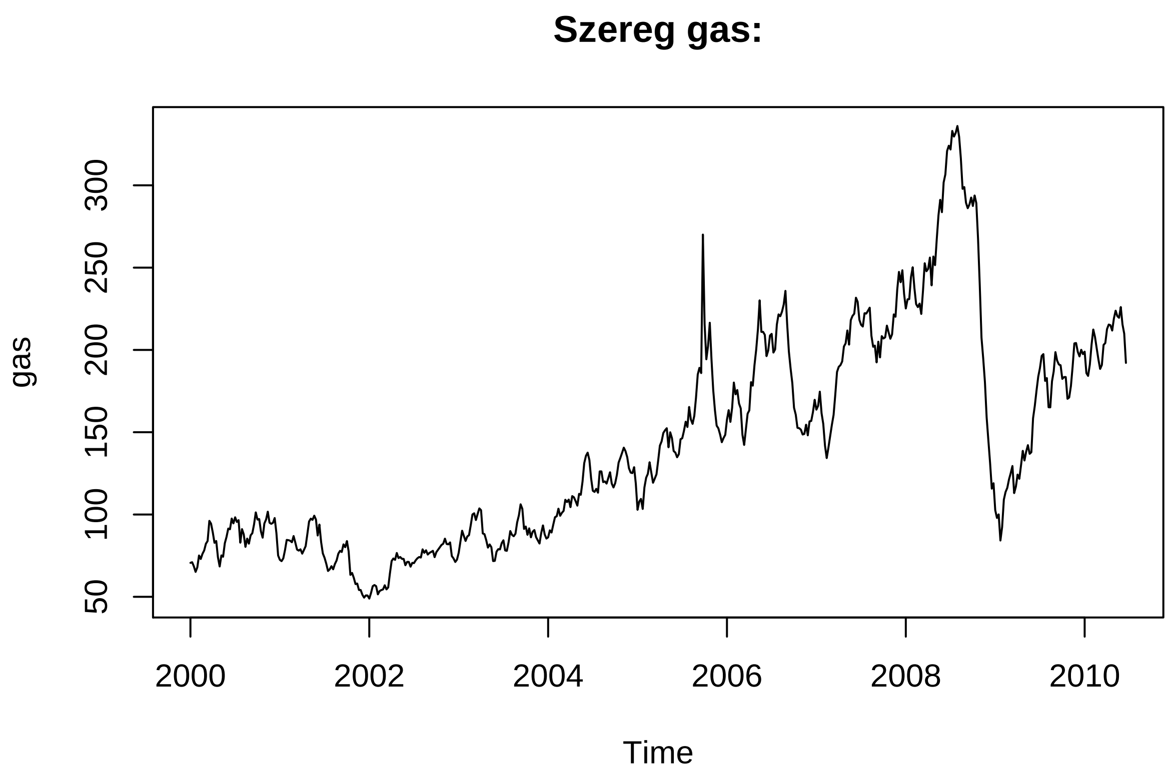
Projekt - Analiza Szeregów Czasowych

Dagmara Sikora i Magdalena Olbryś

1. Nasze dane:

Pracujemy na danych 'gas' z biblioteki 'astsa', opisujących cotygodniowe zmiany cen gazu w latach 2000-2010 w Nowym Jorku. Wyświetlmy kilka początkowych wyrazów oraz wykres.

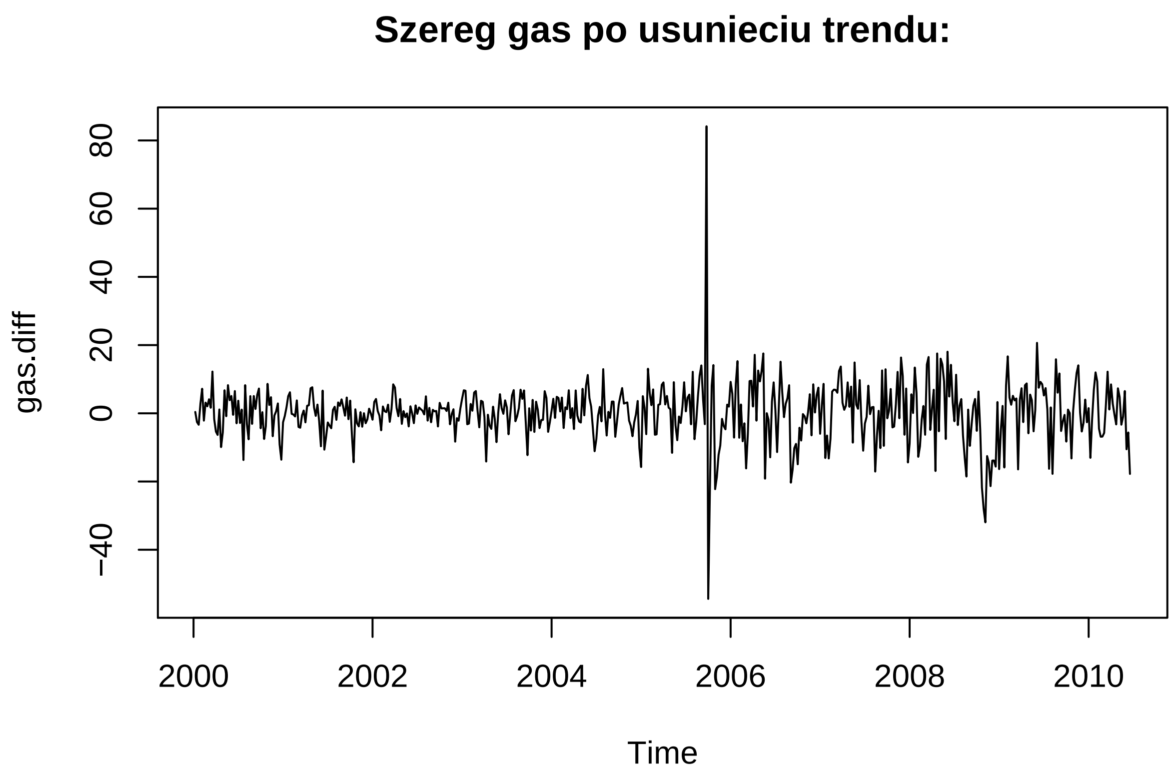
```
## [1] 70.636 71.040 68.490 65.137 67.918 75.117
```



2. Stacjonarność szeregu:

Na powyższym wykresie nie widać sezonowośći, natomiast możemy zaobserwować duży trend, który eliminujemy metodą różnic skończonych:

```r
gas.diff <- diff(gas)
```



Sezonowość prawdopodobnie nie występuje. Sprawdźmy więc za pomocą testu Dickey-Fuller'a czy nasz szereg jest już stacjonarny.

```r
adf.test(diff(gas))
```

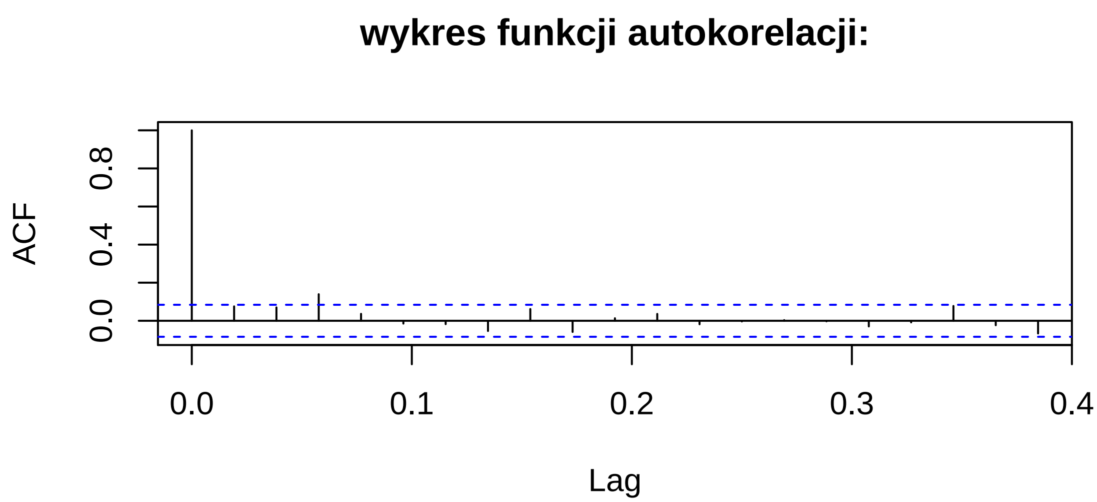
```
## Warning in adf.test(diff(gas)): p-value smaller than printed p-value

##
##  Augmented Dickey-Fuller Test
##
## data:  diff(gas)
## Dickey-Fuller = -7.4801, Lag order = 8, p-value = 0.01
## alternative hypothesis: stationary
```

wartość p jest mała, więc nasz szereg jest stacjonarny już po jednokrotnym zróżnicowaniu.

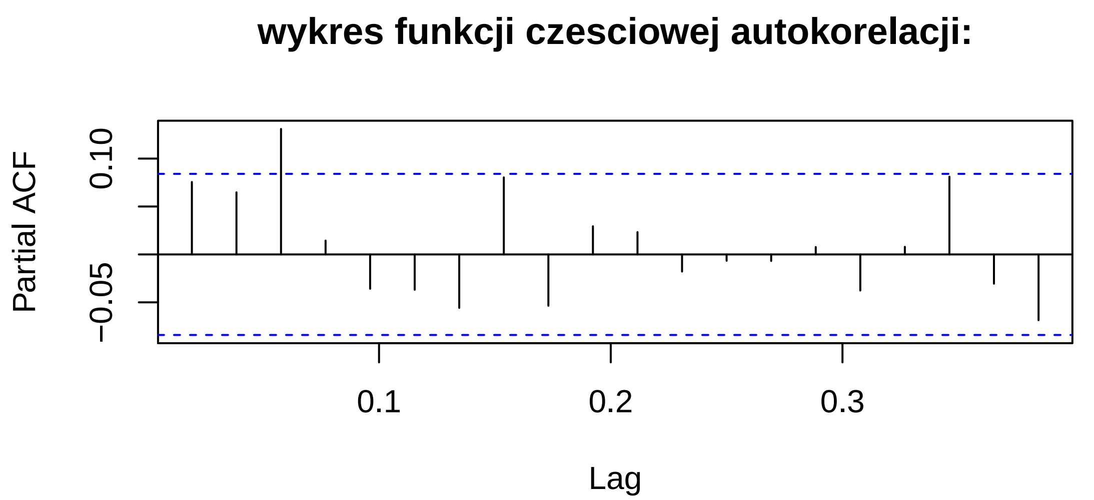
3. Dopasowywanie modelu:

Spróbujmy dopasować model autoregresji(AR(p)) lub średniej ruchomej(MA(q)). W przypadku procesu AR spodziewamy się, że wykres ACF będzie się stopniowo zmniejszał, a jednocześnie PACF powinien mieć gwałtowny spadek. W przypadku MA spodziewamy się czegoś przeciwnego, co oznacza, że: ACF powinien wykazywać ostry spadek, podczas gdy PACF powinien wykazywać geometryczny lub stopniowy trend spadkowy. Z drugiej strony, jeśli zarówno wykresy ACF, jak i PACF wykazują stopniową tendencję malejącą, to do modelowania należy rozważyć proces ARiMR. Potrzebujemy wykresów funkcji autokorelacji(ACF) i częściowej autokorelacji(PACF):

```r
acf(diff(gas), plot = TRUE, lag.max = 20, main = "wykres funkcji autokorelacji:") -> acf
```




```r
pacf(diff(gas), plot = TRUE, lag.max = 20, main = "wykres funkcji czesciowej autokorelacji:") -> pacf
```



Z naszych wykresóW ACF i PACF wynika, że szereg ma zbliżone do zera wartości funkcji autokorelacji, więc wygląda jak biały szum (ciąg niezależnych zmiennych losowych o jednakowych rozkładach prawdopodobieństwa ze skończonymi wartościami średnimi i wariancjami). Sprawdźmy testem Ljung- Boxa:

```r
Box.test(diff(gas), lag = 20, type = "Ljung-Box")$p.value
```

```
## [1] 0.05570119
```

Zbyt mała wartość aby uznać, że nasz szereg jest białym szumem.

Spróbujemy więc dopasować model AR(3).

4. Szacowanie współczynników dla AR(3):

```
d<- diff(gas)

#szacujemy alfa1 Y-W:
wsp.yw <- ar.yw(d, order.max = 3, aic = F)$ar

#teraz regresją:

f <- function(x) {
n <- length(x)
t <- 4:n
b <- coef(lm(x[t] ~ x[t-1] + x[t-2] + x[t-3]))
b[2:length(b)]
}
wsp.regr <- f(d)
```

Wartości współczynnika alpha1 (kolejność: yule-walker, regresja):

```
print(wsp.yw)
```

```
## [1] 0.06221177 0.05552702 0.13092558
```

```
print(wsp.regr)
```

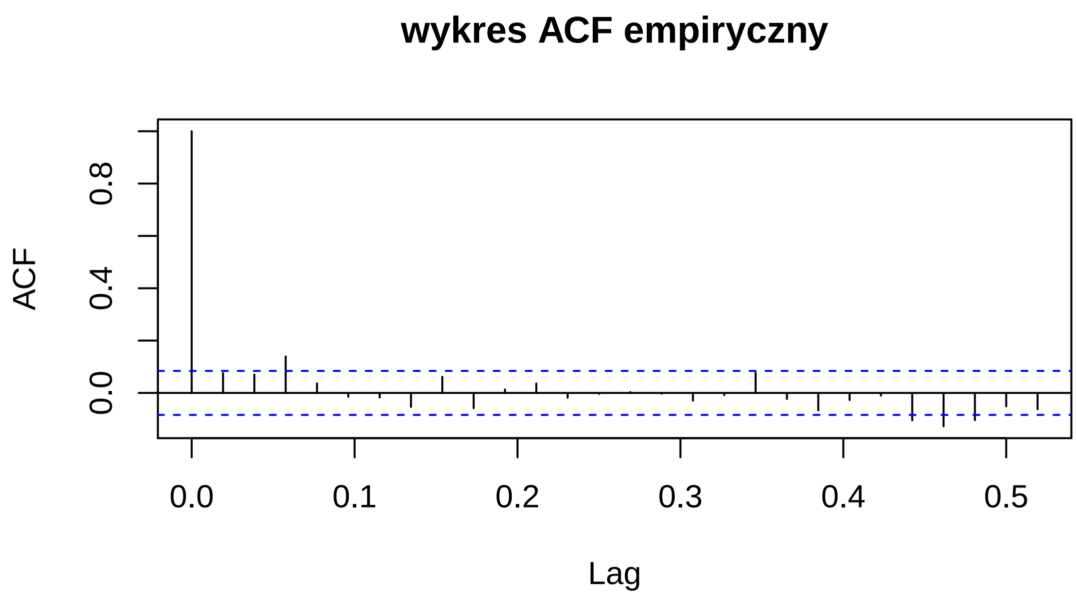
```
##   x[t - 1]   x[t - 2]   x[t - 3]
## 0.06305769 0.05653861 0.13286721
```

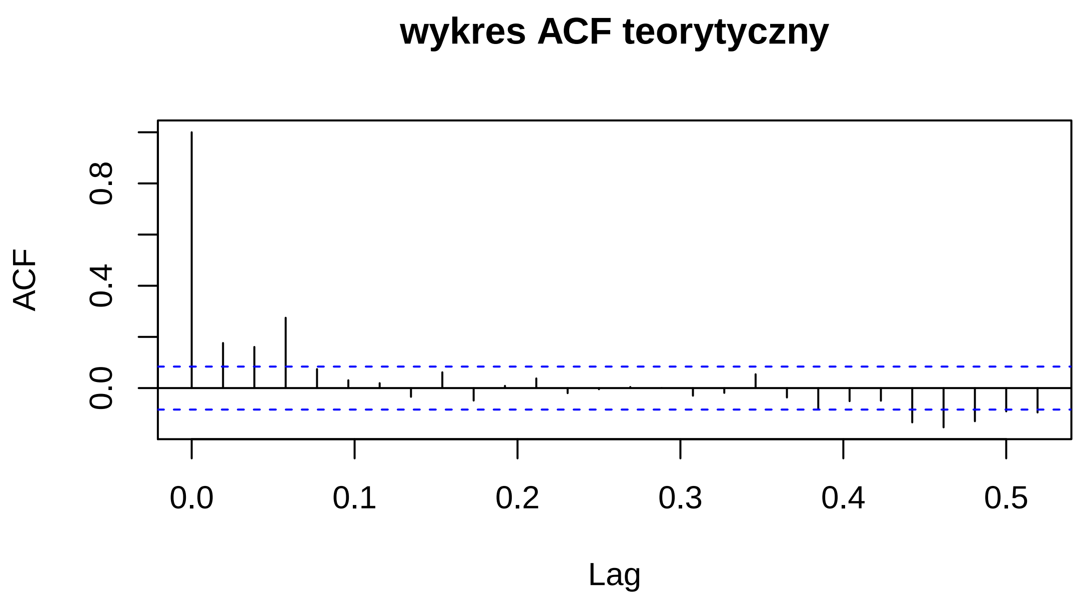
Wartości wychodzą w przybliżeniu takie same. Do dalszej pracy alpha1 obliczone przy pomocy równań Yule-Walkera.

5. Porównanie ACF i PACF z teorytycznymi:

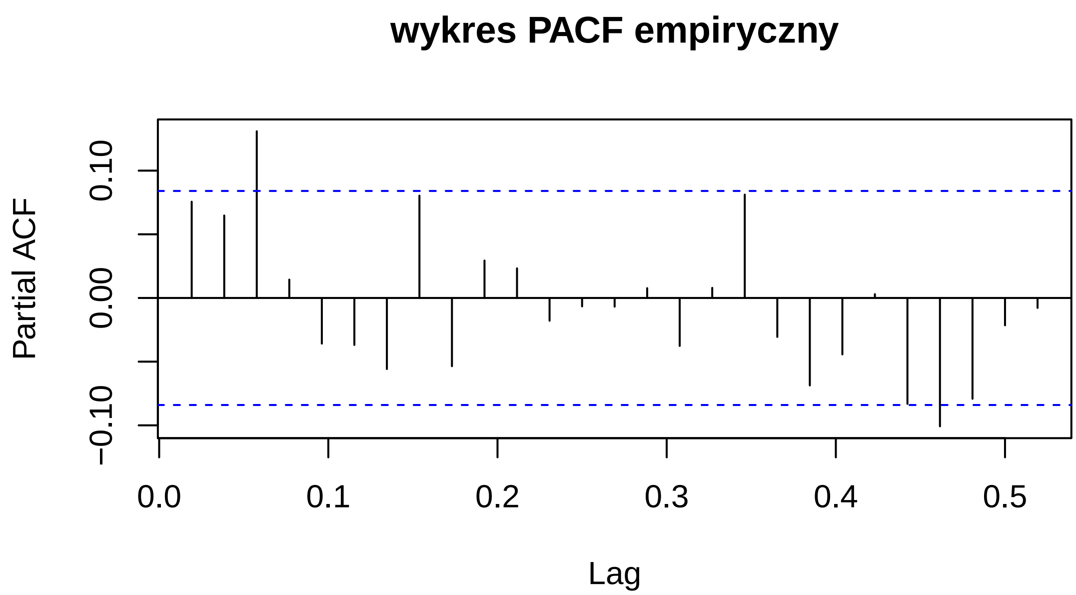
```
sz.gas <- filter(d, filter = c(wsp.yw), method = "r")
```

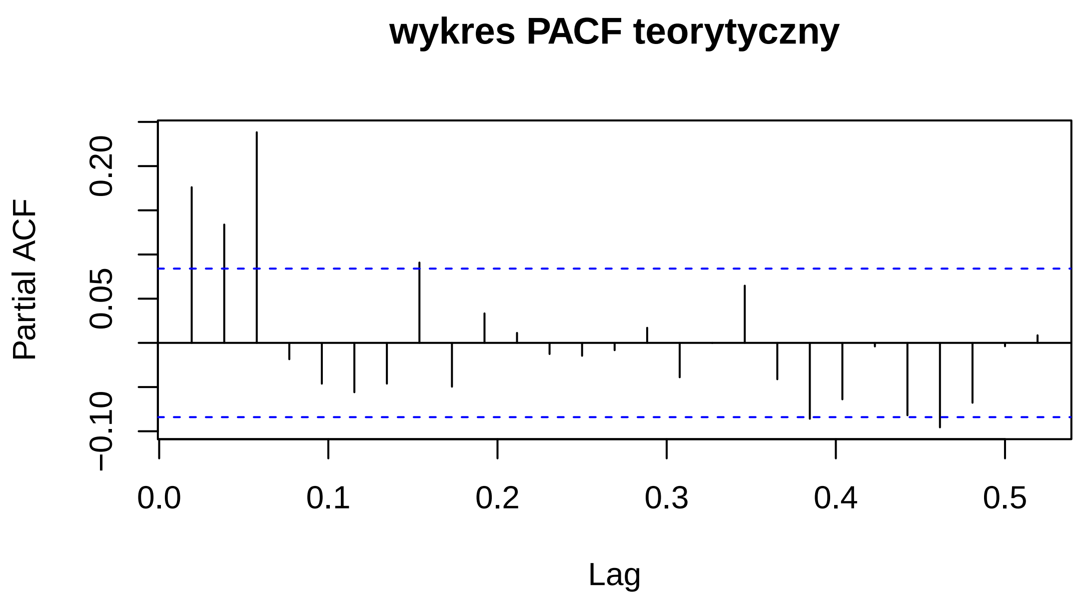
```
par(mfrow = c(2,1))
acf(diff(gas), plot = T,  main = "wykres ACF empiryczny")
acf(sz.gas, plot = T, main = "wykres ACF teorytyczny")
```






```
par(mfrow = c(2,1))
pacf(diff(gas), plot = T,  main = "wykres PACF empiryczny")
pacf(sz.gas, plot = T, main = "wykres PACF teorytyczny")
```

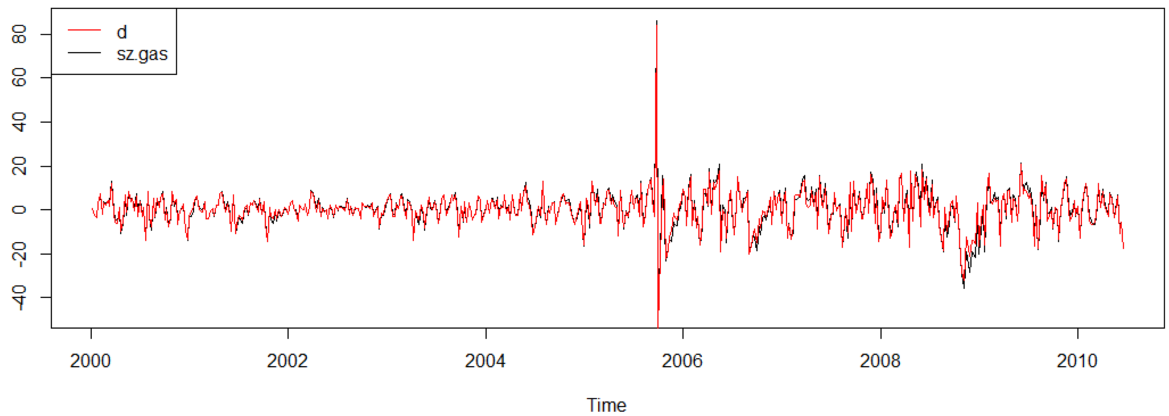




Nasze wykresy funkcji autokorelacji i częściowej autokorelacji są całkiem podobne.

Przypomnijmy, że szereg d to nasz początkowy szereg gas, ale z wyeliminowanym trendem, natomiast sz.gas to szereg zamodelowany za pomocą modelu AR(3).

Zobaczmy te 2 szeregi na jednym wykresie:



6. Reszty:

Spróbujemy dopasować do naszego zróżnicowanego szeregu model ARIMA(3, 0, 0) i jeśli reszty okażą się być białym szumem to zaprognozować zachowanie szeregu w przyszłości.

```
A <- arima(d, order = c(3,0,0))
reszty <- A$residuals

Box.test(reszty, lag = 20, type = "Ljung-Box")$p.value
```

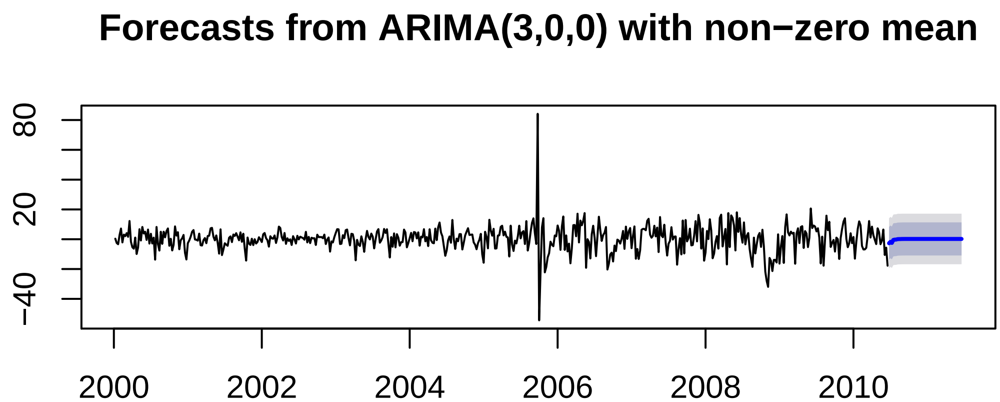
```
## [1] 0.6607131
```

Wartość p w teście Ljung-Box jest duża więc nie mamy podstaw do odrzucenia hipotezy, że nasze reszty są białym szumem.

7. Predykcja:

Spróbujmy przewidzieć ceny paliwa na kolejny rok.

```
A <- arima(d, order = c(3,0,0))
fc <- forecast(A, h= 52)
plot(fc)
```



Całkiem dziwnie to wygląda, więc chyba model którego użyłyśmy nie był do końca dobry.

8. Wnioski i obserwacje:

Nasz szereg po usunięciu trendu wygląda jak biały szum ale test Ljung-Box wskazuje na to że raczej nim nie jest. Według wykresów ACF i PACF mogłyśmy spróbować dopasować model AR(3), ale widocznie nie był to dobry pomysł, bo predykcja nie wygląda najlepiej.