Time Series Analysis of Monthly Ice Cream and Frozen Dessert Industrial Production in the United States

Yujie Ye

Perm: 8161861

PSTAT 174 Time Series

11 December 2022

Abstract

The industrial production of ice cream and frozen dessert is an important factor for relevant enterprises to calculate their profits and estimate the frozen food market. The goal of this study is to predict the ice cream and frozen dessert production by time series models. We applied monthly data from Federal Reserve Bank of St Louis website to analyze the production between 1981.01 and 1992.12 in the U.S.. Employing the Box-Jenkins methodology, we log transformed the original data, and selected a suitable model by checking auto-correlation function (ACF) and partial auto-correlation function (PACF) plots of transformed data. Also, we compared Akaike information criterion (AICc) values for choosing model. After diagnostic checking of our potential model, the final model is $SARIMA(0,1,1)(0,1,1)_{12}$. The model passes tests related to prove the residual of the model follows the assumption of white noise. We successfully forecasted the production of frozen food by our time series model and all predicted values are within the 95% confidence intervals.

1. Introduction

The industrial production of frozen food is considered as an indispensable variable for companies manufacturing frozen food when they develop business plans, because it is related tightly with their profit strategy. Additionally, the industrial production of frozen food could reflect the market conditions of frozen food, providing effective information to government or enterprises to make further strategies. The dataset I applied in this study comes from Federal Reserve Bank of St Louis. It is a monthly data containing total 144 observations. The original dataset was split into two parts: training part includes 132 observations, presenting the industrial production index of ice cream and frozen dessert in the United States from 1981.01 to 1991.12; testing part includes 12 observations, demonstrating the industrial production index of ice cream and frozen dessert in the United States from 1992.01 to 1992.12. The unit of the data is the production index setting the production of 1984 as 100.

1.1 Problem & Results

This study aims to build a time series model to forecast the industrial production of ice cream and frozen desserts in the U.S. from 1992.01 to 1992.12. We applied Box-Cox transformation to make our original data have a more constant variance. After that, we removed seasonal component and trend of data by differencing. Based on model parsimonious assumption, we used ACF/PACF plots and AICc values to choose the most suitable model: $SARIMA(0,1,1)(0,1,1)_{12}$ from several model candidates. In addition, after checking whether the residual of the model follows white noise assumptions by analyzing residual-related plot, histogram, Q-Q plot, ACF/PACF, Shapiro-Wilk normality test, Box-Pierce test, Ljung-Box test and Yule-walker method, we use the final model to forecast. We successfully forecasted the production of frozen food by our time series model and all predicted values are within the 95% confidence intervals.

1.3 Acknowledge

The data source is from Board of Governors of the Federal Reserve System (US). The software used is R 4.2.0 and the packages used are MASS, qpcR, forecast, stats, ggplot2.

2. Methodology

In this section I will present my time series analysis of my dataset, including data analysis, data transformation, model identification, model diagnoses, and forecast.

2.1 Original Data Processing

The time series plot of original data shows a strong seasonal component and an upward trend. There is a cyclical behavior in the ACF and there are significant correlations with values moving proportionally every 12 lags. Therefore, we can see that the period of the seasonal component is given by s = 12. Moreover, the production of ice cream and frozen dessert in winter was much lower than that in summer every year and it increased steadily from 1981 to 1992. Besides, we need transformation to get a stationary series, because we can find the histogram plot of original data does not exhibit a symmetric and normal pattern, but presents a right skewed distribution, and the variance of original data is not constant.

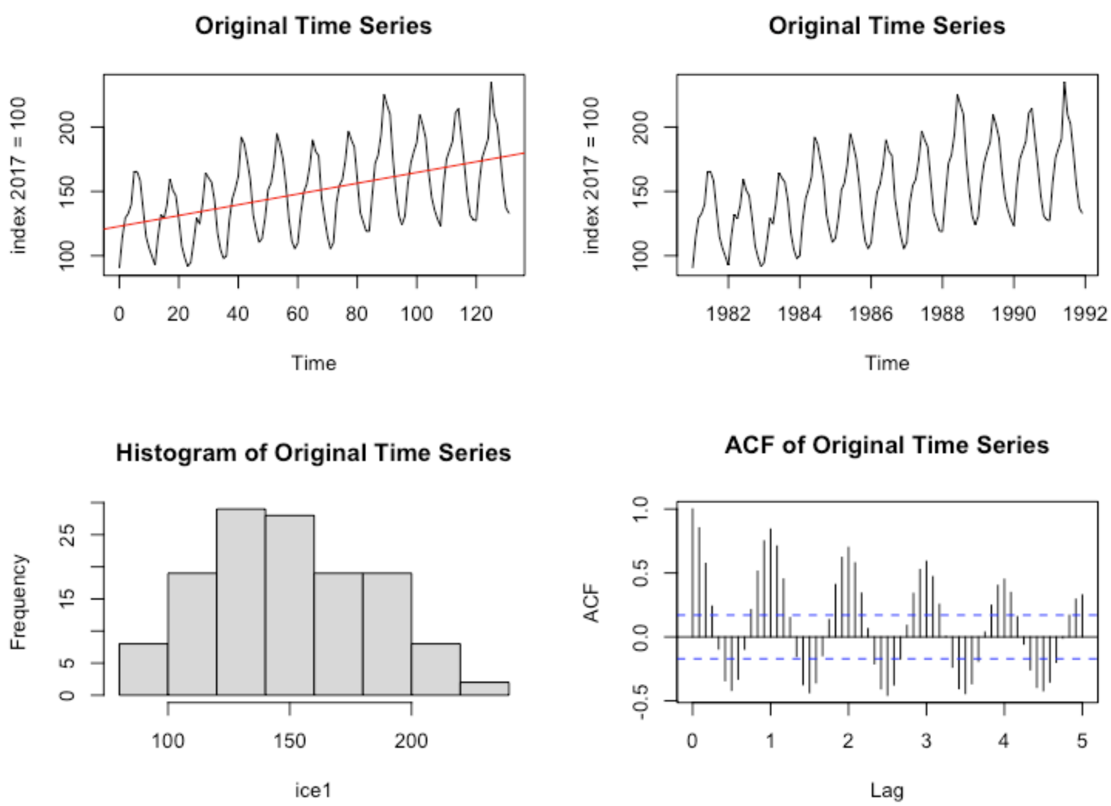


Figure 1: Time series plots, histogram, ACF plot of the original dataset

We applied Box-Cox transformation method to choose the way we transform. The confidence interval of Box-Cox transformation includes 0 and 1, and the log-likelihood is the largest when $\lambda = 0.3434$. Consequently, we can do experiment with log transformation, non-transformation, and box-cox transformation. Compared all three transformation ways, as figure 3 shows, we chose log transformation as the most appropriate method since the variance of the data is more constant than non-transformation, and the histogram plot of log transformation shows a more symmetric normal distribution. As for the Q-Q plot, there is not significant difference among them. In the following report, I used log transformation of data for further analysis.

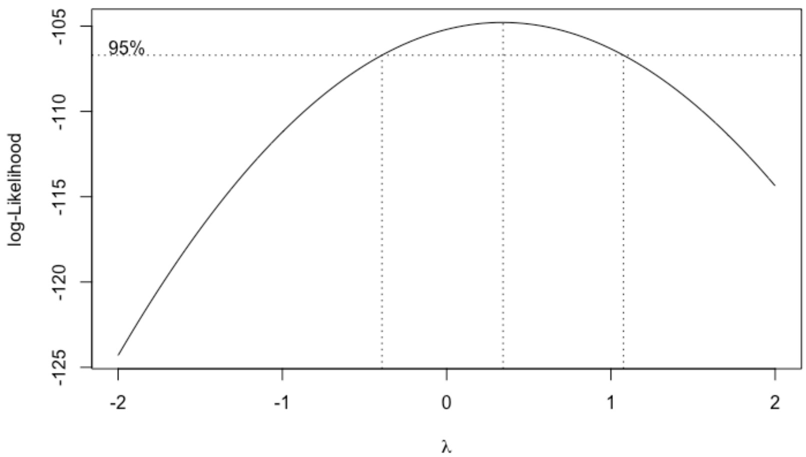


Figure 2: Log-Likelihood of the original dataset

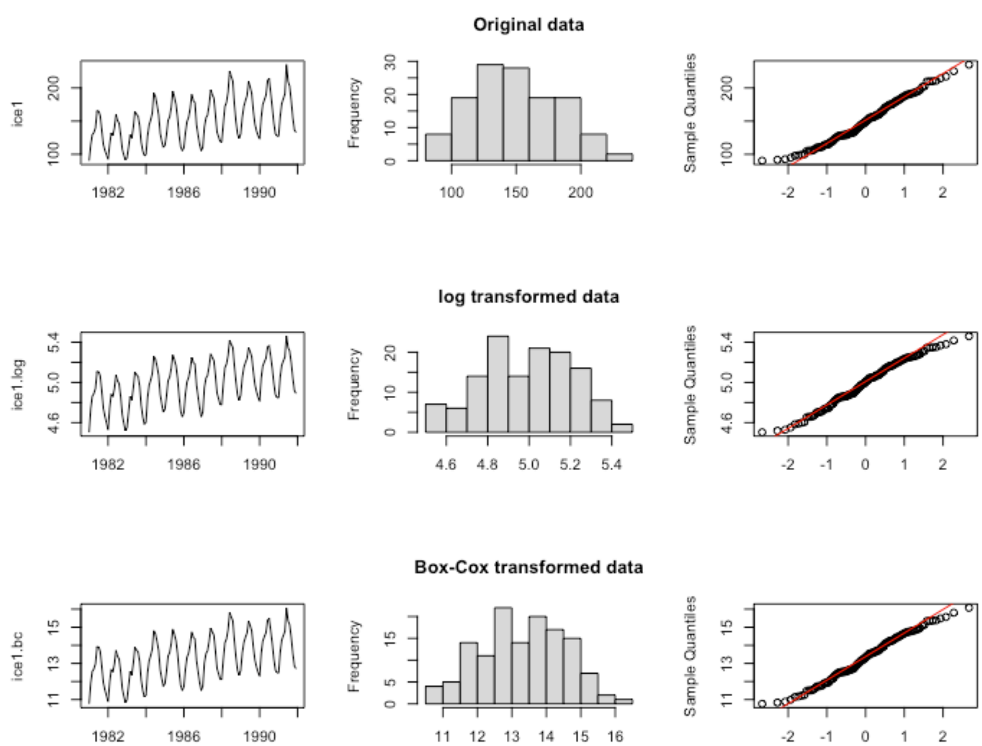


Figure 3: Time series plots, histogram, Q-Q plots of original dataset/ log transformed dataset / Box-Cox transformed dataset

To make our dataset stationary, we need to remove the seasonality and trend of the data. We predict that differencing at lag 12 and lag 1 could make our data to be stationary based on decomposition plot shown in figure 4. As figure 5 shows, by differencing at lag 12, the seasonality of the data is removed, however, a weak periodic trend could still be visible and the mean of it is not equal to 0. We can also find the ACF plot in figure 6 is also show that trend, which presents that it is not stationary. After differencing at lag 1, which helps to remove the trend, the dataset shows stationary and the mean of the data is really close to 0 (figure 5). We can also find there is no decaying pattern and shows stationary in ACF and PACF plots shown in figure 6.

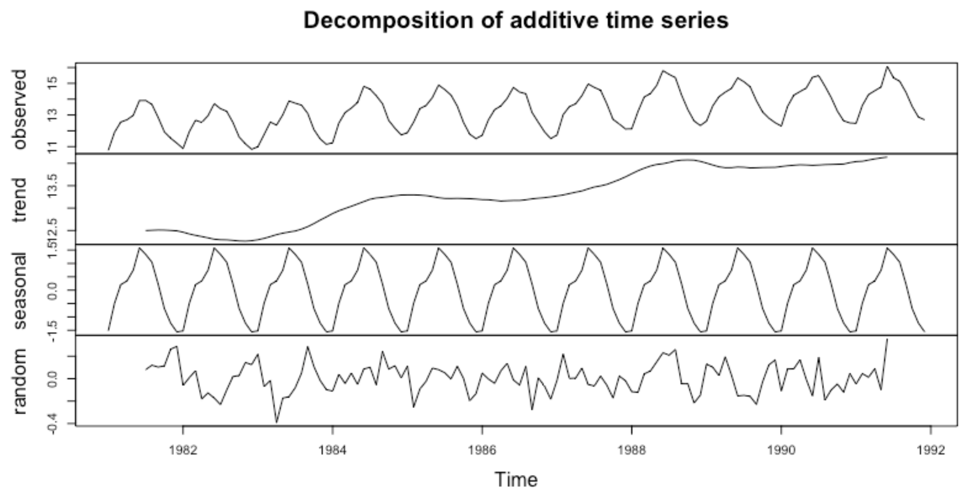


Figure 4: Decomposition of additive time series plot for log transformed dataset

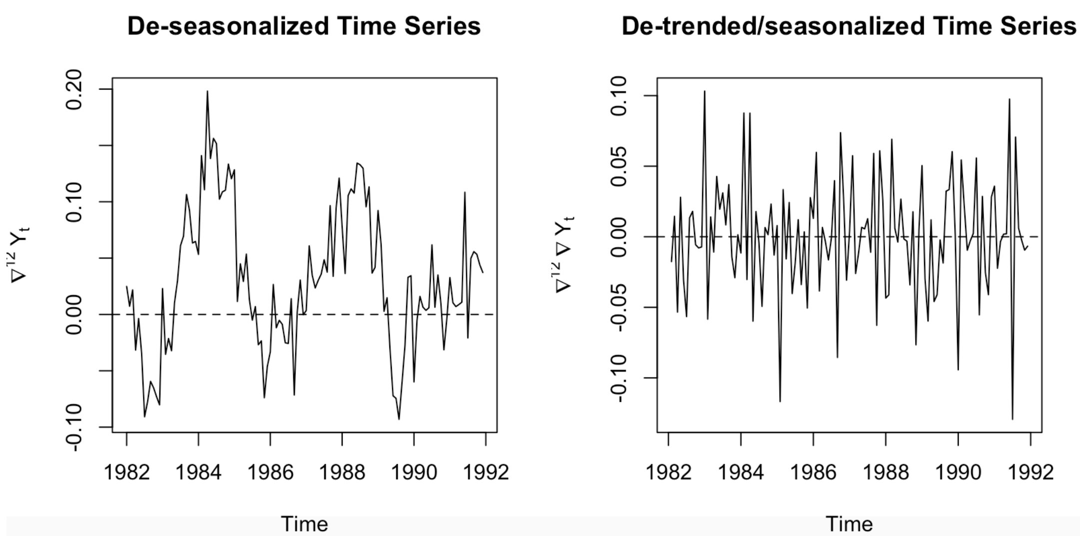


Figure 5: Time series plots of de-seasonalized *(left)* and de-trended/seasonalized *(right)*

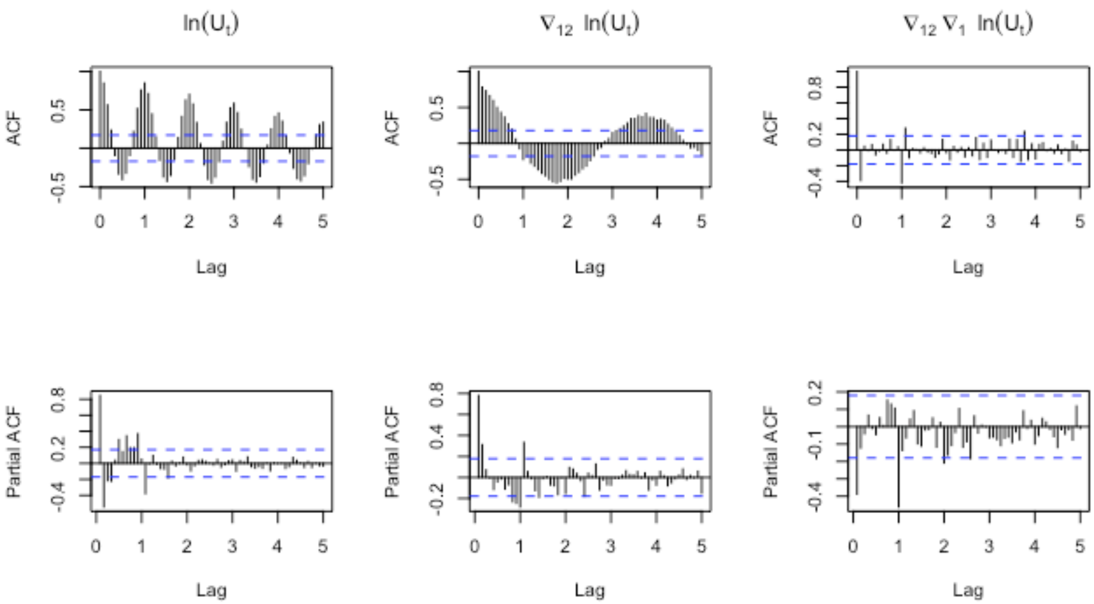


Figure 6: ACF *(top)* / PACF *(bottom)* plots of data differenced at different lags

(Let U_t be the industrial production of ice cream and frozen dessert in the U.S. at time t.)

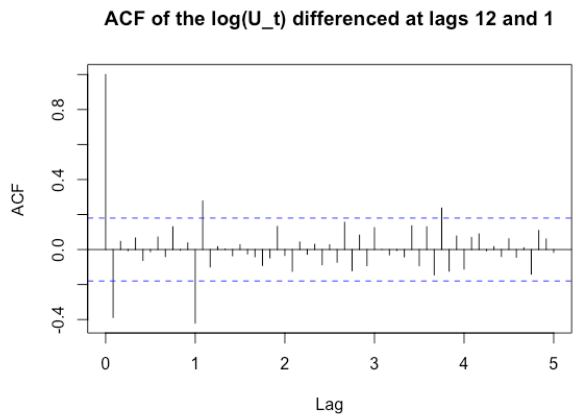
To check if the differencing is sufficient, we kept making differencing at lag 1 and compared the variances among original data, data differencing at lag 12, data differencing at lag 12 and lag 1, and data differencing at lag 12 and twice lag 1. As table 1 shows, we can discover that the variance is minimal when data differencing at lag 12 and only once lag 1, and if we do once more differencing lag 1 the variance is getting larger, which indicates that we do not need further differencing.

	$\ln(U_t)$	$\nabla_{12} ln(U_t)$	$\nabla_{12}\nabla_1 ln(U_t)$	$\nabla_{12}\nabla_1\nabla_1 ln(U_t)$
Variance	0.049221	0.0038787	0.0017117	0.0047897

Table 1: Comparison of log transformed data differencing at different lags

2.2 Model Identification

According to the ACF and PACF of data differencing at lag 12 and lag 1, we could choose a SARIMA model ($SARIMA(p,d,q)(P,D,Q)_s$) for the dataset. We applied one seasonal differencing and one differencing to remove the trend, so we should choose D = 1 at lag s = 12 and d = 1. Based on ACF plot, there is a strong peak at lag 12, which suggests that Q could choose 1. Additionally, the ACF plot shows a value outside of the confidence interval at lag 13, which may be caused by lag 1, so we can choose q = 1. As for PACF plot, there is a strong peak at lag 12 and we can also find a value outside of the confidence interval at lag 24, so we can choose P = 1 or 2. We also can find PACF seems to be tailing off after lag 12, which indicates that P could be 0 as well. Within lag 12 in PACF, we can discover a peak at lag 1 and there is a tailing off within lag 12, we therefore can choose q = 0 or 1.



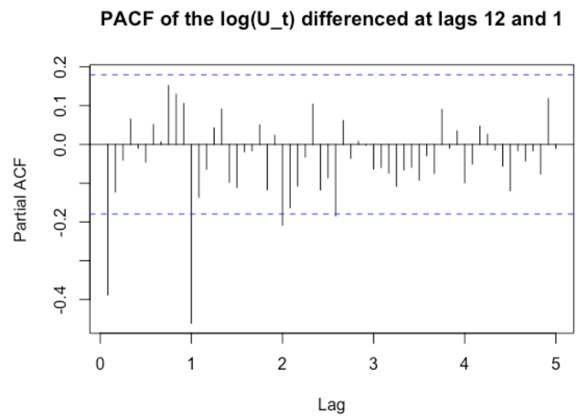


Figure 7: ACF plot and PACF plot of log transformed data after differencing at lag 12 and 1

		AICc
Model 1	SARIMA(0,1,1)(0,1,1)$_{12}$	-482.4616
Model 2	SARIMA(1,1,1)(0,1,1)$_{12}$	-480.7547
Model 3	SARIMA(0,1,1)(1,1,1)$_{12}$	-480.3672
Model 4	SARIMA(1,1,1)(1,1,1)$_{12}$	-478.6279
Model 5	SARIMA(0,1,1)(2,1,1)$_{12}$	-478.2262
Model 6	SARIMA(1,1,1)(2,1,1)$_{12}$	-476.3592

Table 2: Model candidates and their AICc value

2.3 Model Estimation

To choose an appropriate model, we could first compare the Akaike information criterion (AICc) of all possible models. *SARIMA*(0,1,1)(0,1,1)$_{12}$ has the lowest AICc value (-482.4616) and *SARIMA*(1,1,1)(0,1,1)$_{12}$ has the second lowest AICc value (-480.7547). After we adjust the model by 95% confidence interval test, we could find model *SARIMA*(0,1,1)(0,1,1)$_{12}$ still has the lowest AICc value and it is a parsimonious model. Therefore, we will choose model *SARIMA*(0,1,1)(0,1,1)$_{12}$ for further analysis.

			Coefficient		AICc
Model 1	SARIMA(0,1,1)(0,1,1)$_{12}$	ma1	sma1		-482.4616
		-0.4049	-0.8213		
		(0.0780)	(0.1183)		
Model 2	SARIMA(1,1,1)(0,1,1)$_{12}$	ar1	ma1	sma1	-480.7547
		-0.1207	-0.3109	-0.8205	
		(0.1880)	(0.1733)	(0.1191)	
Model 2	SARIMA(1,1,1)(0,1,1)$_{12}$	ar1	ma1	sma1	-464.4631
		0	0	-1.1470	
		0	0	(0.1996)	

Table 3: Models with highest and second highest AICc values and their coefficients

2.4 Diagnostic Checking

Let U_t be the industrial production of ice cream and frozen dessert in the U.S. at time t.

Denote X_t as $\nabla_1\nabla_{12}\ln(U_t)$.

$$\nabla_1\nabla_{12}\ln(U_t): X_t = (1-0.4049_{(0.0780)}B)(1-0.8213_{(0.1183)}B^{12})Z_t$$

$$\hat{\sigma}_Z^2 = 0.0008616$$

Initially, based on the model we choose, it is a MA model, which is always stationary. It is also invertible since the roots of Zt is 2.4697 and 1.2176 and both absolute values of roots are larger than 1, which indicates that roots are outside unit circle.

Secondly, we should check the model whether satisfy the assumption that the residual is a white noise. No trend, no seasonality, and no discernible variance change can be found in the residual plot, and its mean is pretty close to 0. It shows that the residual is stationary. Additionally, the histogram plot of residual seems to have a heavy tail, but the mean of the histogram is close to 0. The Q-Q plot of the residual seems a "S" shape, also indicating that the distribution of residual has a heavy tail. However, there is a contradiction that although the Q-Q plot seems to slightly violate the normality assumption of the residual, when we take the Shapiro-Wilk normality test, the p-value of it is greater than 0.05, which means we fail to reject the null hypothesis that the residual of the model has normality at significant level 0.05. Therefore, we still can use the 95% confidence interval to forecast. Our goal is to prove the residual is a white noise, therefore the contradiction of normality does not affect our model estimation.

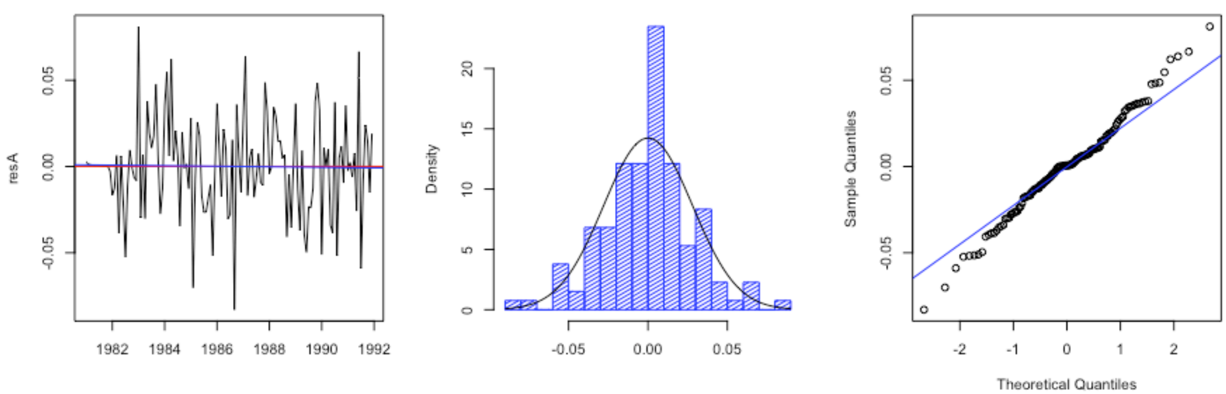


Figure 8: Time series plot *(left)*; histogram *(middle)*; Q-Q plot *(right)* of the residual of the model *SARIMA*$(0,1,1)(0,1,1)_{12}$

Moreover, to further check the white noise assumption, we applied Box-Pierce test and Ljung-Box test. Since $\sqrt{n}$ (root of the amount of observations: $\sqrt{132}$) is approximately equal to 11.49, we can set lag = 11 to perform model diagnostics. Also, we have two parameters in the model, so we set fit degree of freedom is 2 when we do the tests. The p-values of Box-Pierce test and Ljung-Box test are 0.5946 and 0.5425 respectively, and both are greater than 0.05, which means we fail to reject the null hypothesis that the residual is white noise at significance level 0.05.

We can also observe the ACF plot and PACF plot of the residual to check whether the residual is white noise. In the ACF plot, the value at lag 17 is outside the 95% confidence interval, however, since the value is just slightly beyond the interval, it can be considered as 0. Similarly, in PACF plot, there is a value at lag 9 which is just slightly outside the 95% confidence interval, so we can consider it as 0 as well. As for the outside value at lag 17 in PACF plot, firstly, it is kind of far from now, so the effect of it is not that significant. Secondly, as we check the white noise assumption by estimating AR model with the Yule-Walker method, the selected order for AR model is 0, which still can indicate that the residual is a white noise. Additionally, we also employed Ljung-Box test on the squared residuals to check there is no non-linear relationship in the residuals.

Test Type	p-value
Shapiro-Wilk Normality Test	0.3748 (greater than 0.05)
Box-Pierce Test	0.4992 (greater than 0.05)
Box-Ljung Test	0.4475 (greater than 0.05)
Box-Ljung Test (square)	0.9164 (greater than 0.05)
Yule-Walker	AR order selected 0; $\hat{\sigma^2} = 0.000785$

Table 4: Residual normality/ white noise/ non-linear checking

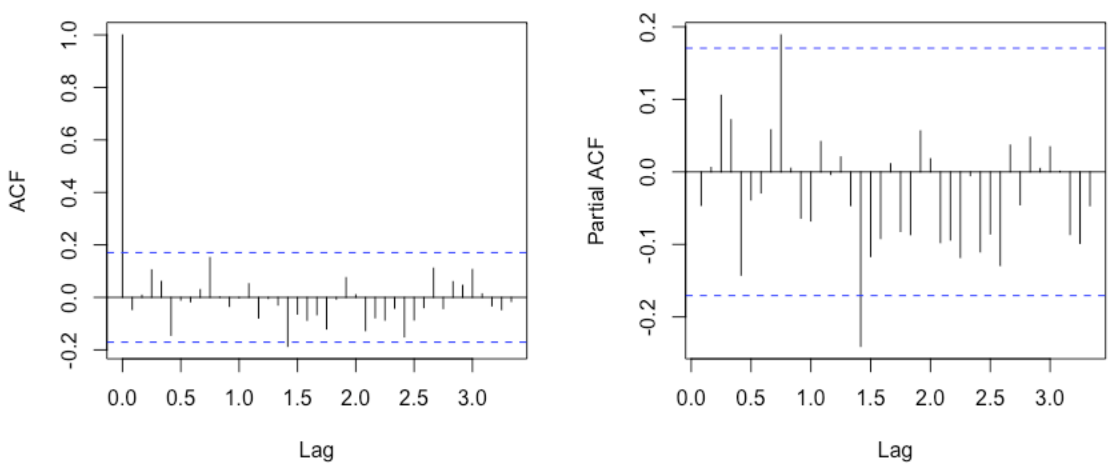


Figure 9: ACF and PACF of the model residual

2.5 Forecasting

We forecasted the transformed data and the original data by our time series model, showing the predicted value of ice cream and frozen dessert industrial production from 1992.01 to 1992.12

and 95% confidence intervals. As we can see in the figure 10 and 11, all prediction points are within the 95% confidence intervals.

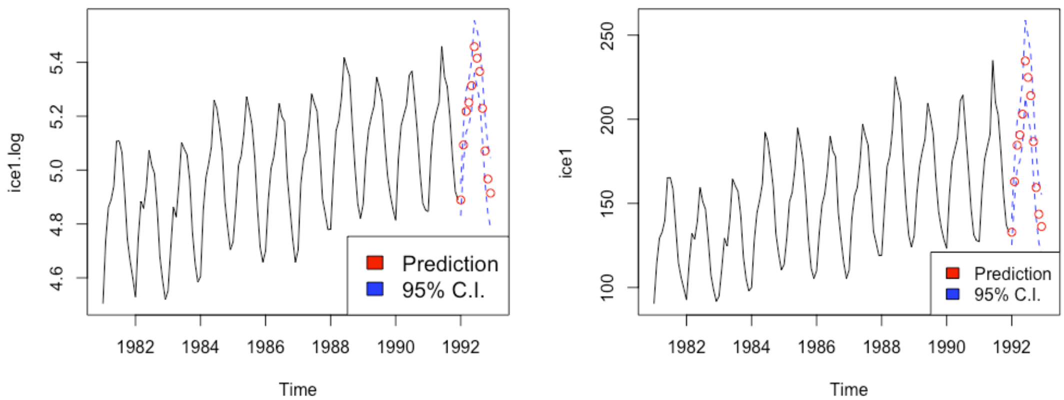


Figure 10: Forecast on the log transformed data *(left)* / original data *(right)*

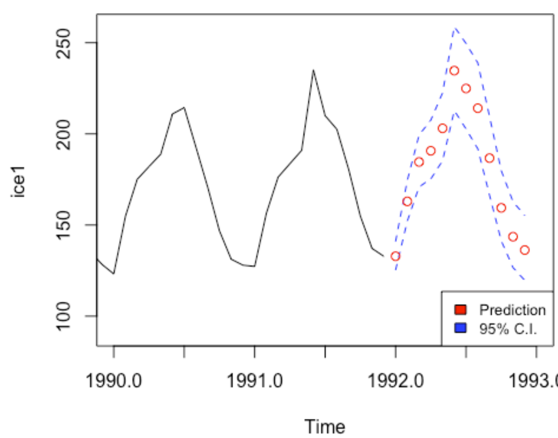


Figure 11: Zoomed in forecast on original data

Moreover, as shown in figure 11, the prediction values are closed to the actual values, illustrating the suitability and sufficiency of our time series model in the ice cream and frozen dessert industrial production case.

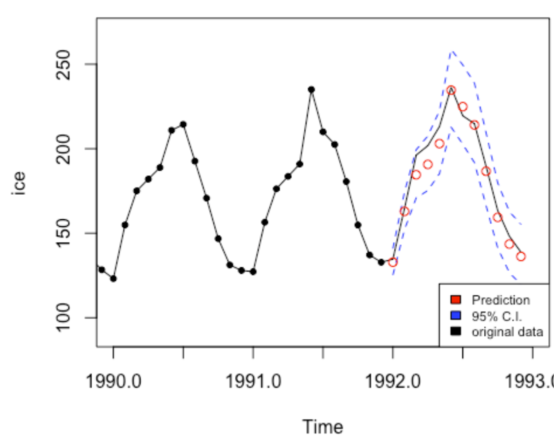


Figure 12: Zoomed in forecast on original data with testing part dataset

3. Conclusion

Based on the dataset presenting the industrial production of the ice cream and frozen dessert in the United State between 1981.01 and 1991.12, we successfully build a SARIMA model to forecast the industrial production of ice cream and frozen dessert between 1992.01 and 1992.12.

Let U_t be the industrial production of ice cream and frozen dessert in the U.S. at time t.

Denote X_t as $\nabla_1\nabla_{12}\ln(U_t)$.

$$\nabla_1\nabla_{12}\ln(U_t): X_t = (1-0.4049_{(0.0780)}B)(1-0.8213_{(0.1183)}B^{12})Z_t$$

$$\hat{\sigma}_Z^2 = 0.0008616$$

We initially applied Box-Cox transformation to get a more constant variance in the time series plot of original dataset. Based on the analysis of Box-Cox transformation, we decided to use log transformation. After that, we removed seasonal component and trend of the log transformed data by differencing at lag 12 and lag 1. We selected our final model by observing ACF and PACF plots of log transformed data after removing seasonality and trend. Then we diagnosed our model by checking Box-Pierce, Ljung-Box and Yule-Walker method to find out the residual of the model follows white noise assumptions and it is not non-linear. Besides, by running Shapiro-Wilk normality test, we failed to reject the null hypothesis that the residual of the model has normality at significant level 0.05. As for the forecasting, we got predicted values by our model prediction, and all predicted values are within the 95% confidence intervals. We can also find that the predicted values are close to the actual values.

Great thanks to Prof. Raya Feldman and my TA Lihao Xiao for their kind help and support for my final project. Also, thanks to myself for never giving up during this stressful quarter. Thanks to all people supporting me this quarter.

4. References

Brockwell, Peter J, and Richard A Davis. Introduction to Time Series and Forecasting. Cham Springer International Publishing, 2016.

Shumway, Robert H, and David S Stoffer. Time Series Analysis and Its Applications : With R Examples. Cham, Switzerland: Springer, 2017.

Feldman, Raya. “PSTAT 174: Lecture 15,” 2022.

5. Appendix

PSTAT 174 Code

Yujie Ye

Packages used

```
library(MASS)
library(forecast)
library(qpcR)
library(stats)
library(ggplot2)
```

```
#data
setwd("/Users/johnnys/Downloads/PSTAT 174")

frozen = read.table("IPN31152N.csv",
                    sep=",", header=FALSE, skip=1)

fro <- frozen[1:132,]
fro.test <- frozen[133:144,]

#plot original
ice = ts(frozen[,2], start = c(1981,1), end = c(1993,1), frequency = 12)
reg <- lm(V2 ~ as.numeric(1:132), data = fro)
plot(V2 ~ as.numeric(1:132), data = fro); abline(reg)
op = par(mfrow = c(1,2))
ice1 = ts(fro[,2], start = c(1981,1), end = c(1991,12), frequency = 12)
ice1_1 = ts(fro[,2], start = c(0), end = c(131), frequency = 1)
ts.plot(ice1_1, ylab = "index 2017 = 100"); abline(reg,col = "red")
ts.plot(ice1, ylab = "index 2017 = 100")
#histogram / ACF/PACF of original data

op = par(mfrow = c(1,3))
hist(ice1, main = "")
acf(ice1,lag.max = 60,main = "")
pacf(ice1,lag.max = 60,main = "")
```

```
#find lambda to transform (constant variables)
t = 1:length(ice1)
fit = lm(ice1 ~ t)
bcTransform = boxcox(ice1 ~ t,plotit = TRUE)

lambda = bcTransform$x[which(bcTransform$y == max(bcTransform$y))]
ice1.bc = (1/lambda)*(ice1^lambda-1)
ice1.log = log(ice1)
op= par(mfrow=c(2,2))
ts.plot(ice1, main = "Original Times Series")
```

```r
ts.plot(ice1_1, main = "Original Times Series"); abline(reg,col = "red")
ts.plot(ice1.bc, main = "Box-Cox Transform")
ts.plot(ice1.log, main = "Log Transform")

#Compare before and after log
par(mfrow=c(3,3))
plot.ts(ice1,xlab = "", main = "")
hist(ice1, xlab = "", main="Original data")
qqnorm(ice1, main = "", xlab = "")
qqline(ice1, col = "red")

plot.ts(ice1.log,xlab = "", main = "")
hist(ice1.log, xlab = "", main = "log transformed data")
qqnorm(ice1.log, main = "", xlab = "")
qqline(ice1.log, col = "red")

plot.ts(ice1.bc,xlab = "", main = "")
hist(ice1.bc, xlab = "", main = "Box-Cox transformed data")
qqnorm(ice1.bc, main = "", xlab = "")
qqline(ice1.bc, col = "red")
```

```r
# de-seasonal and de-trend
y2 = diff(ice1.log, 12)
y21 = diff(y2, 1)
y211 = diff(y21, 1)
ts.plot(y2, main = "De-seasonalized Time Series",
        ylab = expression(nabla^{12}~Y[t]))
abline(h = 0,lty = 2)

ts.plot(y21,main= "De-trended/seasonalized Time Series",
        ylab = expression(nabla^{12}~nabla~Y[t]))
abline(h = 0,lty = 2)

c(Ut = var(ice1.log),delta_12 = var(y2),delta_1_12 = var(y21)
  ,deta_1_1_12 = var(y211))
#compare acf and pacf
op= par(mfrow=c(2,3))
acf(ice1.log,lag.max = 60,main = expression(ln(U[t])))
acf(y2,lag.max = 60,main = expression(nabla[12]~~ln(U[t])))
acf(y21,lag.max = 60,main = expression(nabla[12]~nabla[1]~~ln(U[t])))
pacf(ice1.log,lag.max = 60,main = "")
pacf(y2,lag.max = 60,main = "")
pacf(y21,lag.max = 60,main = "")

#ACF and PACF of de-trend&season
decomp <- decompose(ts(as.ts(ice1.bc),
                       start=c(1981,1), end=c(1991,12), frequency = 12))
plot(decomp)
op= par(mfrow=c(1,2))
acf(y21,lag.max = 60,main = "ACF of the log(U_t) differenced at lags 12 and 1")
pacf(y21,lag.max = 60,main = "PACF of the log(U_t) differenced at lags 12 and 1")
```

```r
# D = 1; s = 12; d = 1; Q = 1; P = 0 or 1 or 2; q = 1; p = 1 or 0(tail off)
fit_1 <- arima(ice1.log, order = c(0,1,1), seasonal = list(order = c(0,1,1))
               ,method = "ML")
fit_2 <- arima(ice1.log, order = c(1,1,1), seasonal = list(order = c(0,1,1))
               ,method = "ML")
fit_3 <- arima(ice1.log, order = c(0,1,1), seasonal = list(order = c(1,1,1))
               ,method = "ML")
fit_4 <- arima(ice1.log, order = c(1,1,1), seasonal = list(order = c(1,1,1))
               ,method = "ML")
fit_5 <- arima(ice1.log, order = c(0,1,1), seasonal = list(order = c(2,1,1))
               ,method = "ML")
fit_6 <- arima(ice1.log, order = c(1,1,1), seasonal = list(order = c(2,1,1))
               ,method = "ML")

fit_1;fit_2;fit_3;fit_4;fit_5;fit_6
c(fit_1 = AICc(fit_1),fit_2 = AICc(fit_2),fit_3 = AICc(fit_3),
  fit_4 = AICc(fit_4),fit_5 = AICc(fit_5),fit_6 = AICc(fit_6))
fit_1
fit_2

fit_2_2 <-arima(ice1.log, order = c(1,1,1), seasonal = list(order = c(0,1,1)),
                fixed = c(0,0,NA)
               ,method = "ML")

c(fit_1 = AICc(fit_1),fit_2_2 = AICc(fit_2_2))
```

$$\nabla_1\nabla_{12}ln(U_t) : X_t = (1 - 0.4049_{0.0780}B)(1 - 0.8213_{(0.1183)}B^{12})Z_t\hat{\sigma_Z}^2 = 0.0008616$$

```r
#check Invertible
polyroot(c(1,-0.4049))
polyroot(c(1,-0.8213))
```

MA: stationary

```r
# Diagnose
modelA <- arima(ice1.log, order = c(0,1,1), seasonal = list(order = c(0,1,1))
               ,method = "ML")

op= par(mfrow=c(1,3))
resA <- residuals(modelA)
plot.ts(resA, xlab = ""); abline(h=mean(resA),col = "red")
fitt <- lm(resA ~ seq(1981,1992,1/12)[1:length(resA)]); abline(fitt,col="blue")

hist(resA, density = 20, breaks = 20, col = "blue",main ="", xlab = "", prob = TRUE);

qqnorm(resA, main = ""); qqline(resA, col="blue")
######
op= par(mfrow=c(1,2))
acf(resA, lag.max = 40)
pacf(resA, lag.max = 40)

######
```

```r
#Tests
shapiro.test(resA)
Box.test(resA, lag = 11, type = c("Box-Pierce"), fitdf = 2)
Box.test(resA, lag = 11, type = c("Ljung-Box"), fitdf = 2)
ar(resA, aic = TRUE, order.max = NULL, method = c("yule-walker"))
Box.test((resA)^2, lag = 11, type = c("Ljung-Box"), fitdf = 0)
```

Forecast

```r
op= par(mfrow=c(1,2))
fit_1_1 <- arima(ice1.log, order = c(0,1,1), seasonal = list(order = c(0,1,1),period=12)
                 ,method = "ML")
# Create confidence interval
pred.tr <- predict(fit_1_1, n.ahead = 12)
U.tr = pred.tr$pred + 2*pred.tr$se
L.tr = pred.tr$pred - 2*pred.tr$se
# Forecast on original data
ts.plot(ice1.log, xlim = c(1981,1993), xlab="",
        ylim = c(min(ice1.log), max(U.tr)))
lines(U.tr, col = "blue", lty = "dashed")
lines(L.tr, col = "blue", lty = "dashed")
points(pred.tr$pred,col = "red", pch = 1)
legend("bottomright", c("Prediction","95% C.I."),
       fill = c("red", "blue"), cex = 1.25)

pred.orig <- exp(pred.tr$pred)
U = exp(U.tr)
L = exp(L.tr)
ts.plot(ice1, xlim= c(1981,1993), ylim= c(min(ice1), max(U)))
lines(U, col="blue", lty="dashed")
lines(L, col="blue", lty="dashed")
points(pred.orig,col = "red", pch = 1)
legend("bottomright", c("Prediction","95% C.I."),
fill = c("red", "blue"), cex = 1)
```

Forecast with test data

```r
pred.orig <- exp(pred.tr$pred)
ts.plot(ice, xlim= c(1981,1993),ylim=c(80,300))
lines(U, col="blue", lty="dashed")
lines(L, col="blue", lty="dashed")
points(pred.orig,col = "red", pch = 1)
points(ice1,col = "black", pch = 20)
legend("bottomright", c("Prediction","95% C.I.","original data"),
       fill = c("red", "blue","black"), cex = 0.7)
```

Zoom in: Forecast on original data and Forecast on original data with actual values

```r
op= par(mfrow=c(1,2))
pred.orig <- exp(pred.tr$pred)
U = exp(U.tr)
L = exp(L.tr)
```

```
ts.plot(ice1, xlim= c(1990,1993), ylim= c(min(ice1), max(U)))
lines(U, col="blue", lty="dashed")
lines(L, col="blue", lty="dashed")
points(pred.orig,col = "red", pch = 1)
legend("bottomright", c("Prediction","95% C.I."),
fill = c("red", "blue"), cex = 0.8)

pred.orig <- exp(pred.tr$pred)
ts.plot(ice, xlim= c(1990,1993),ylim=c(90,270))
lines(U, col="blue", lty="dashed")
lines(L, col="blue", lty="dashed")
points(pred.orig,col = "red", pch = 1)
points(ice1,col = "black", pch = 20)
legend("bottomright", c("Prediction","95% C.I.","original data"),
       fill = c("red", "blue","black"), cex = 0.7)
```