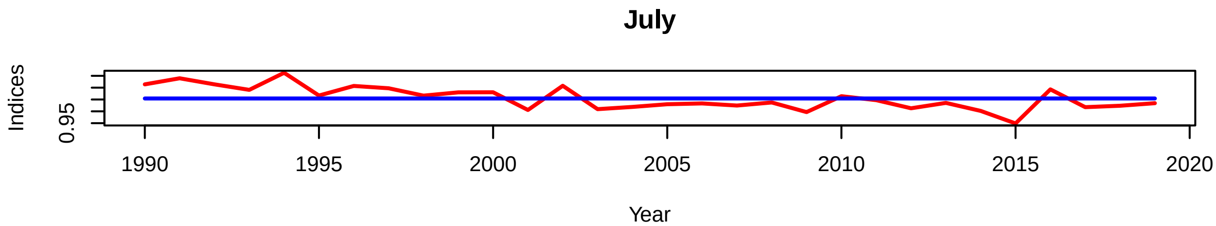

**February**

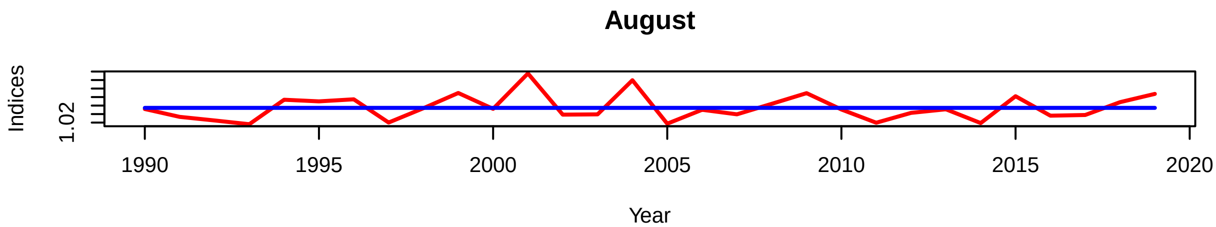


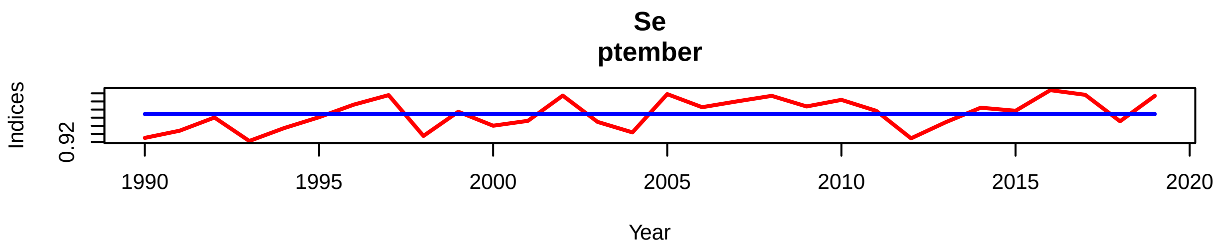

**March**



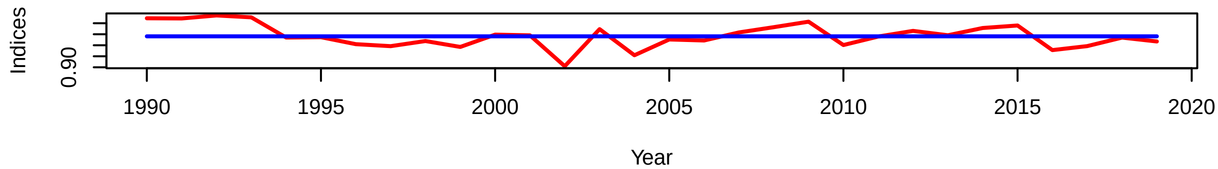

April
Indices
1.00
1990
1995
2000
2005
2010
2015
2020
Year
May
Indices
1.00
1990
1995
2000
2005
2010
2015
2020
Year
June
Indices
1.02
1990
1995
2000
2005
2010
2015
2020
Year


July
Indices
0.95
1990
1995
2000
2005
2010
2015
2020
Year

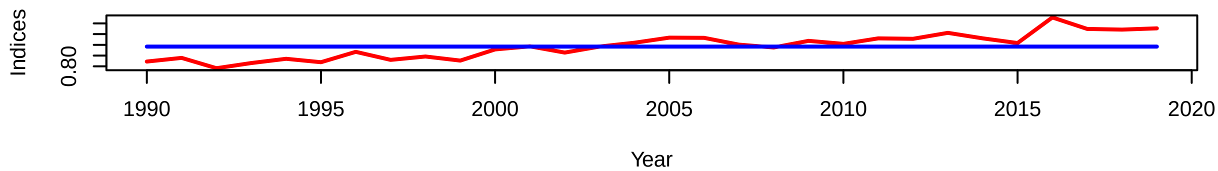

August
Indices
1.02
1990
1995
2000
2005
2010
2015
2020
Year


Se
ptember
Indices
0.92
1990
1995
2000
2005
2010
2015
2020
Year

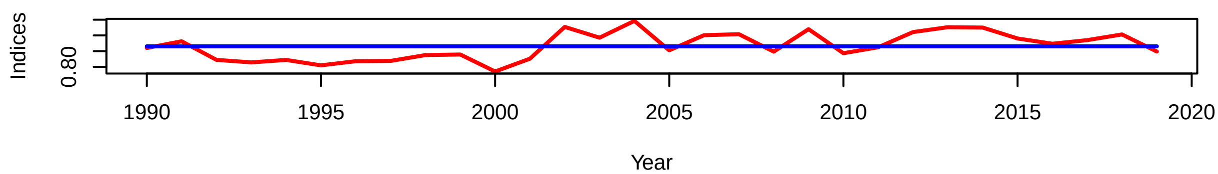
**October**




**November**




**December**




(iii) Perform a residual analysis for your ARIMA model, examining the plot of the residuals vs. time, the residual normal quantile plot, the residual acf and pacf, and the residual spectral density (along with Bartlett's test). Has the ARIMA model you've fit produced reduction to white noise? Discuss carefully.

**Discussion:**

Yes, the ARIMA model produced reduction to white noise.

**Residuals vs. Time Plot** - The residuals vs. time plot shows a flat residuals across time, suggesting that trend has been sufficiently captured. We also do not see fluctuations which indicate the presence of seasonality, suggesting that seasonality has also been captured by our mode.

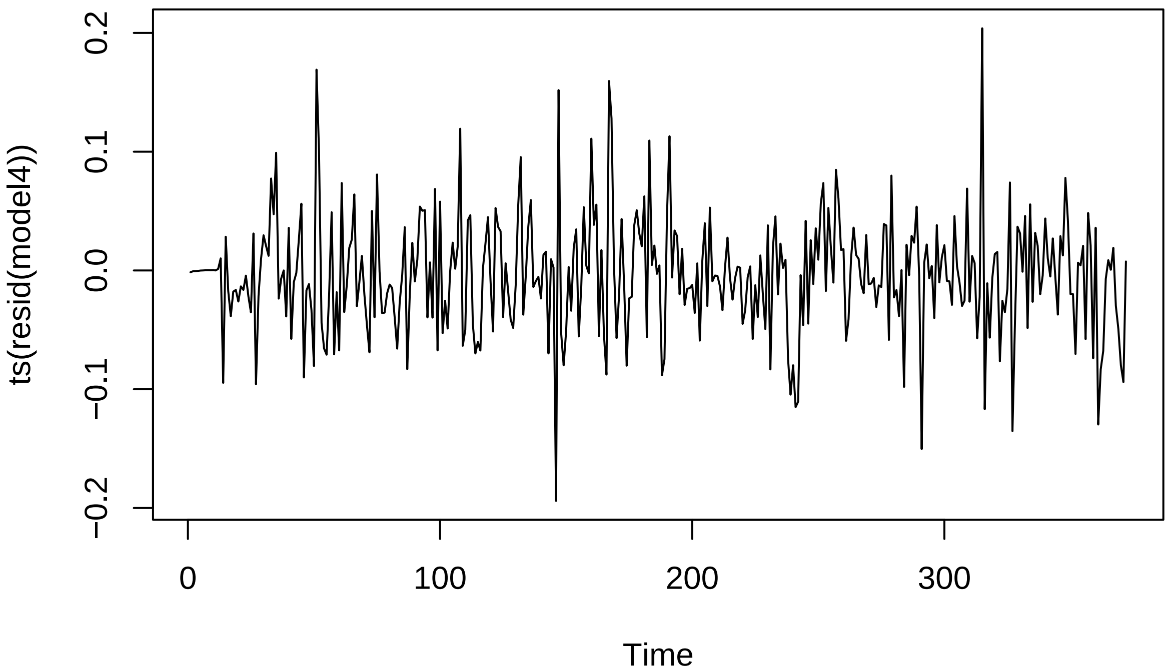
**Normal Quantile Plot** - The majority of residuals fall on the normal quantile plot line, which suggest that we cannot reject the null hypothesis that the residuals are normally distributed.

**ACF / PACF plots** - The ACF and PACF plots indicate no remaining autocorrelation structures in the model.

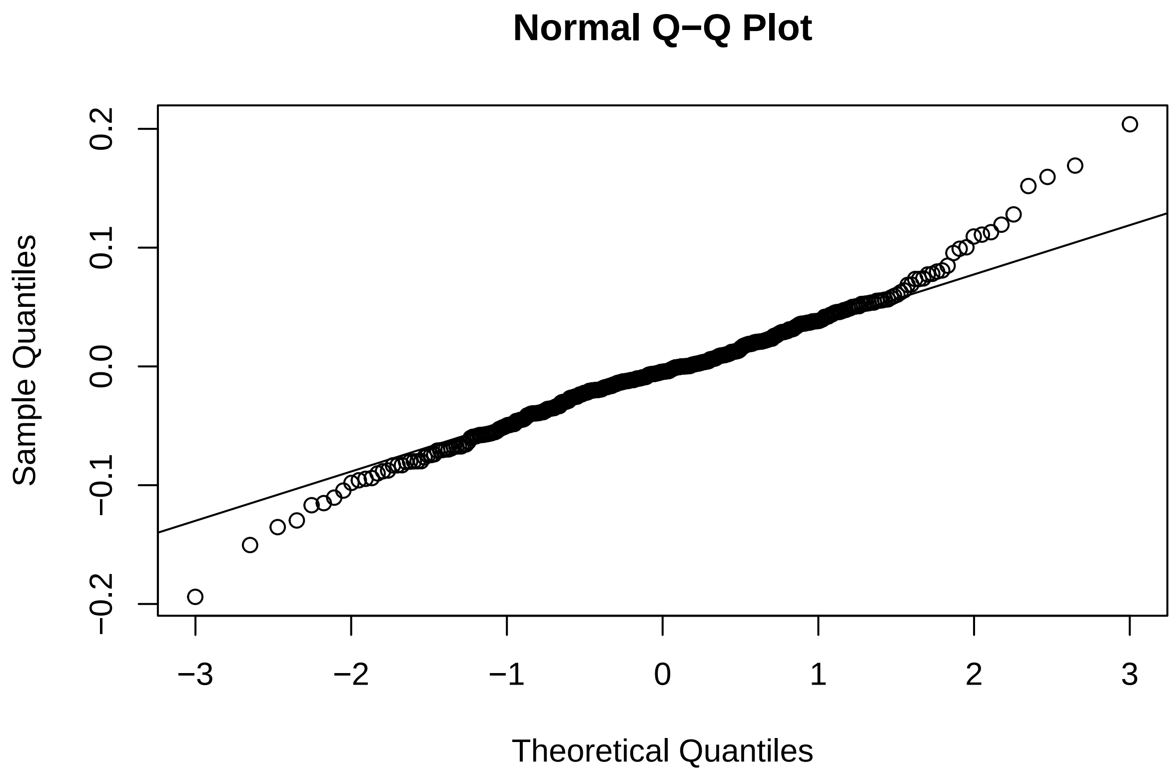
**Residual Spectral Density** - The spectral density plot shows that the distance between the highest and lowest peak of the spectral density graph is less than twice the upper half of the blue measurement line above the notch, therefore we can conclude that the model has been reduced to white noise. Also, the plot indicates no significant seasonal or calendar effect within the residuals, as indicated by the lack of prominence in the blue and red dashed lines in the graph.

**Bartlett's B test** - Bartlett's test shows a P-value of 1, a clear indication that the model has been reduced to white noise.

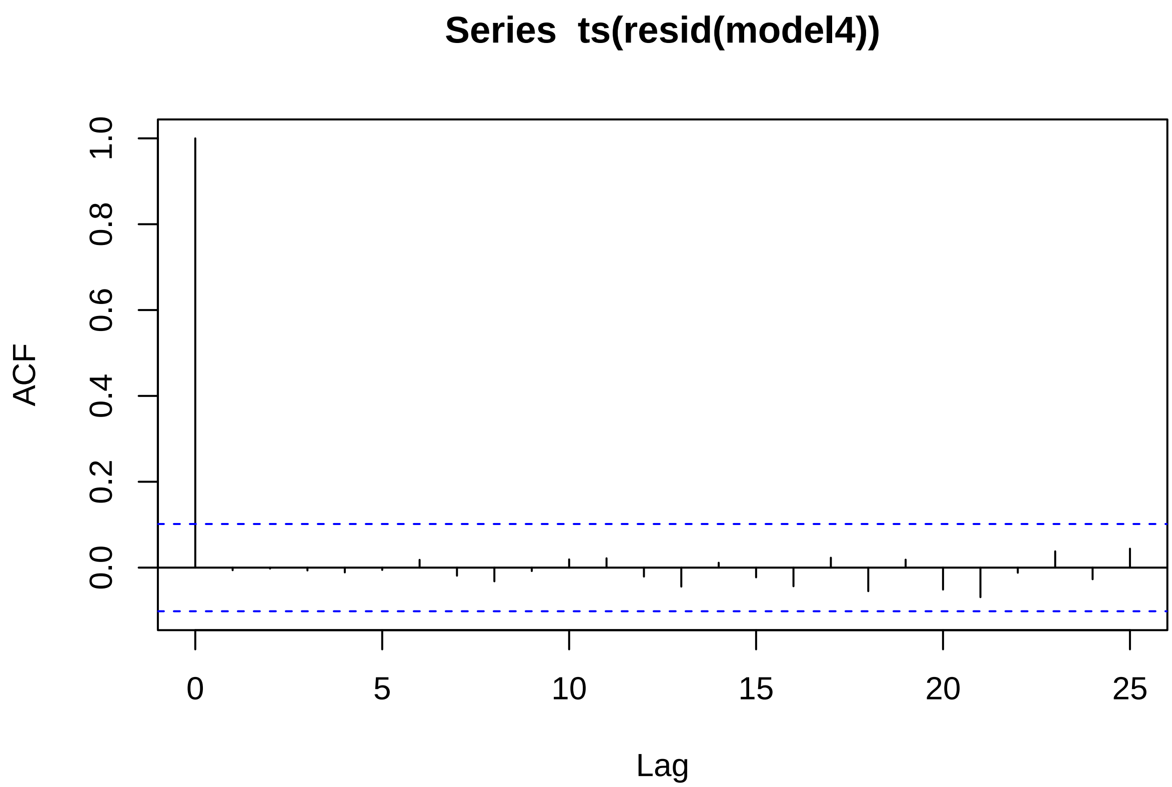
```
plot(ts(resid(model4)))
```



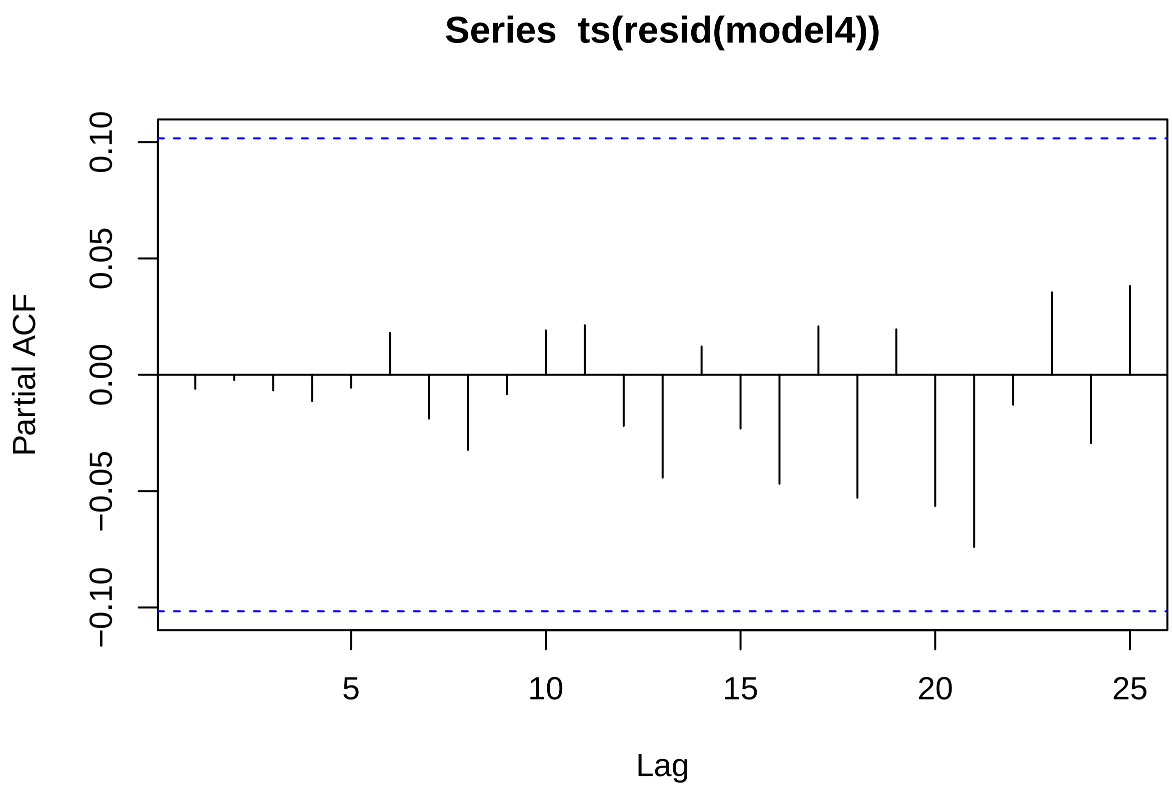

```
qqnorm(resid(model4))
qqline(resid(model4))
```



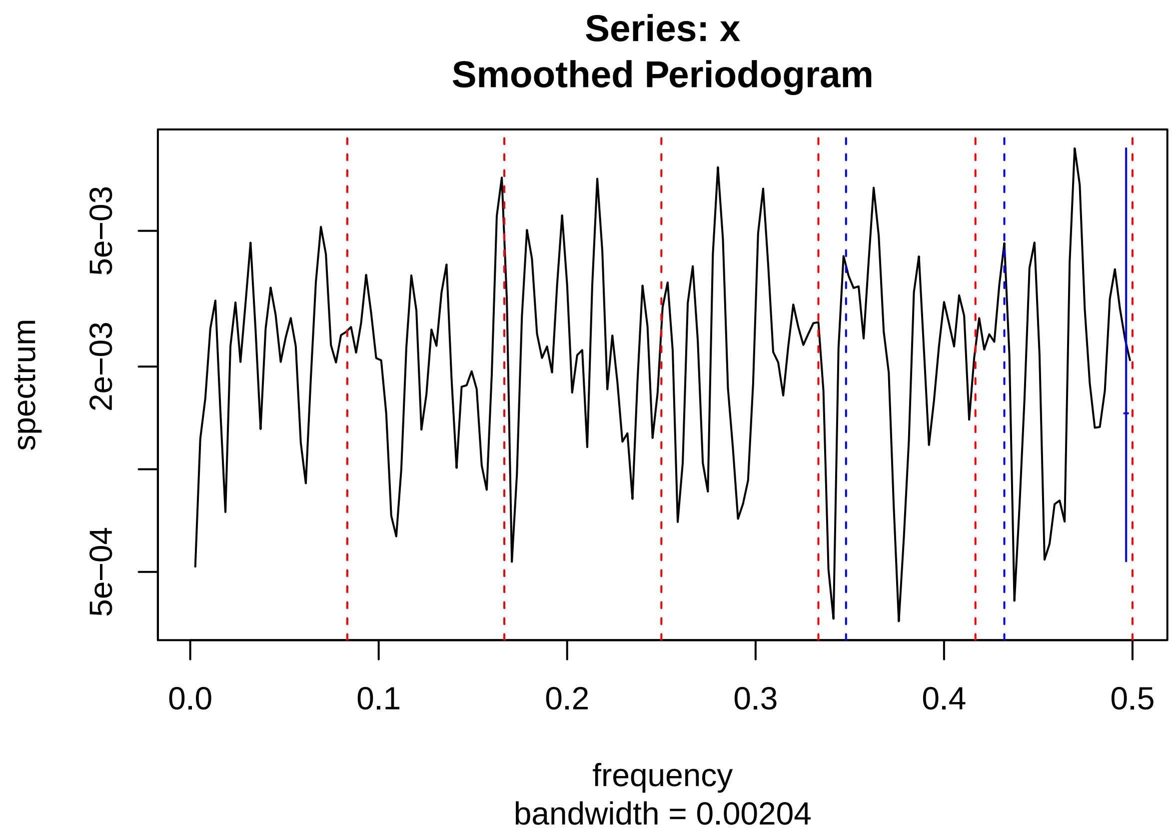

```
acf(ts(resid(model4)))
```




```
pacf(ts(resid(model4)))
```




```
spectrum(resid(model4), span=3)
abline(v=c(1/12,2/12,3/12,4/12,5/12,6/12),col="red",lty=2)
abline(v=c(0.348,0.432),col="blue",lty=2)
```




```
bartlettB.test(ts(resid(model4)))
```

```
##
##  Bartlett B Test for white noise
##
## data:
## = 0.30128, p-value = 1
```

4. Give some brief concluding remarks. What has the analysis revealed about imports from China to the U.S. during the years 1989 to 2019?

**Discussion:** Overall, imports from China have dramatically increased from 1989 to 2019. The seasonality of these imports change during that same time span. The indices analysis from part three indicates that months such as March and November sees more imports relative to other months over the years whereas months such as February and July see less imports relative to other months in the same span. We also see that historically economic downturns causes imports from China to dip.