Time Series Analysis on U.S. Suicide Rate 1920-2015

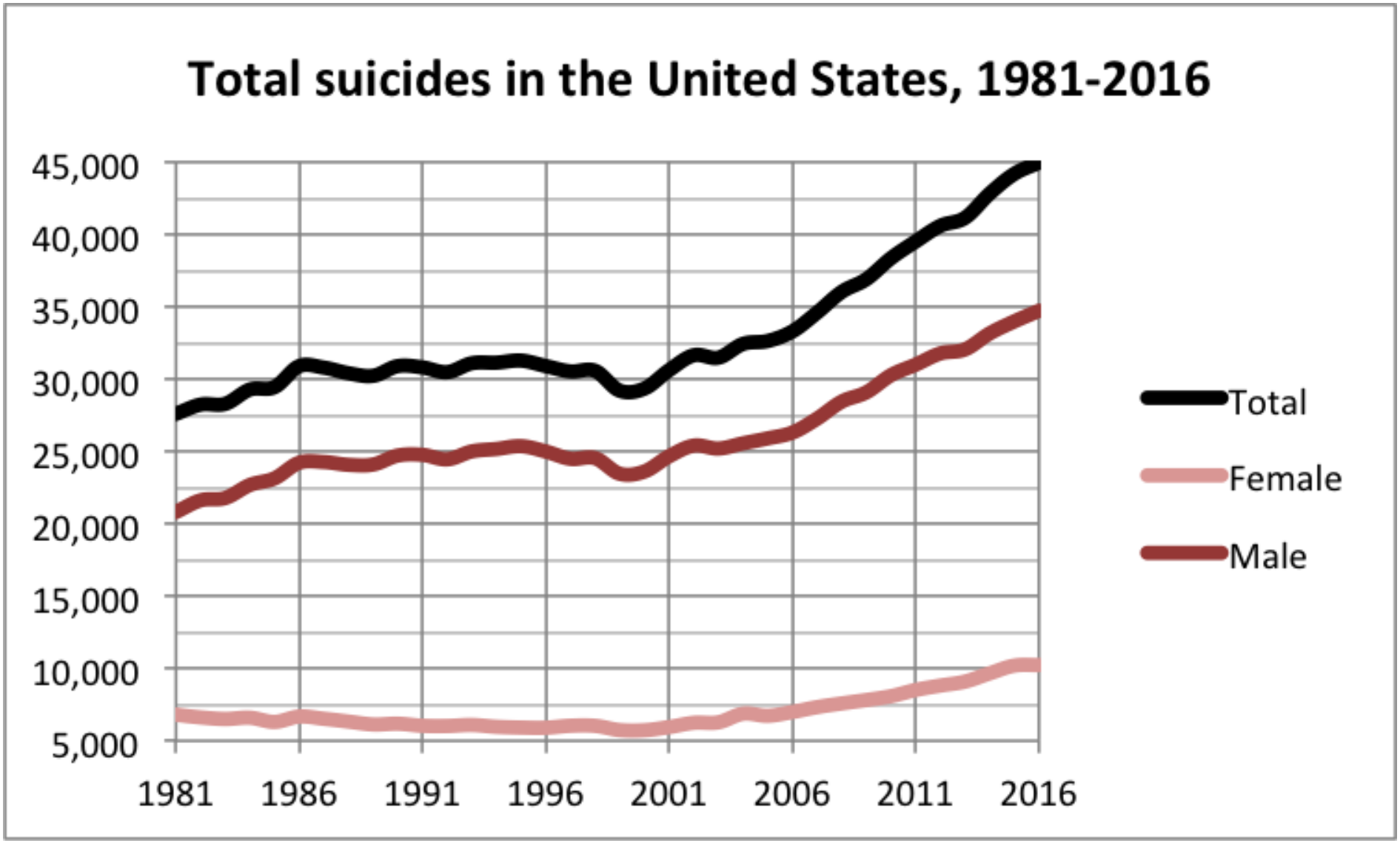
Xiran Wang

4/15/2019

Presentation Structure

- 1.Motivation: Suicide Issues in the U.S.
- 2.Data set: Annual U.S. Suicide Rate 1920-2015
- 3.Models: Candidate Models and Model Diagnostics
- 4.Forcast: 1-step Ahead Predictions
- 5.Forcast: 10-steps Ahead Predictions for 2016-2025

Suicide Issues in the U.S.



Suicide Issues in the U.S.

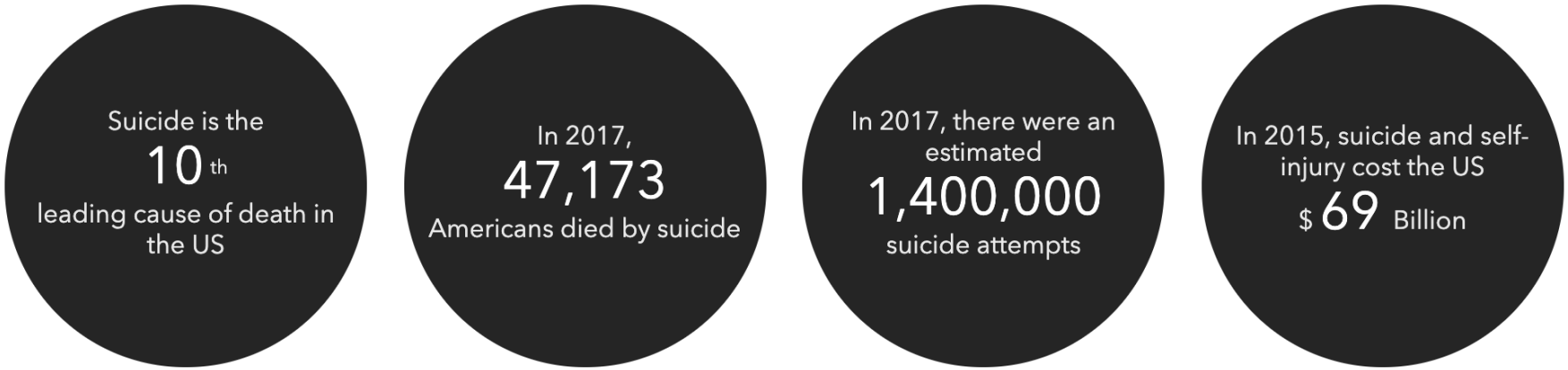
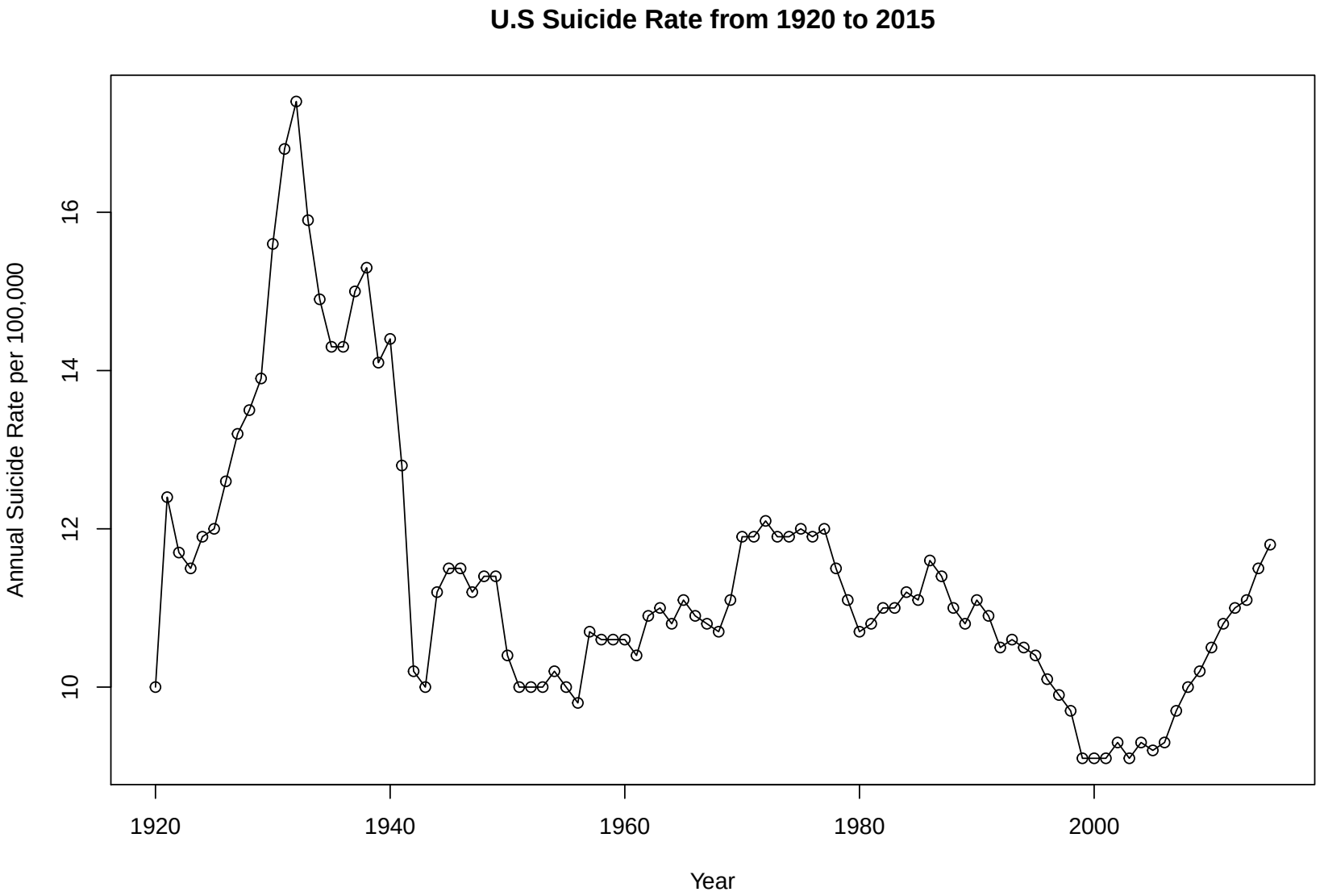
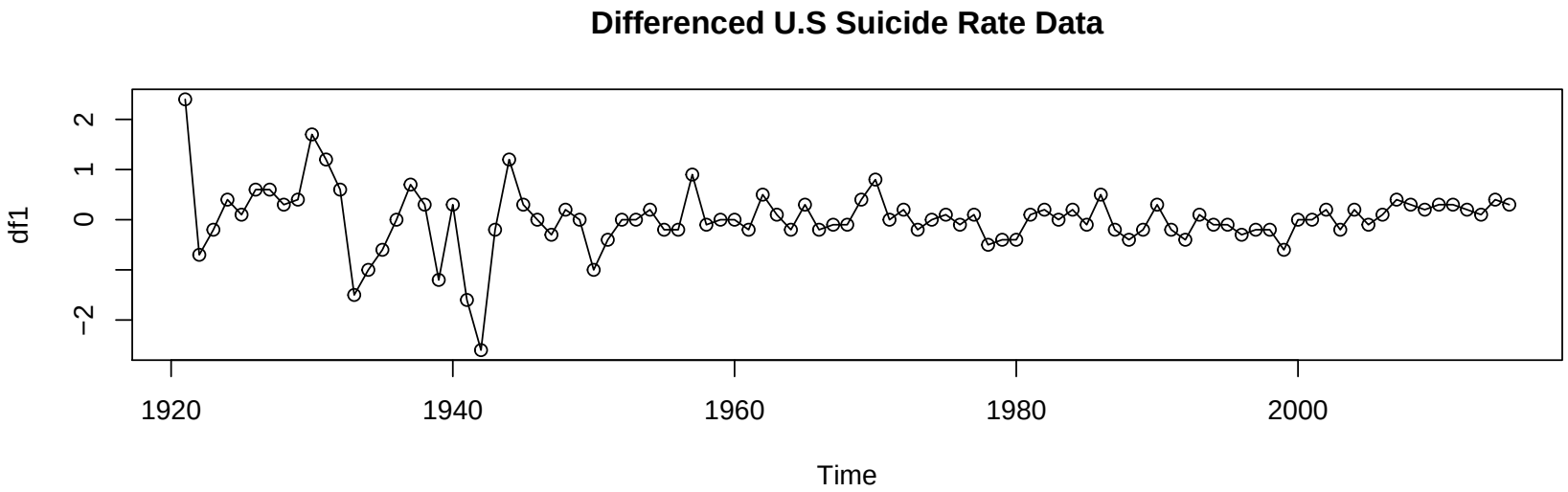


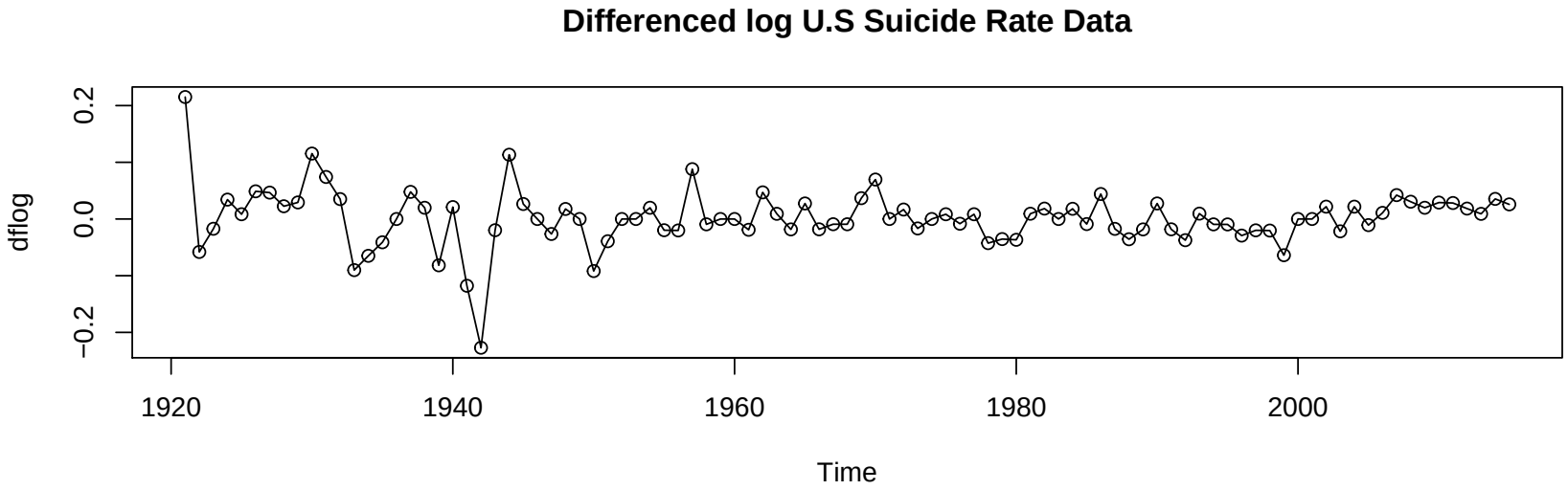
Figure 2: Image from American Fundation for Suicide Prevention

The data: Annual U.S. Suicide Rate 1920-2015

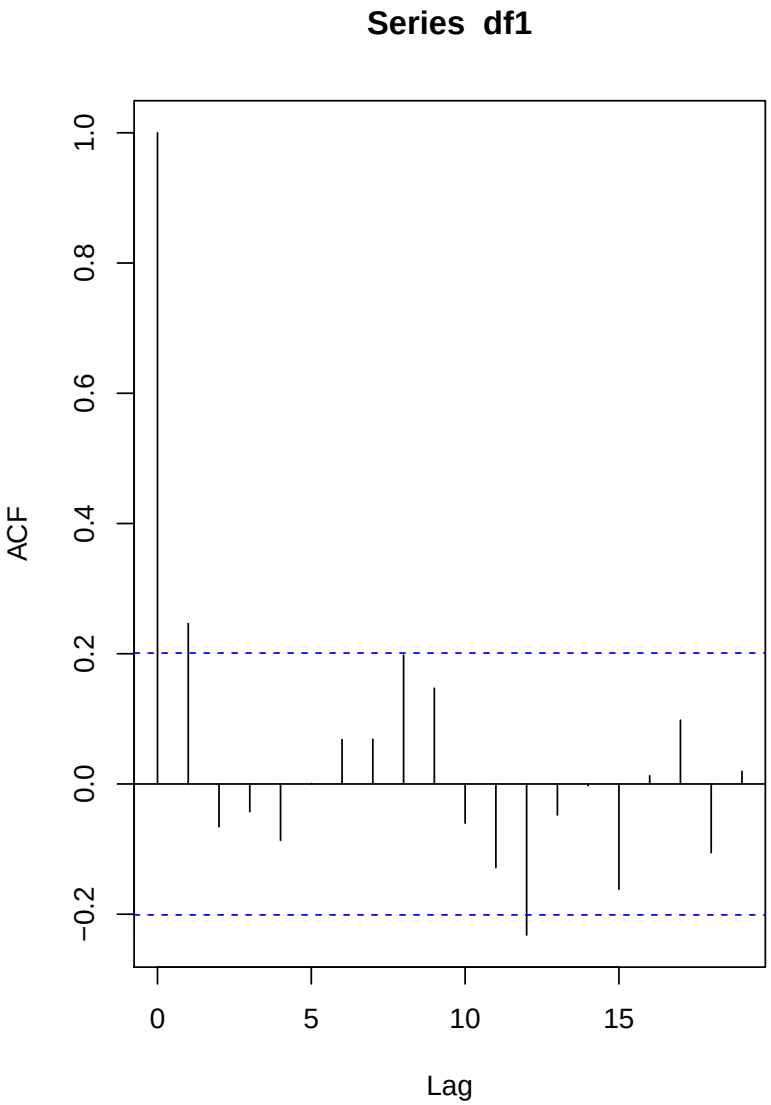


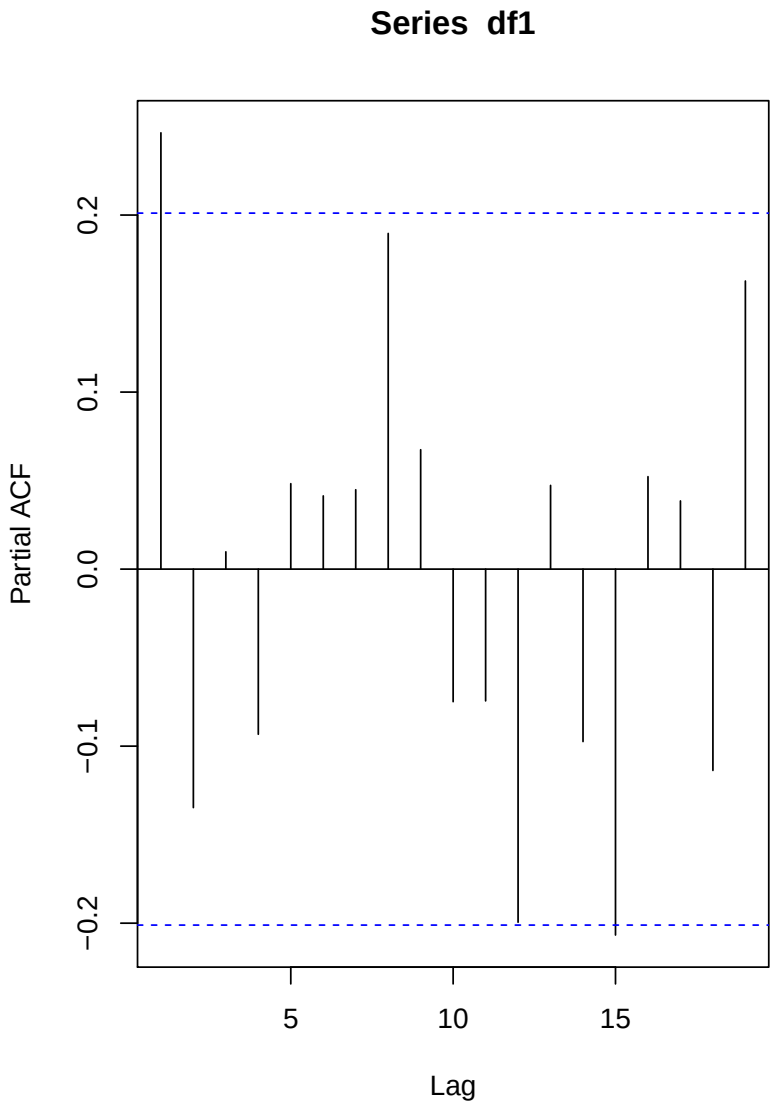
Model Specification





Model Specification





Model Specification

Overall, arima(0,1,1) has the best fit based on AIC and BIC.

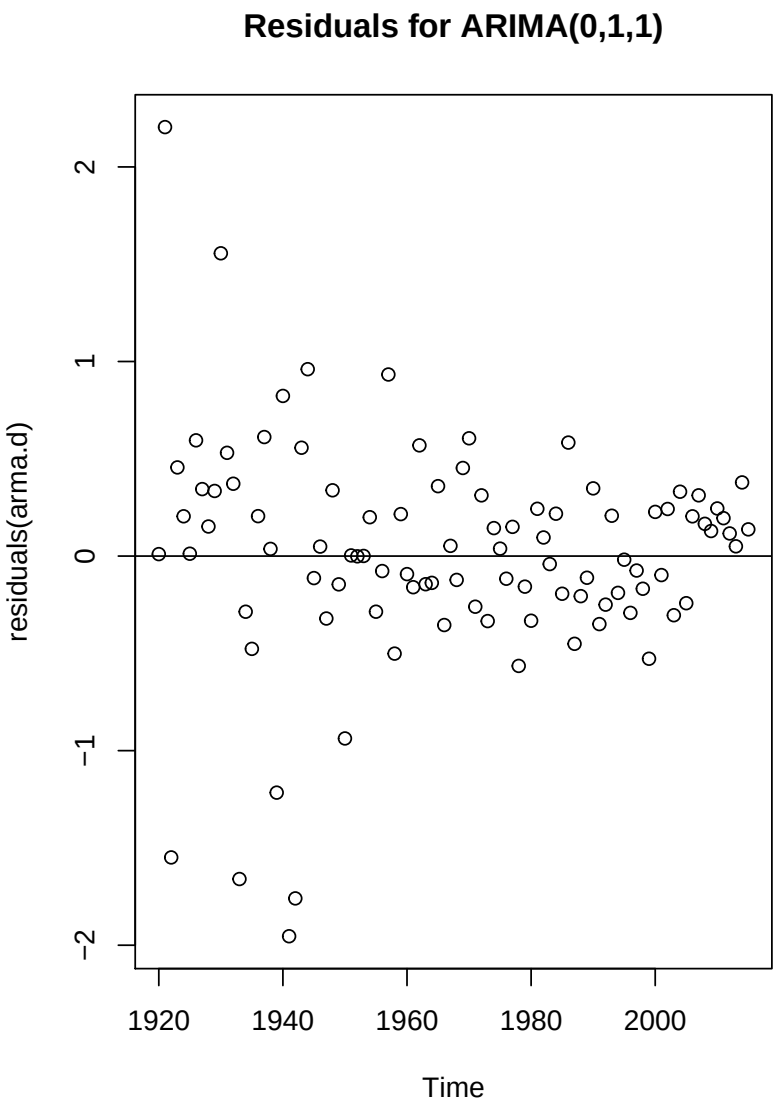
```
##                ar1 coef    ma1 coef  AIC       BIC
## ARIMA(1,1,0) 0.2908627   NA         172.4452 177.553
## ARIMA(0,1,1) NA          0.4301368 169.1625 174.2703
## ARIMA(1,1,1) -0.2022397 0.5912562 170.4317 178.0933
```

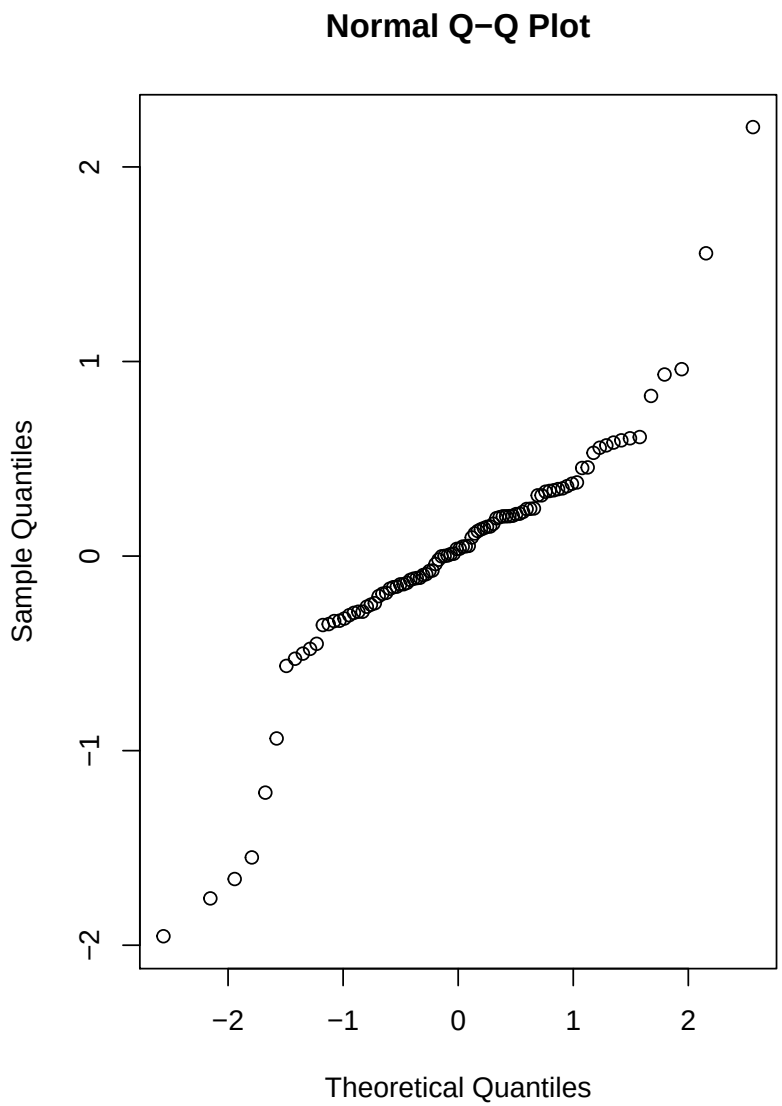
Model diagnostics

```
arma.d<-arima(sr, order=c(0,1,1),method='ML')
# ARIMA(0,1,1) using maximum likelihood

Box.test(arma.d$resid,type="Ljung-Box")
```

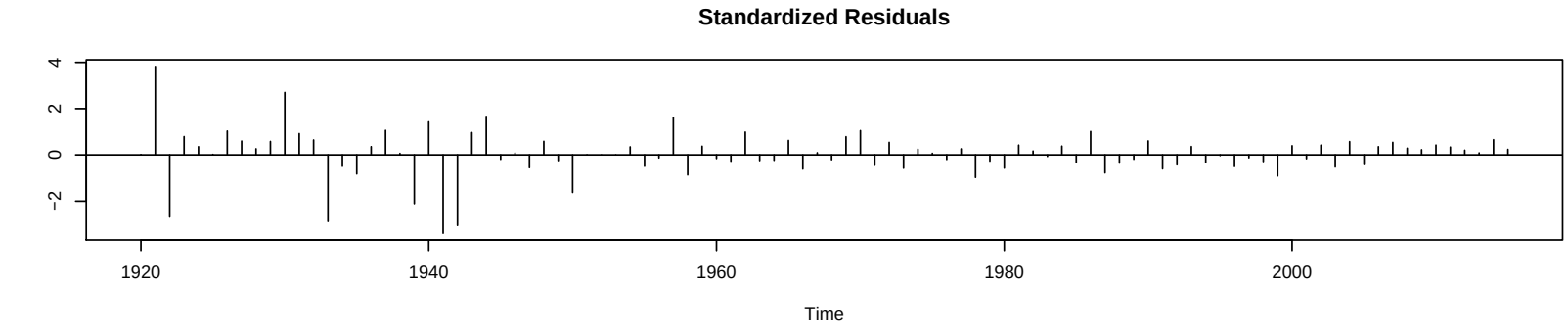
```
##
##  Box-Ljung test
##
## data:  arma.d$resid
## X-squared = 1.3346, df = 1, p-value = 0.248
```

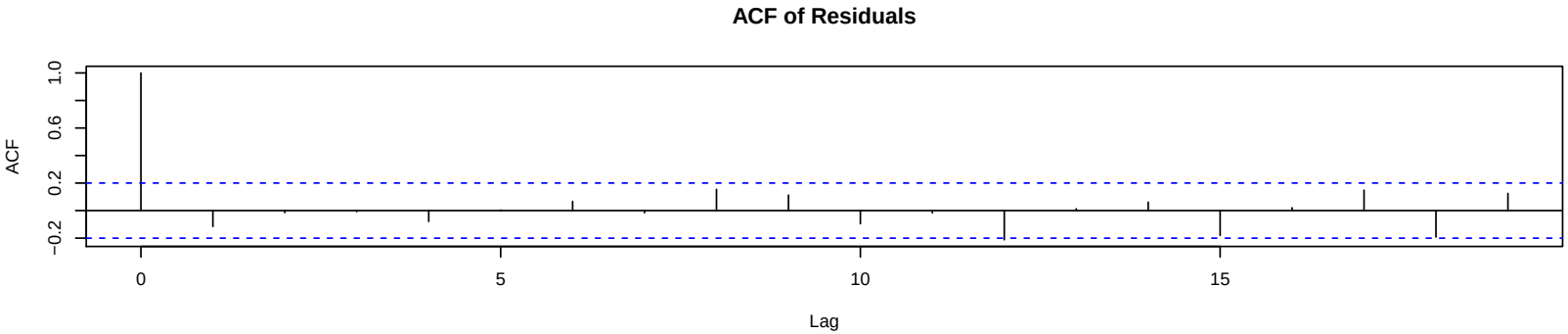
Model diagnostics

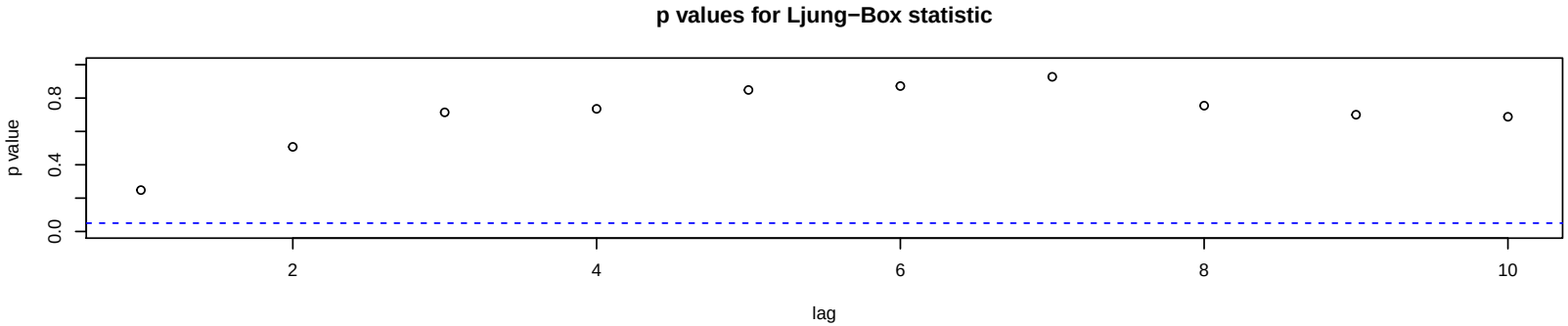




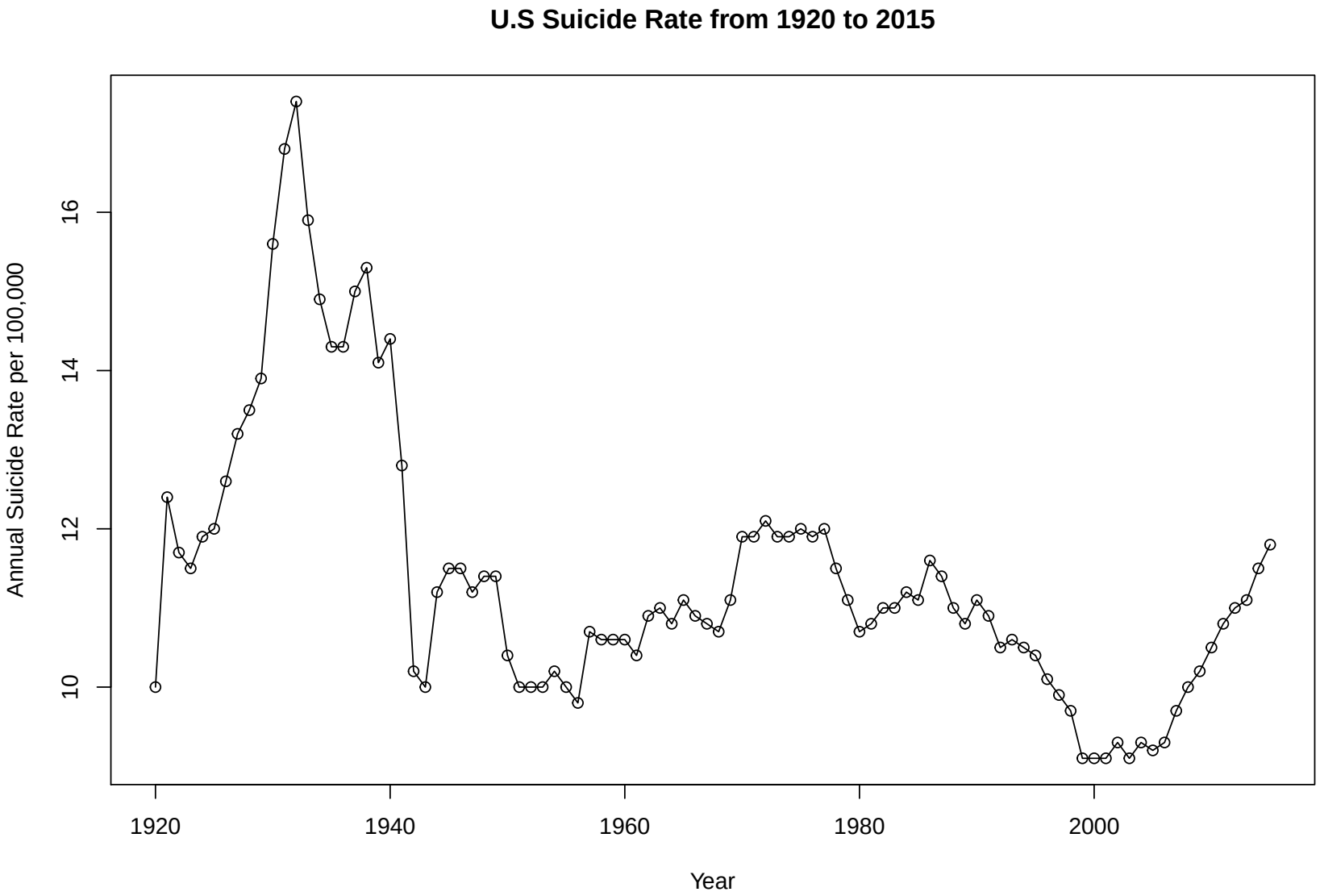
Model diagnostics







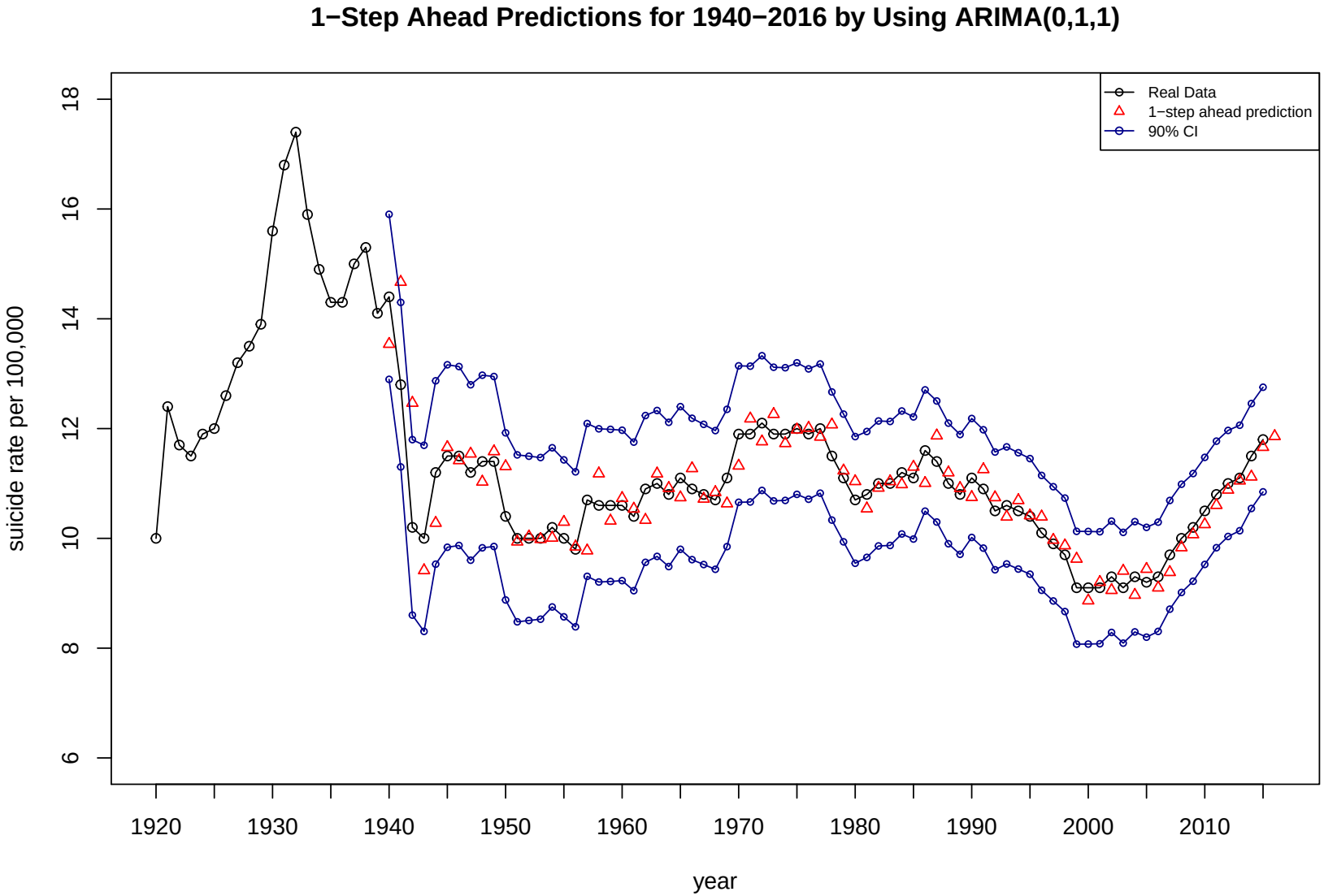
Forcast: 1-step ahead prediction



Forcast: 1-step ahead prediction

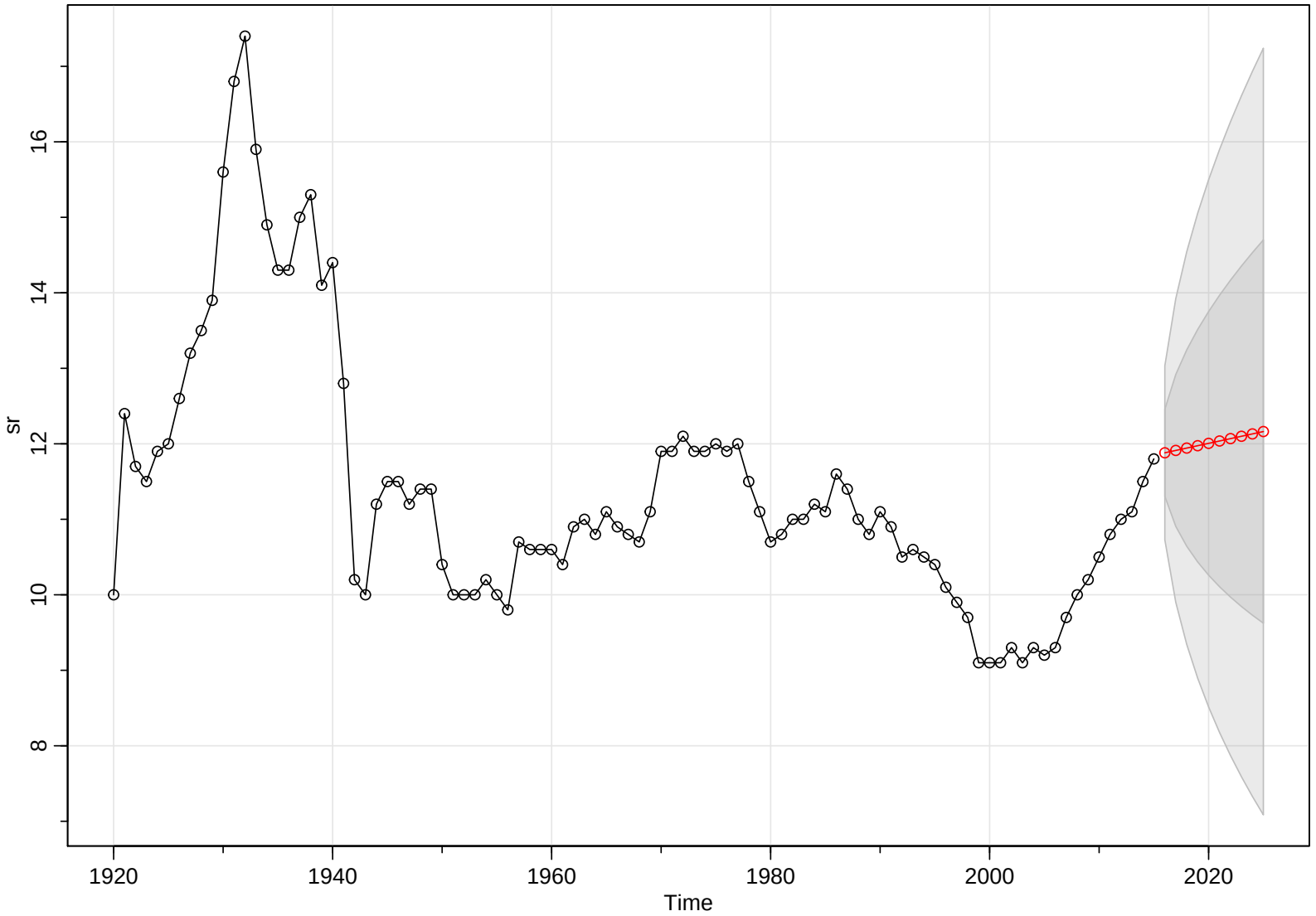
```
pred<-function(data,initial.box=19,prediction=0){

   for (i in 1: (nrow(suicide)- initial.box + prediction) ) {
  pre.data <- suicide[1:(initial.box+i),] #20 years for first step

  ts.data <- ts(pre.data[,2] , start=c(1920), frequency=1) #Times series data
  arma <- arima( ts.data, order=c(0,1,1) )
  pre.1step <-predict(arma,n.ahead = 1)
  estimate[i] <- pre.1step$pred
  s.e[i] <- pre.1step$se
  lowerCI[i] <- suicide.ts[20+i] + s.e[i]*qnorm(alpha/2)
  upperCI[i] <- suicide.ts[20+i] - s.e[i]*qnorm(alpha/2)
   }

est.ts<-ts(estimate,start = 1920+initial.box+1)
cilow.ts<-ts(lowerCI,start = 1920+initial.box+1)
ciup.ts<- ts(upperCI,start = 1920+initial.box+1)
avg.sd<-mean(s.e)
avg.diff<-mean( (estimate[1:76]-ts.data[21:96]) )

table<-data.frame(avg.sd,avg.diff)
  plot(pre.data,type="o",xlim=c(1920,2016+prediction),ylim=c(6,18),
       cex=0.9,xaxt='n',ylab="suicide rate per 100,000",
       xlab="year", lty=1,main="1-Step Ahead Predictions for 1940-2016 by Using ARIMA(0,1,1)")
  axis(1,at=seq(1920,2016+prediction,by=5) )
  points(est.ts,col='red',pch=2,cex=0.8,lty=1)
  lines(cilow.ts,type='o',col='darkblue',cex=0.6,lty=1)
  lines(ciup.ts,type='o',col='darkblue',cex=0.6,lty=1)
  legend("topright", legend=c("Real Data","1-step ahead prediction","90% CI"),
         col=c("black","red", "darkblue"), cex=0.7,lty = c(1,0,1),pch = c(1,2,1) )

  return(table)
```

Forcast: 1-step ahead prediction




```
##       avg sd    avg diff
```

Forcast: 10-steps ahead prediction for 2016-2025



$pred

Forcast: 10-steps ahead prediction for 2016-2025

```
$pred
Time Series:
Start = 2016
End = 2025
Frequency = 1
 [1] 11.88087 11.91215 11.94344 11.97472 12.00600 12.03728 12.06856 12.09985 12.13113
[10] 12.16241

$se
Time Series:
Start = 2016
End = 2025
Frequency = 1
 [1] 0.5761048 1.0055429 1.3001294 1.5393363 1.7460745 1.9308019 2.0993370 2.2553127
 [9] 2.4011779 2.5386759
```

Figure 3: 10-step ahead predictions 2016-2025 by using ARIMA(0,1,1)

Last...

Thank you for listening!!