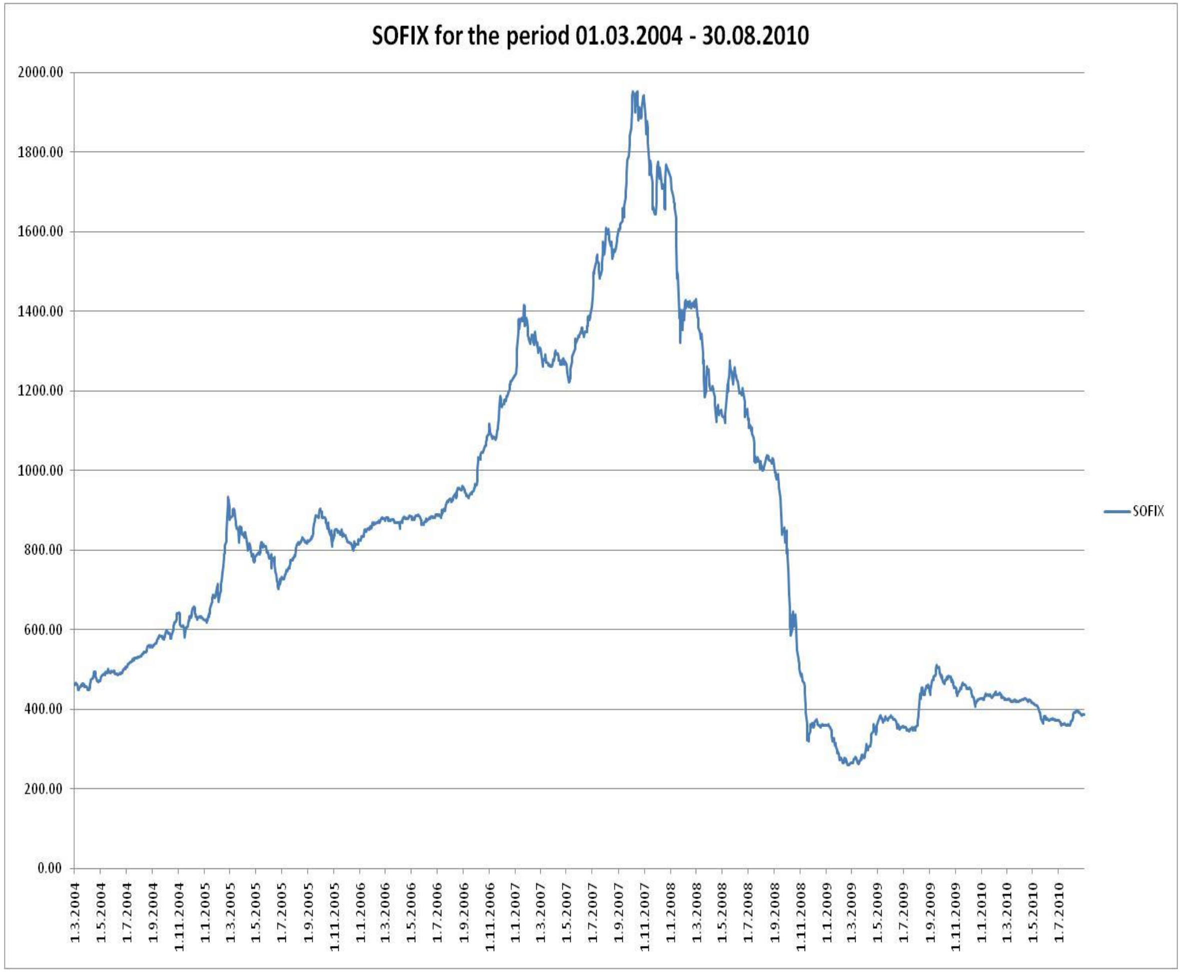
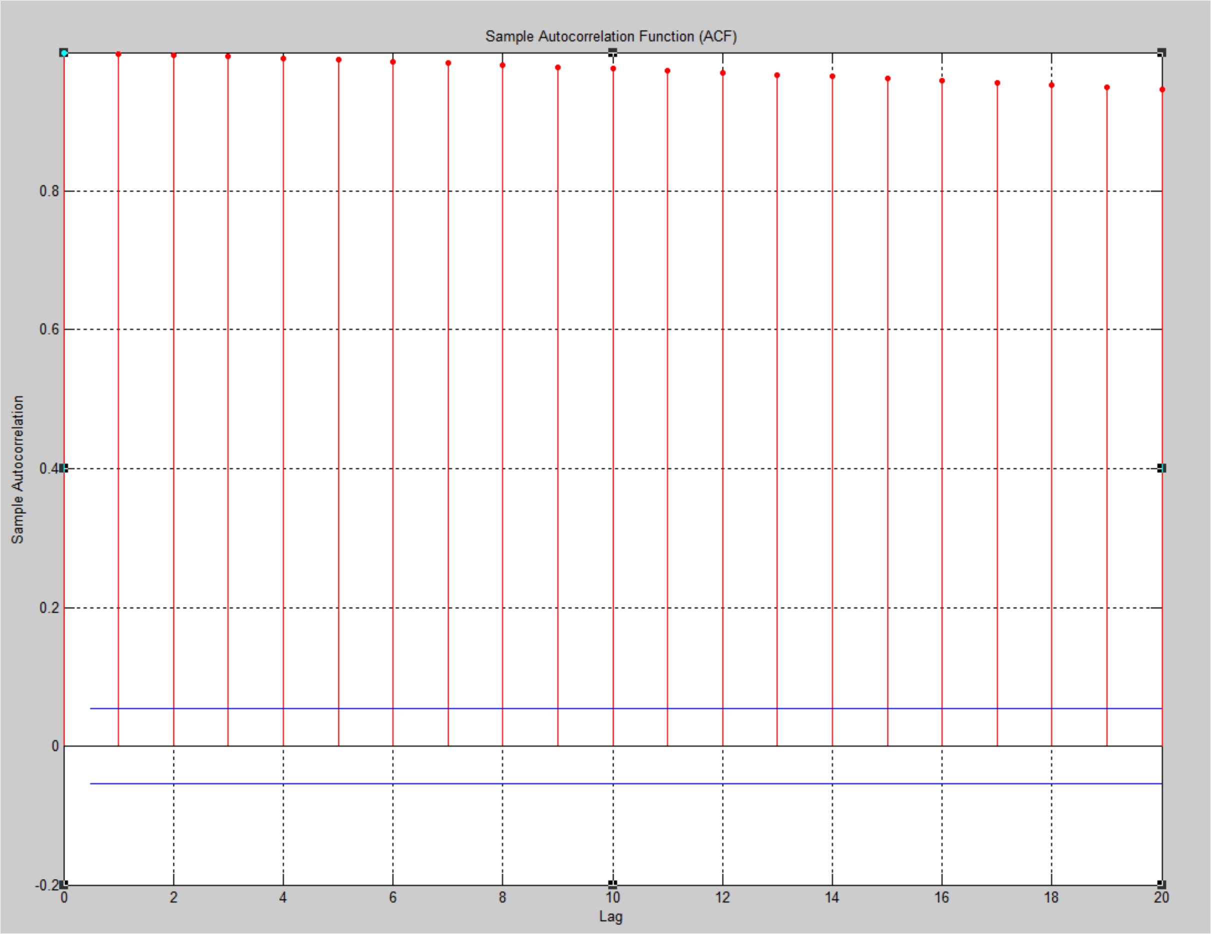
SOFIX Modeling with ARIMA

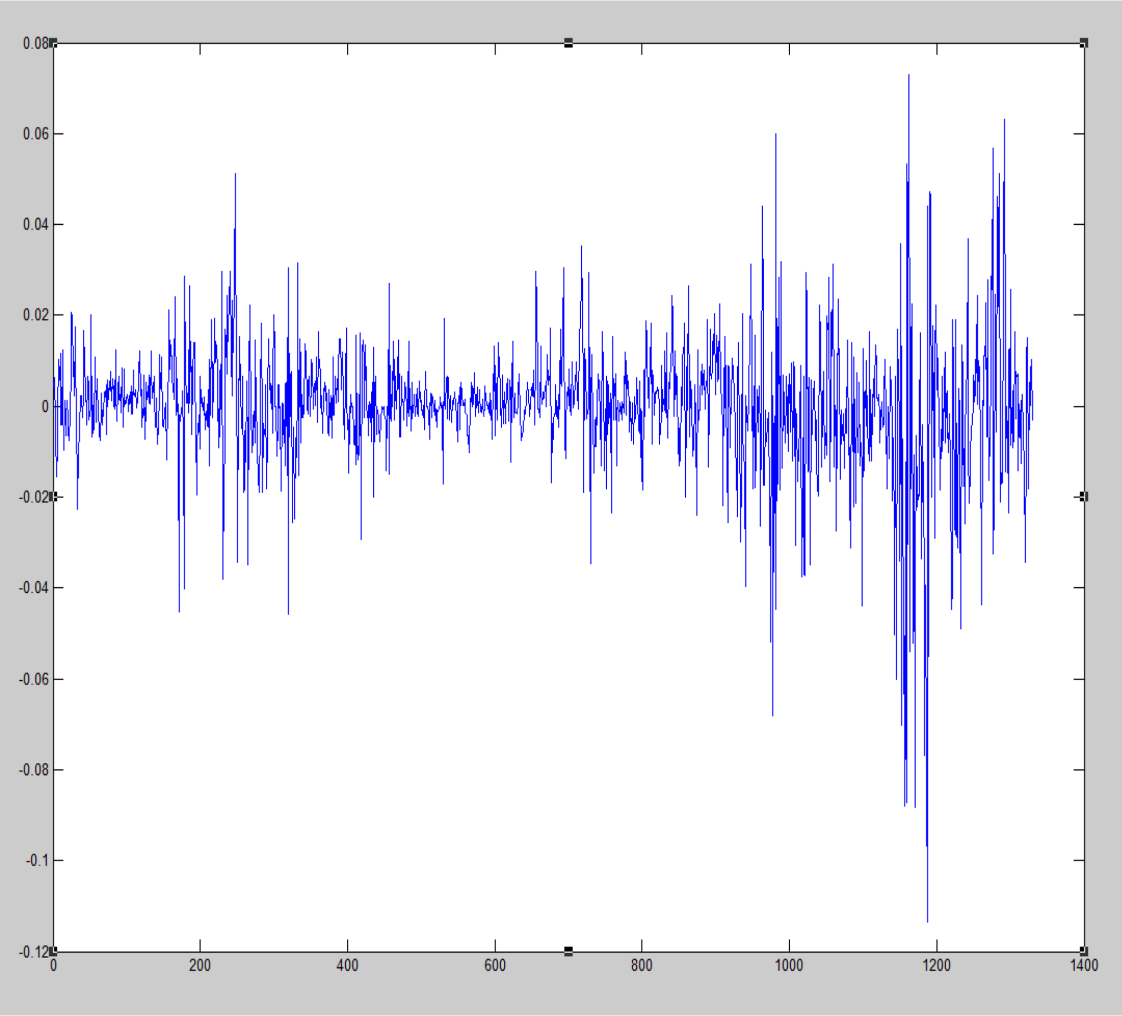
SOFIX plot, 01.03.2004 – 30.08.2010



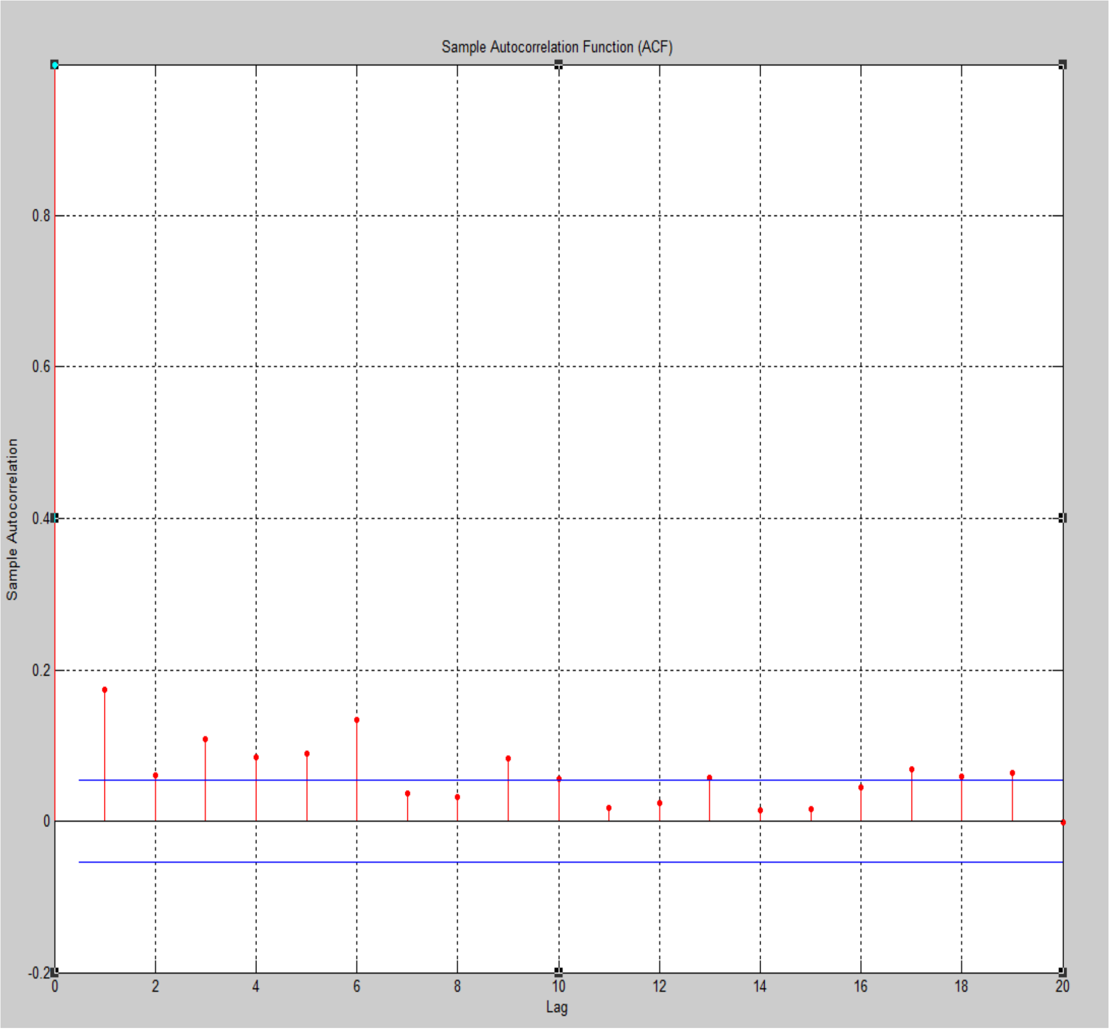
Autocorrelation function for the TS



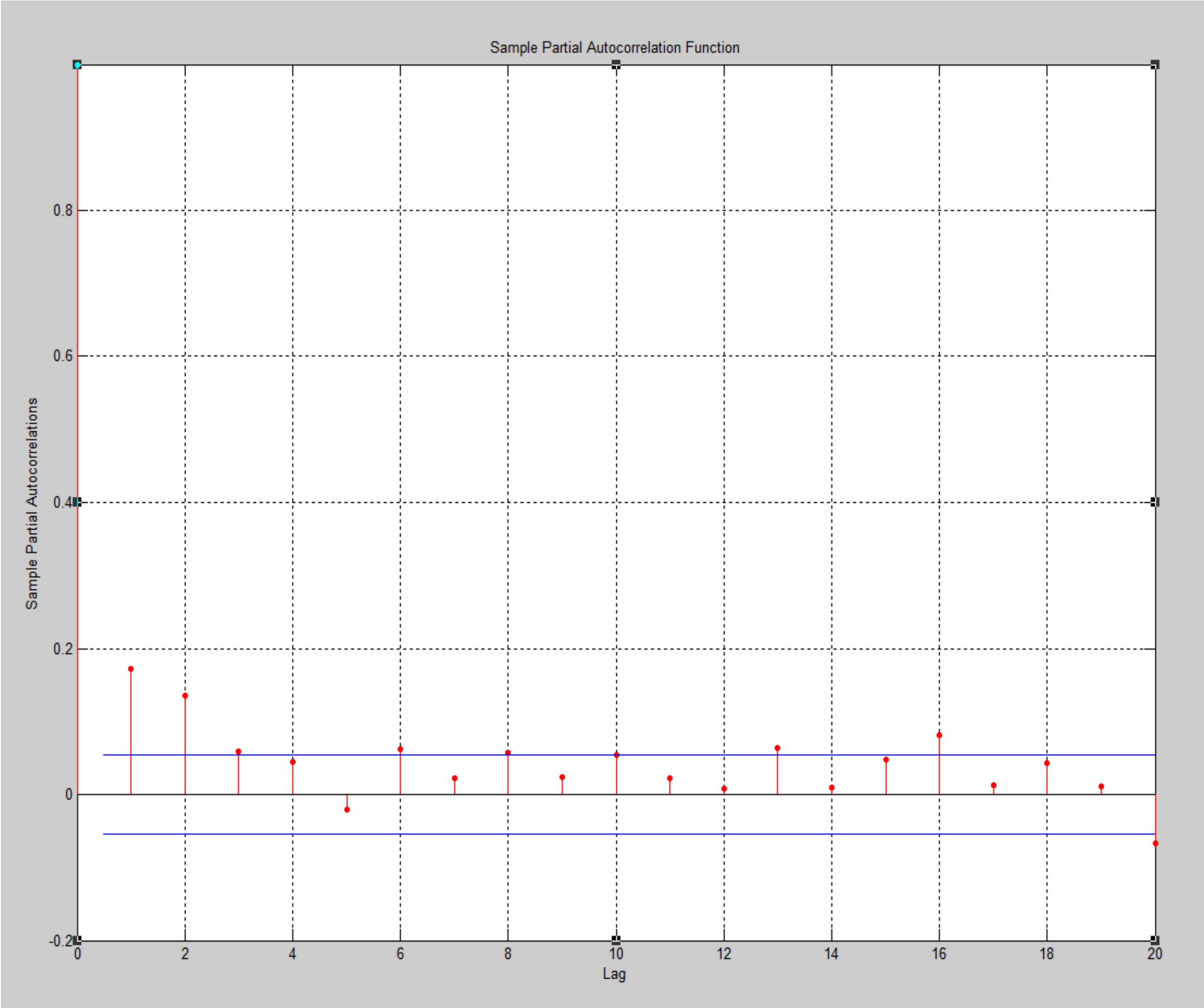
Differenced TS



Autocorrelation function for differenced TS



Partial autocorrelation function for differenced TS



Structure of the models in the initial set

- Order of integration: d=1
- Max AR order: R=3
- Max MA order: M=9
- The Initial set:
  Three pure AR models
  Nine pure MA models
  27 mixed models
- Total number – 39 models

Models estimation and significance testing

- Ten models are with fully significant coefficients
- Namely: MA(1), MA(2), MA(3), MA(4), ARMA(1,1) , ARMA(1,2) , ARMA(1,4) , ARMA(2,1) , ARMA(3,3) , ARMA(3,5)

MA(1)

Model: MA(1)			
Parameter	Value	Standard Error	T Statistic
C	-0.00019219	0.00049929	-0.3849
MA(1)	0.13646	0.011728	11.6359
K	0.00022287	4.0515E-06	55.0089

MA(2)

Model: MA(2)			
Parameter	Value	Standard Error	T Statistic
C	-0.00019027	0.00054945	-0.3463
MA(1)	0.13601	0.012723	10.6897
MA(2)	0.12347	0.012871	9.5935
K	0.00021903	4.0866E-06	53.5975

MA(3)

Model: MA(3)			
Parameter	Value	Standard Error	T Statistic
C	-0.00018906	0.00057957	-0.3262
MA(1)	0.13515	0.013332	10.1374
MA(2)	0.13579	0.013043	10.4109
MA(3)	0.079106	0.014056	5.6278
K	0.00021754	4.1885E-06	51.9374

MA(4)

Model: MA(4)			
Parameter	Value	Standard Error	T Statistic
C	-0.00018682	0.000641	-0.2914
MA(1)	0.14269	0.013125	10.8718
MA(2)	0.13933	0.013394	10.4026
MA(3)	0.087589	0.014048	6.2351
MA(4)	0.066876	0.014274	4.685
K	0.00021649	4.17E-06	51.9645

ARMA(1,1)

Model: ARMA(1,1)			
Parameter	Value	Standard Error	T Statistic
C	-1.4334E-05	0.000079039	-0.1813
AR(1)	0.92543	0.015592	59.3542
MA(1)	-0.83344	0.022558	-36.9457
K	0.00021549	3.9157E-06	55.0318

ARMA(1,2)

Model: ARMA(1,2)			
Parameter	Value	Standard Error	T Statistic
C	-7.7295E-06	0.000048246	-0.1602
AR(1)	0.96228	0.013072	73.6159
MA(1)	-0.83943	0.019301	-43.4914
MA(2)	-0.059504	0.013811	-4.3083
K	0.00021483	3.9316E-06	54.6411

ARMA(1,4)

Model: ARMA(1,4)			
Parameter	Value	Standard Error	T Statistic
C	-0.00032385	0.0010276	-0.3151
AR(1)	-0.69429	0.099098	-7.0061
MA(1)	0.83732	0.10251	8.1681
MA(2)	0.23549	0.023707	9.9333
MA(3)	0.17977	0.024176	7.4358
MA(4)	0.11515	0.01541	7.4721
K	0.00021607	4.2753E-06	50.5395

ARMA(2,1)

Model: ARMA(2,1)			
Parameter	Value	Standard Error	T Statistic
C	-6.3606E-06	0.000040622	-0.1566
AR(1)	1.0465	0.022846	45.8063
AR(2)	-0.077167	0.015008	-5.1418
MA(1)	-0.91512	0.020212	-45.2766
K	0.00021468	3.9342E-06	54.5678

ARMA(3,3)

Model: ARMA(3,3)			
Parameter	Value	Standard Error	T Statistic
C	-3.9203E-06	0.00002875	-0.1364
AR(1)	0.63461	0.12992	4.8848
AR(2)	0.74692	0.15611	4.7847
AR(3)	-0.4005	0.14252	-2.8101
MA(1)	-0.50563	0.13546	-3.7326
MA(2)	-0.70597	0.15405	-4.5828
MA(3)	0.27054	0.1357	1.9936
K	0.0002132	3.9706E-06	53.6962

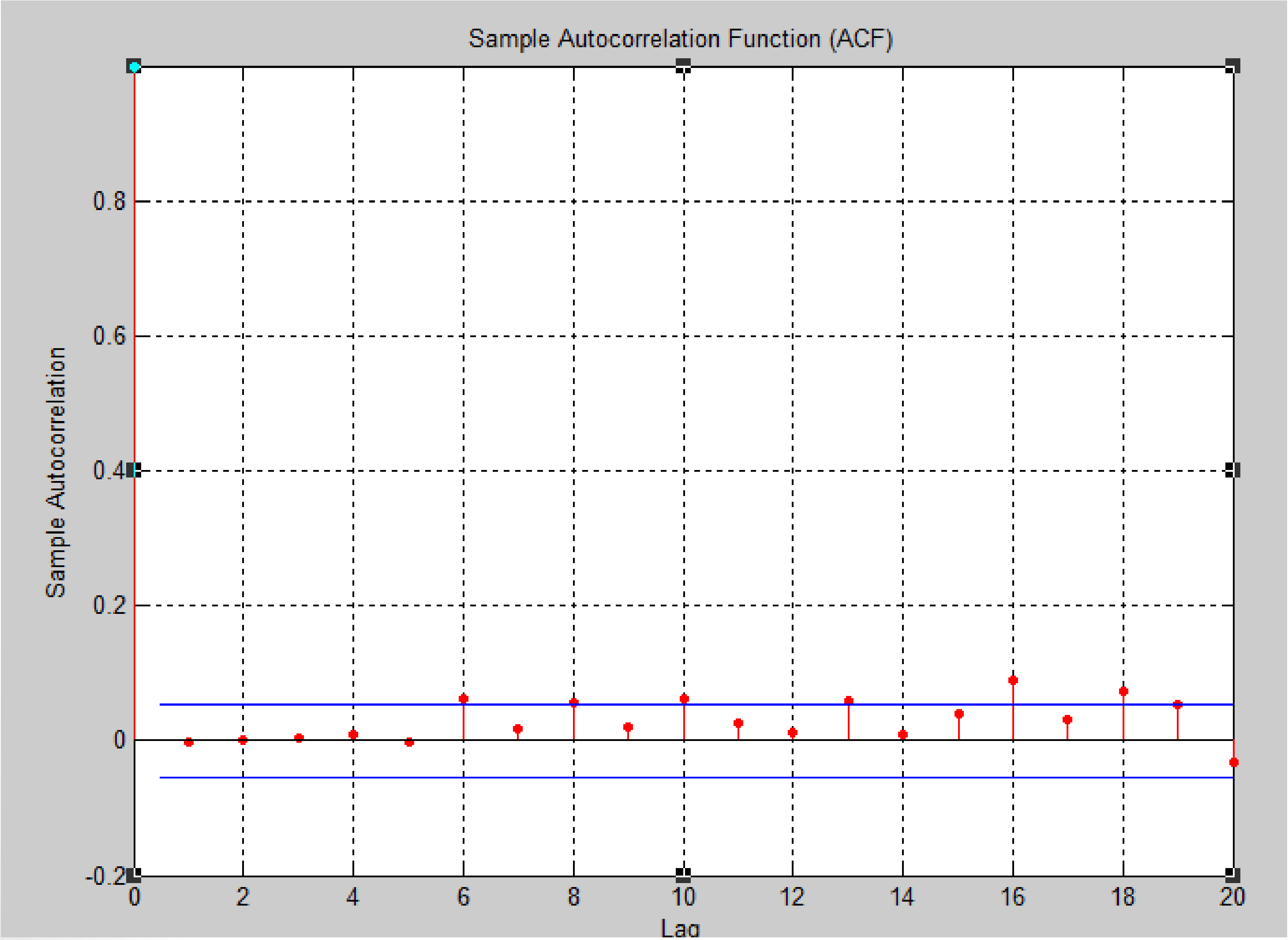
ARMA(3,5)

Model: ARMA(3,5)			
Parameter	Value	Standard Error	T Statistic
C	-1.66E-05	0.00011721	-0.1413
AR(1)	-0.62533	0.053515	-11.6852
AR(2)	0.69684	0.028703	24.2775
AR(3)	0.84808	0.046416	18.2714
MA(1)	0.75378	0.055704	13.5319
MA(2)	-4.98E-01	0.036074	-13.8139
MA(3)	-0.81061	0.048654	-16.6607
MA(4)	-0.10614	0.021388	-4.9628
MA(5)	-0.094801	0.017265	-5.491
K	0.00021204	4.07E-06	52.096

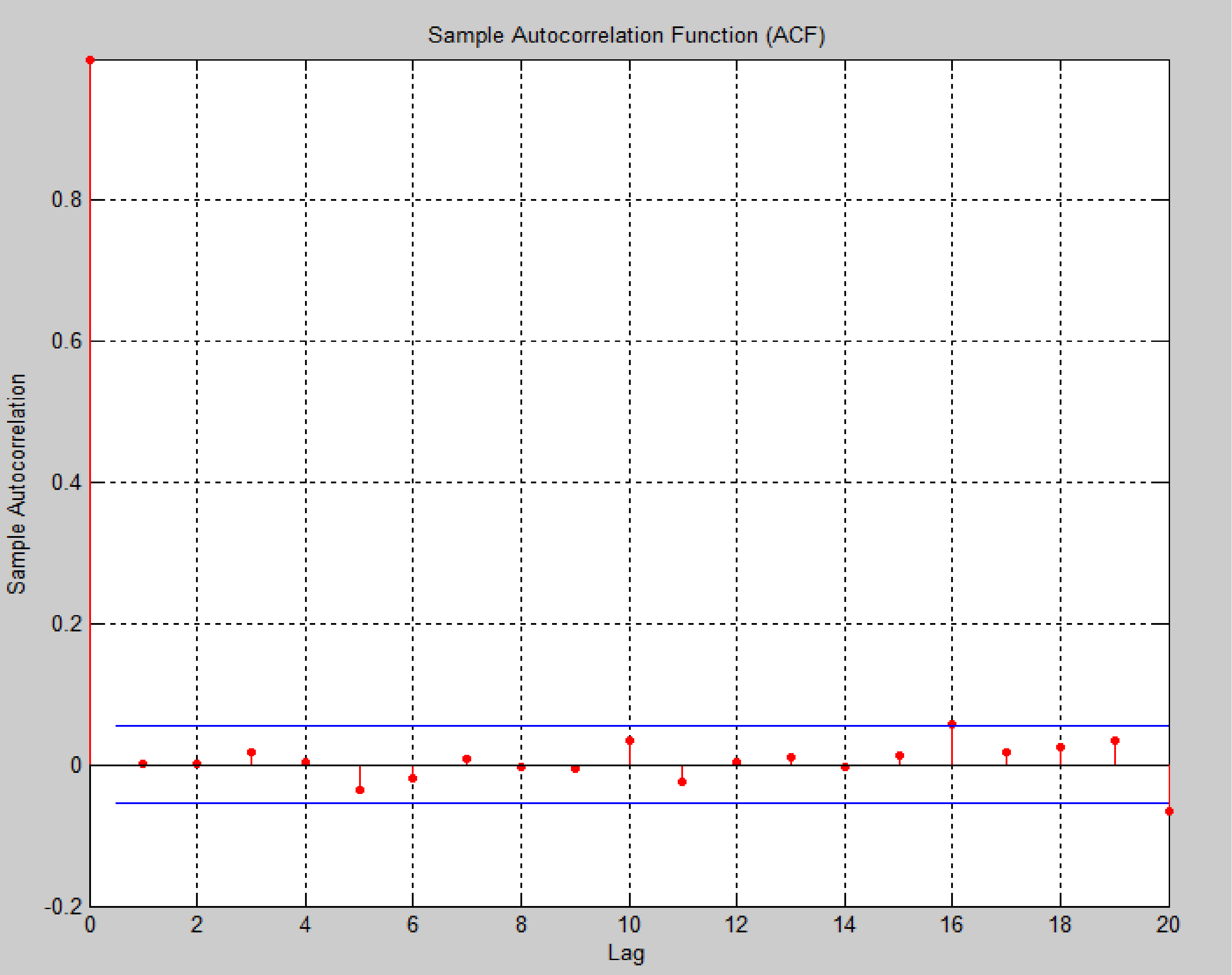
AIC and BIC for model selection

	MA(1)	MA(2)	MA(3)	MA(4)	ARMA(1,1)
AIC	-5.41E+03	-5.4325E+03	-5.4416 E+03	-5.481E+03	-5.45E+03
BIC	-6.31E+03	-6.3108E+03	-6.3108E+03	-6.3108 E+03	-6.07E+03
	ARMA(1,2)	**ARMA(1,4)**	**ARMA(2,1)**	**ARMA(3,3)**	**ARMA(3,5)**
AIC	-5.458E+03	-5.450E+03	-5.4592E+03	-5.4684E+03	-5.4757E+03
BIC	-5.8489E+03	-5.5047E+03	-5.8498E+03	-5.3890E+03	-5.1731E+03

MA(4) test for adequacy



ARMA(3,5) test for adequacy



The Final choice: ARMA(3,5)

Forecasting with ARIMA(R,d,M) models

Forecast with minimal square error

$$\hat{Y}_t(h) = E(Y_{t+h} \mid Y_t, Y_{t-1}, \ldots, Y_1)$$

Forecast calculation

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots \phi_R Y_{t-R} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots \theta_M e_{t-M}$$

$$\begin{aligned}\hat{Y}_t(1) &= E(Y_{t+1} \mid Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_1) = \\ &= C + \phi_1 Y_t + \phi_2 Y_{t-1} + \cdots \phi_R Y_{t-R+1} + \theta_1 \hat{e}_t + \theta_2 \hat{e}_{t-1} + \cdots \theta_M \hat{e}_{t-M+1}\end{aligned}$$

$$\hat{Y}_t(2) = C + \phi_1 \hat{Y}_t(1) + \phi_2 Y_t + \cdots \phi_R Y_{t-R+2} + \theta_2 \hat{e}_t + \cdots \theta_M \hat{e}_{t-M+2}$$

Forecast calculation

$$\hat{Y}_t(h) = \phi_1 \hat{Y}_t(h-1) + \cdots + \phi_h Y_t + \cdots + \phi_R Y_{t-R+h}$$
$$+ \theta_h \hat{e}_t + \cdots + \theta_M \hat{e}_{t-M+h}$$

$$If \quad h > R \; and \; h > M \, , then$$

$$\hat{Y}_t(h) = \phi_1 \hat{Y}_t(h-1) + \cdots + \phi_R \hat{Y}_t(h-R)$$

Forecast calculation

$$d = 1$$

$$\hat{Y}_t(h) = Y_t + \Delta\hat{Y}_t(1) + \Delta\hat{Y}_t(2) + \cdots + \Delta\hat{Y}_t(h)$$

$$d = 2$$

$$\hat{Y}_t(h) = Y_t + [\Delta Y_t + \Delta^2\hat{Y}_t(1)] + [\Delta Y_t + \Delta^2\hat{Y}_t(1) + \Delta^2\hat{Y}_t(2)] + \cdots$$
$$+ [\Delta Y_t + \Delta^2\hat{Y}_t(1) + \Delta^2\hat{Y}_t(2) + \cdots \Delta^2\hat{Y}_t(h)]$$

Forecast error

$$Y_t = \phi^{-1}(B)\theta(B)e_t = \psi(B)e_t = \sum_{j=0}^{\infty} \psi_j e_{t-j}$$

$$Y_{t+h} = \psi_0 e_{t+h} + \psi_1 e_{t+h-1} + \cdots + \psi_{h-1} e_{t+1} + \sum_{j=0}^{\infty} \psi_{h+j} e_{t-j}$$

$$\hat{Y}_t(h) = \sum_{j=0}^{\infty} \psi_{h+j} e_{t-j}$$

$$Er_t(h) = \psi_0 e_{t+h} + \psi_1 e_{t+h-1} + \cdots + \psi_{h-1} e_{t+1}$$

$$D[Er_t(h)] = (\psi_0^2 + \psi_1^2 + \cdots + \psi_{h-1}^2)\sigma_e^2$$

Forecasting with AR(1)

$$Y_t = C + \phi_1 Y_{t-1} + e_t$$

$$\hat{Y}_t(1) = E(Y_{t+1} \mid Y_t, Y_{t-1}, \ldots, Y_1) = C + \phi_1 Y_t$$

$$\hat{Y}_t(2) = C + \phi_1 \hat{Y}_t(1) = C(1 + \phi_1) + \phi_1^2 Y_t$$

$$\hat{Y}_t(h) = C(1 + \phi_1 + \phi_1^2 + \cdots + \phi_1^{h-1}) + \phi_1^h Y_t$$

$$\lim_{h\to\infty} \hat{Y}_t(h) = C\sum_{j=0}^{\infty} \phi_1^j = \frac{c}{1-\phi_1} = E(Y_t)$$

$$Er_t(h) = e_{t+h} + \phi_1 e_{t+h-1} + \cdots + \phi_1^{h-1} e_{t+1}$$

$$D[Er_t(h)] = (1 + \phi_1^2 + \cdots + \phi_1^{2h-2})\sigma_e^2$$

Forecasting with MA(1)

$Y_t = C + e_t + \theta_1 e_{t-1}$

$\hat{Y}_t(1) = E(Y_{t+1}|Y_t, Y_{t-1}, \ldots, Y_1) = C + \theta_1 \hat{e}_t$

$\hat{Y}_t(h) = E(Y_{t+h}|Y_t, Y_{t-1}, \ldots, Y_1) = E(C + e_{t+h} + \theta_1 e_{t+h-1}) =$

$= C, \quad for\ h > 1$

$D[Er_t(h)] = E\{[Y_{t+h} - \hat{Y}_t(h)]^2\} = E[(e_{t+h} + \theta_1 e_{t+h-1})^2] =$

$= (1 + \theta_1^2)\sigma_e^2, \quad for\ h > 1$

Forecasting with ARMA(1,1)

$$Y_t = C + \phi_1 Y_{t-1} + e_t + \theta_1 e_{t-1}$$

$$\hat{Y}_t(1) = C + \phi_1 Y_t + \theta_1 \hat{e}_t$$

$$\hat{Y}_t(2) = C + \phi_1 \hat{Y}_t(1) = C(1 + \phi_1) + \phi_1^2 Y_t + \phi_1 \theta_1 \hat{e}_t$$

$$\hat{Y}_t(h) = C(1 + \phi_1 + \phi_1^2 + \cdots \phi_1^{h-1}) + \phi_1^h Y_t + \phi_1^{h-1} \theta_1 \hat{e}_t$$

Example 1

- We have the following model: $dY_t = 0.1dY_{t-1} + e_t + 0.2e_{t-1}$
- And the data:

T	Y
-1	
0	-11
1	2
2	5
3	-1
4	13
5	4

- Calculate the forecast for time 6 and 7 having in mind that *d* is 1, namely we have ARIMA(1,1,1) model
- Example for P7

Example 2

For the following model:

$$Y_t = 0.3Y_{t-1} + 0.6Y_{t-2} - 0.2Y_{t-3} + 2 + e_t - 0.9e_{t-1}$$

Calculate the forecast for 50 intervals in the future.

```
arma31=arima('c',2, 'var' ,1, 'ar',[0.3 0.6 -0.2],'ma',-0.9)
y=simulate(arma31,1000);
Y=y(951:end);
```

Example 2

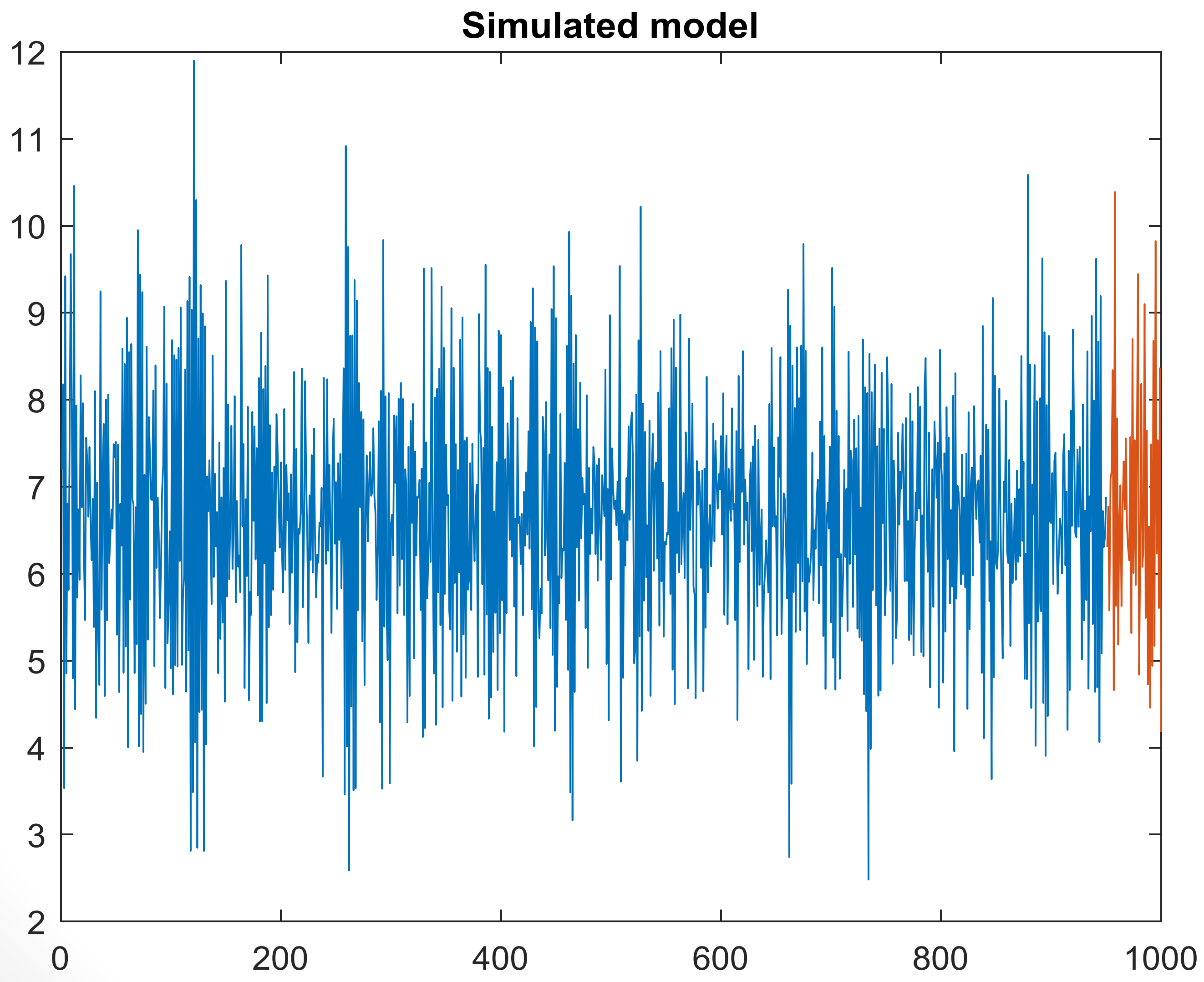
```
[Yf,YMSE] = forecast(arma31,50,Y);



figure,plot(y(1:950)),hold on,plot([951:1000],Yf)


figure, plot(y(900:950)),hold on, plot([51:100],Yf)
xticklabels({'900','920','940','960','980','1000'})
```

Example 2



Example 2 Forecast