Financial Time Series Forecasting

Yiqing Liu, Wenzhe Zhang

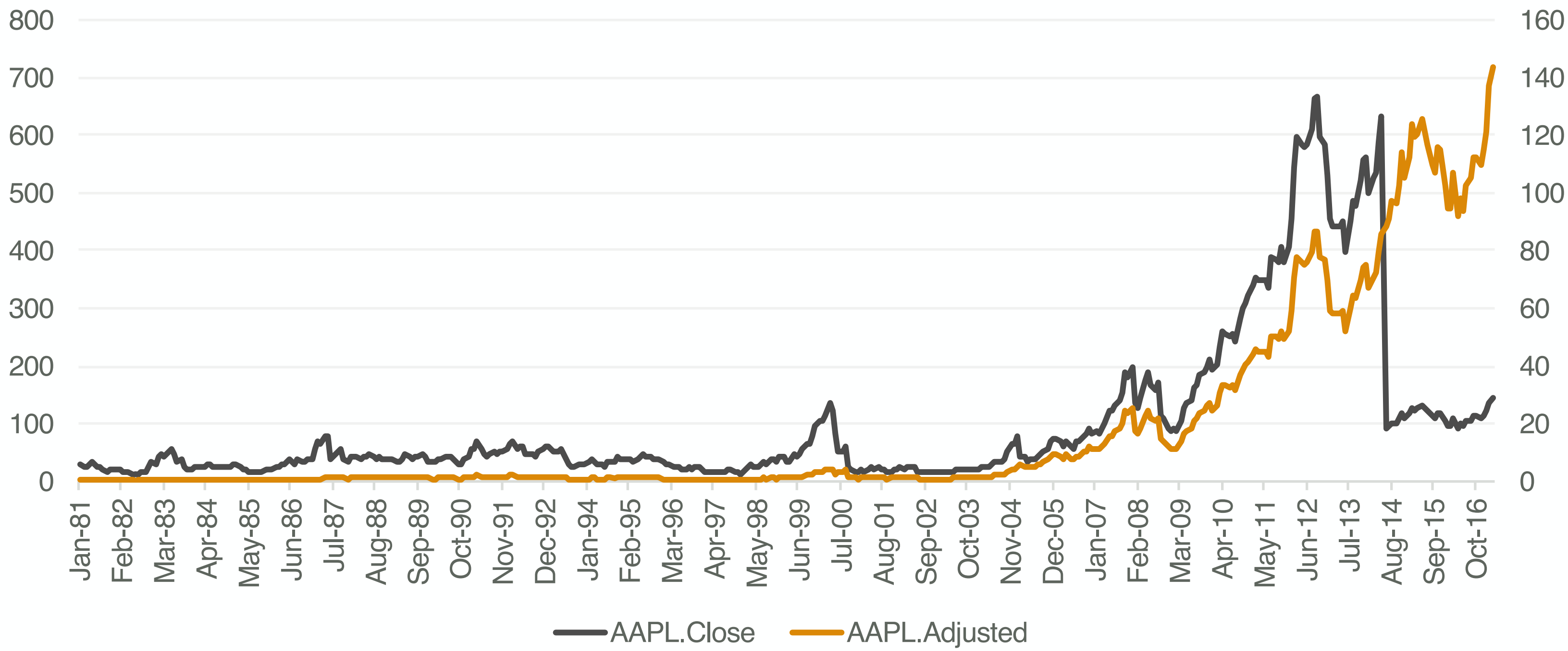
Background
Data
SARIMA
LSTM
SVM
Conclusions

Background

- Time Series Forecast
  - Tracking of a particular variable over time
  - Stable, Recurring Patterns, Seasonal Swings...
- Why Do Forecast?
  - Risk Hedging
  - Investment Planning
- Critical Questions
  - Estimating Accuracy of a Forecast RMSE MAE
  - Meeting the Criteria for Timeliness Monthly
- Methods
  - Random Walk SARIMA
  - Artificial Neural Networks LSTM
  - Support Vector Machine RBF Linear Polynomial

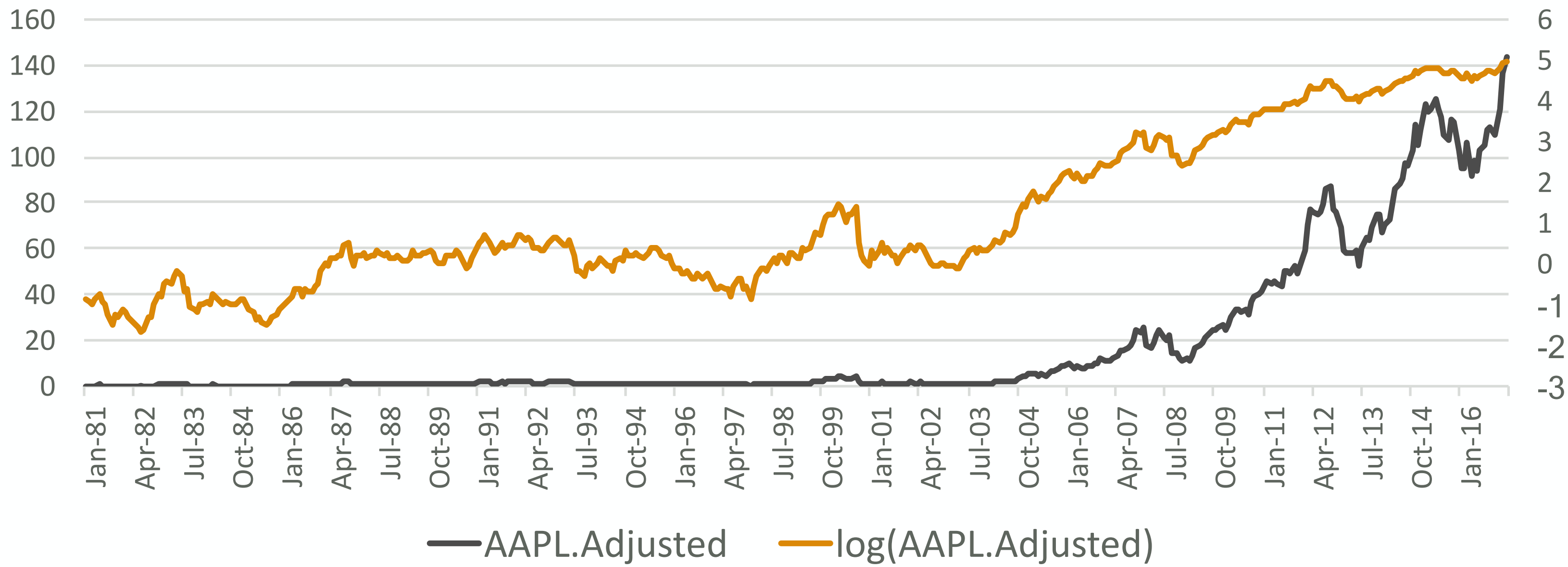
Data

- ➢ Apple Stock Monthly Price(AAPL)
  - ➢ Training Data: 1981/01 - 2015/12
  - ➢ Testing Data: 2016/01 - 2017/03

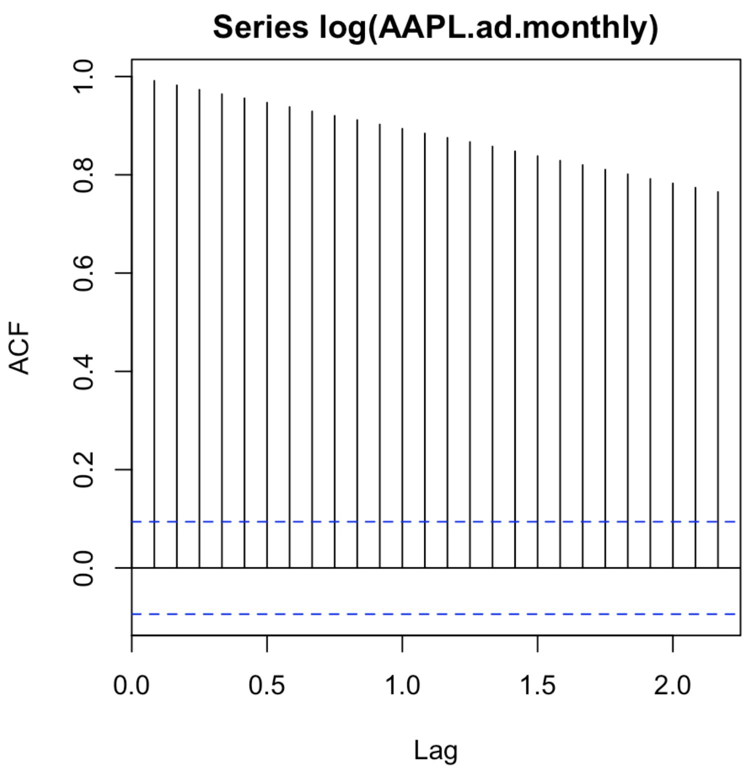


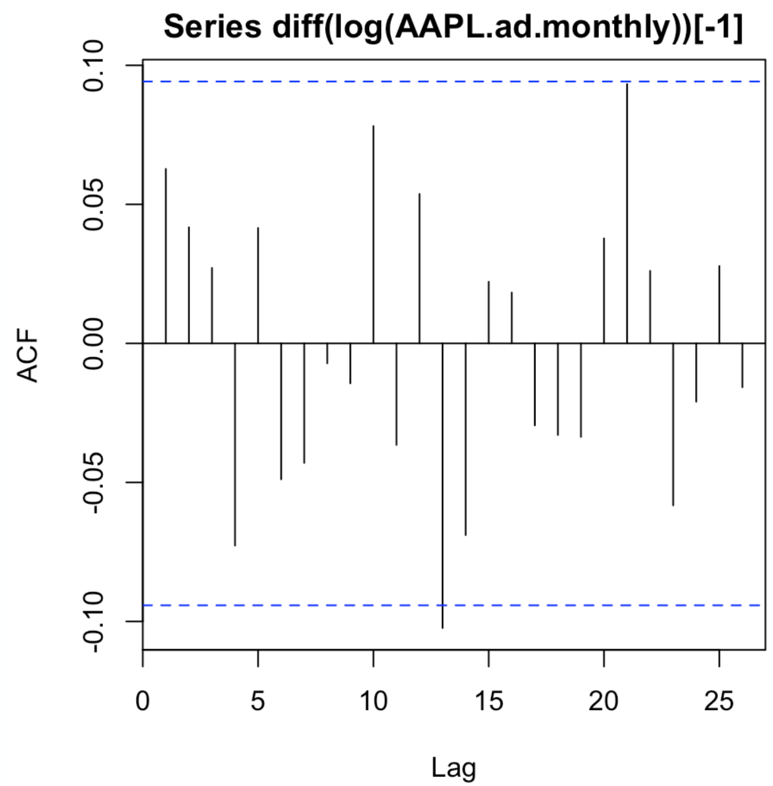
Data

- Variability grows with time
  - log transformation
  - Stabilize variance



SARIMA

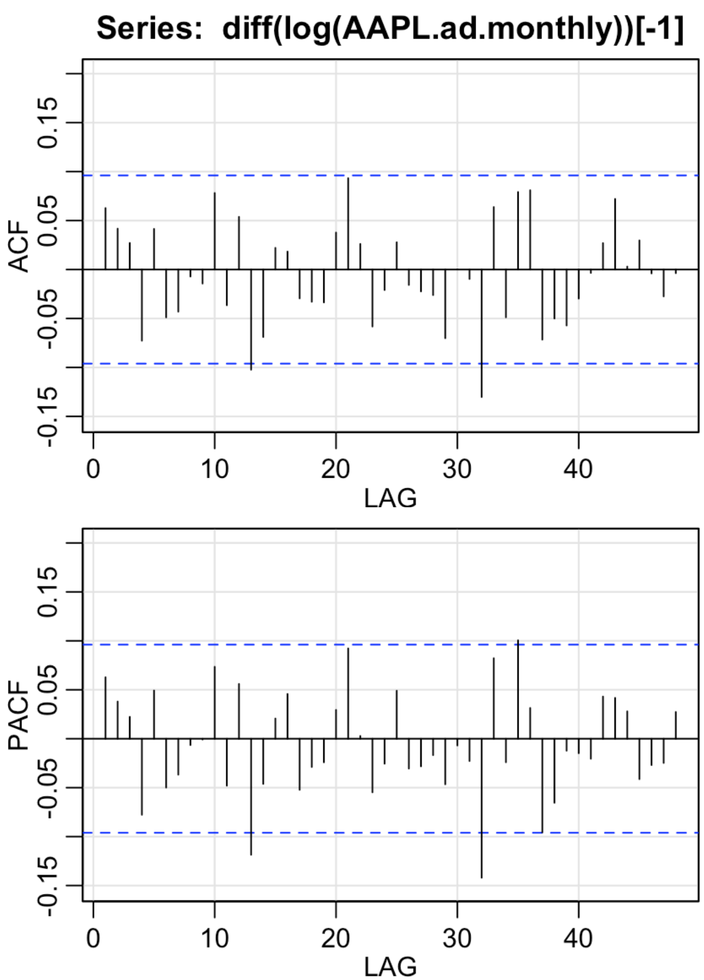




A Slow decay in sample ACF

Differencing

SARIMA



AR/MA

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	o	o	o	o	o	o	o	o	o	o	o	o	x	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	x	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	x	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	x	o	o	o	o
4	x	x	x	o	o	o	o	o	o	o	o	o	x	o	o	o	o
5	x	x	x	o	o	o	o	o	o	o	o	o	x	o	o	o	o
6	x	x	o	x	x	x	o	o	o	o	o	o	o	o	o	o	o
7	x	o	o	x	x	o	x	o	o	o	o	o	x	o	o	o	o
8	x	o	o	x	o	o	x	x	o	o	o	o	o	o	o	o	o
9	o	o	x	x	o	o	x	o	x	o	o	o	o	o	o	o	o
10	x	x	x	x	x	x	x	x	o	x	o	o	o	o	o	o	o
11	x	x	x	x	o	x	x	x	o	x	o	o	o	o	o	o	o
12	x	x	x	x	o	x	x	x	x	x	o	x	o	o	o	o	o
13	x	x	x	x	x	x	x	x	x	x	o	o	x	o	o	o	o
14	x	x	x	x	x	x	o	x	o	o	o	o	x	x	o	o	o
15	x	x	o	x	o	o	o	x	x	o	o	o	x	o	o	o	o
16	x	x	o	x	o	x	x	x	x	o	o	o	x	x	o	o	o

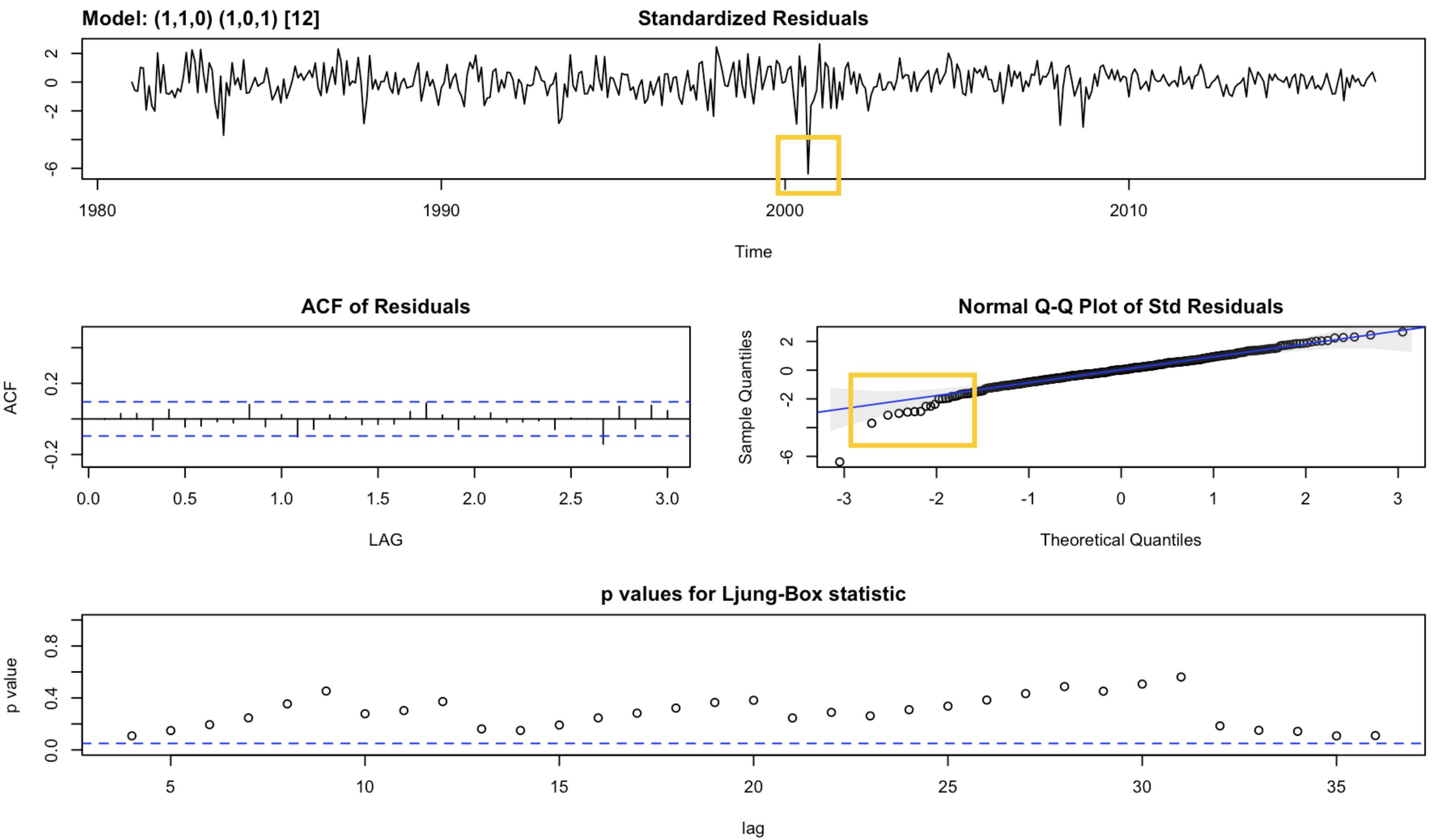
Seasonality

Low-order

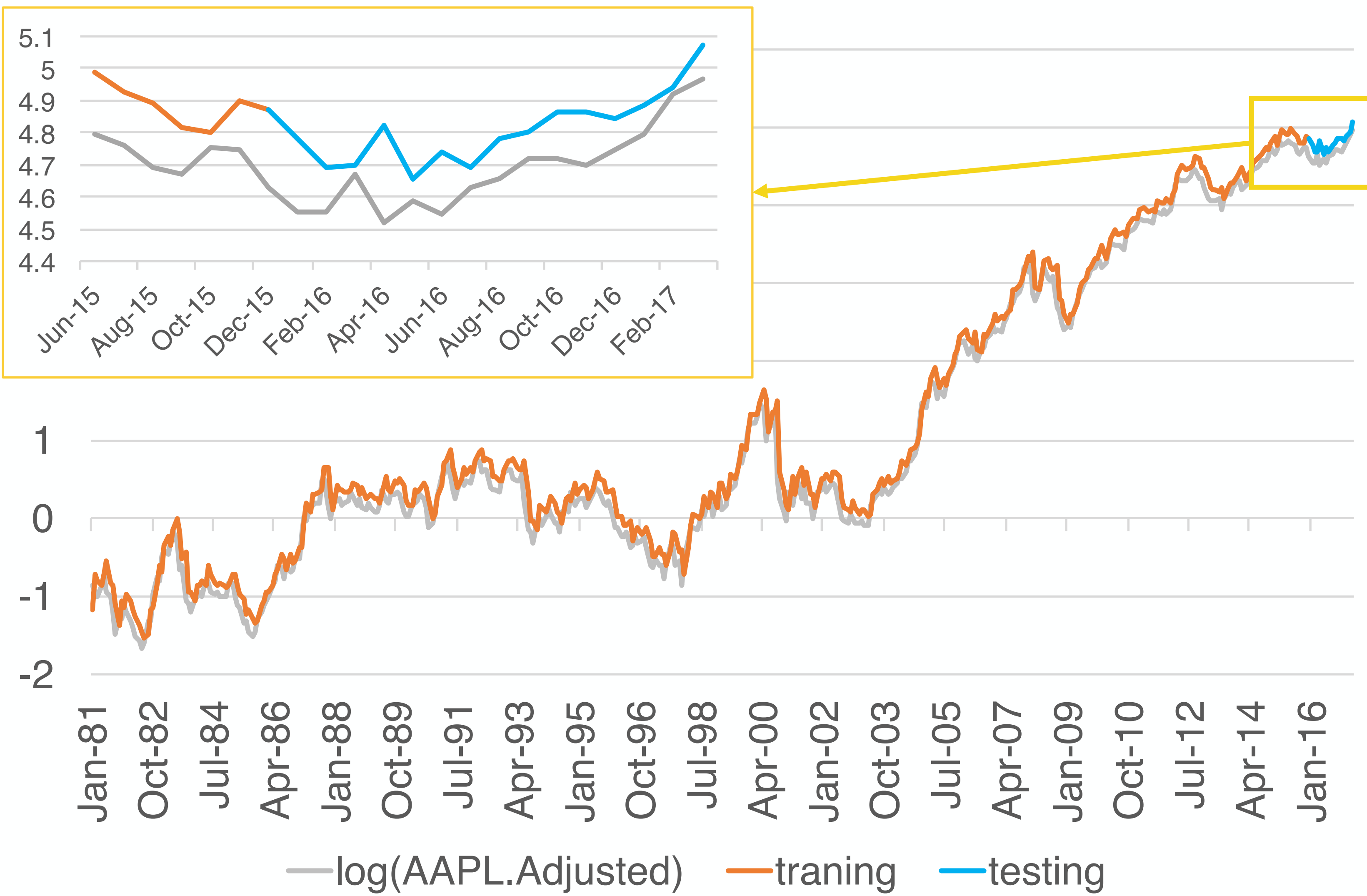
SARIMA

Model	AIC	RMSE	MAE
ARIMA(1,1,0)(1,1,0)[12]	-308.72	0.1636	0.1205
ARIMA(1,1,0)(2,0,0)[12]	-490.59	0.1358	0.1010
ARIMA(2,1,0)(1,0,0)[12]	-490.84	0.1357	0.1008
ARIMA(1,1,1)(0,0,1)[12]	-490.88	0.1357	0.1007
ARIMA(1,1,0)(2,0,1)[12]	-491.69	0.1349	0.1002
ARIMA(1,1,0)(1,0,0)[12]	-492.31	0.1358	0.1010
ARIMA(1,1,0)(1,0,1)[12]	-493.26	0.1350	0.1004
mean	**-465.47**	**0.1395**	**0.1035**

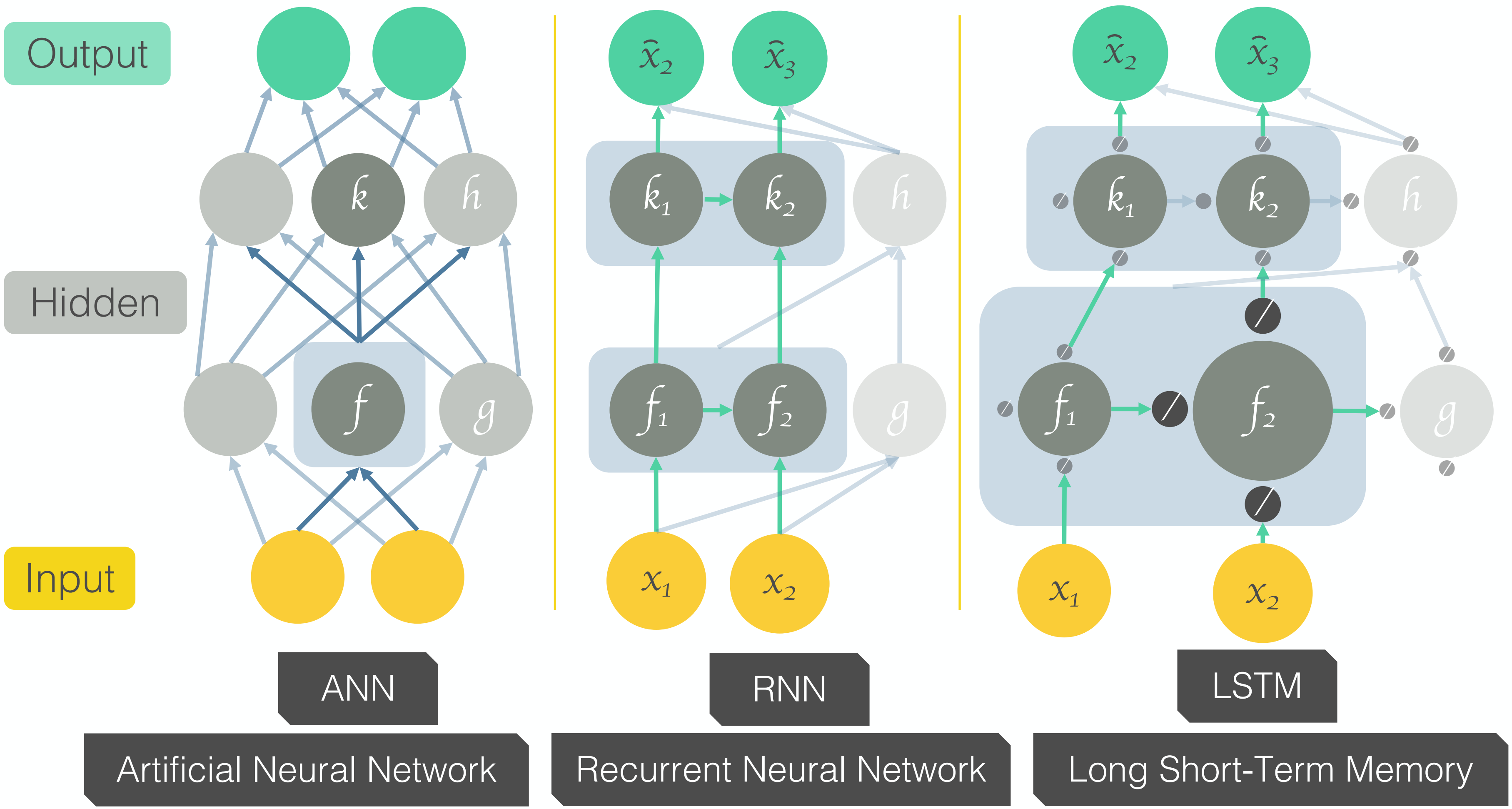
SARIMA


Model: (1,1,0) (1,0,1) [12]
Standardized Residuals
Time
ACF of Residuals
ACF
LAG
Normal Q-Q Plot of Std Residuals
Sample Quantiles
Theoretical Quantiles
p values for Ljung-Box statistic
p value
lag

SARIMA

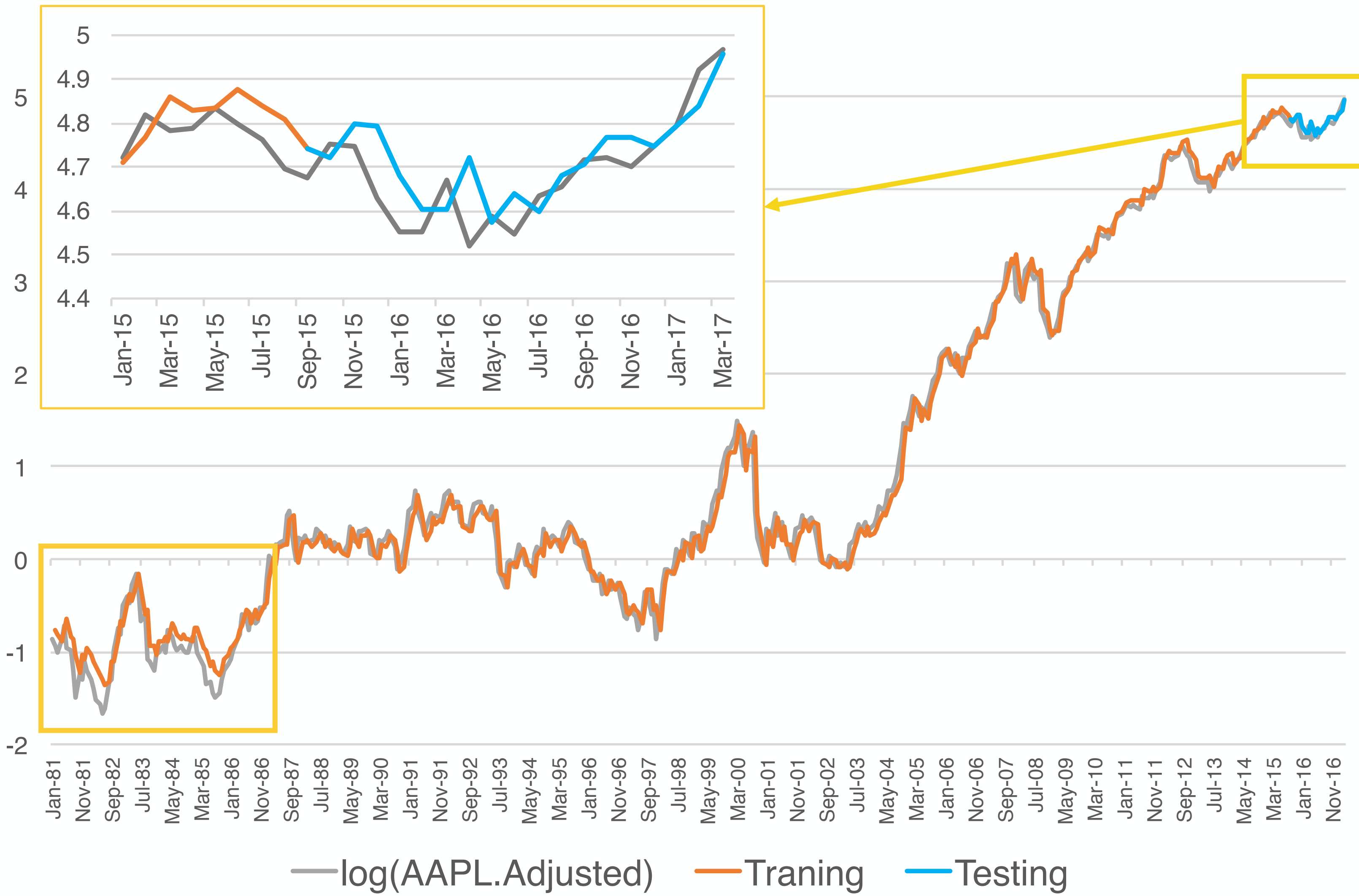

5.1
5
4.9
4.8
4.7
4.6
4.5
4.4
Jun-15
Aug-15
Oct-15
Dec-15
Feb-16
Apr-16
Jun-16
Aug-16
Oct-16
Dec-16
Feb-17
1
0
-1
-2
Jan-81
Oct-82
Jul-84
Apr-86
Jan-88
Oct-89
Jul-91
Apr-93
Jan-95
Oct-96
Jul-98
Apr-00
Jan-02
Oct-03
Jul-05
Apr-07
Jan-09
Oct-10
Jul-12
Apr-14
Jan-16
log(AAPL.Adjusted)
traning
testing

LSTM



Lack of Memory sucks when you’re working with anything time series related

LSTM



5
4.9
4.8
4.7
4.6
4.5
4.4
Jan-15
Mar-15
May-15
Jul-15
Sep-15
Nov-15
Jan-16
Mar-16
May-16
Jul-16
Sep-16
Nov-16
Jan-17
Mar-17
5
4
3
2
1
0
-1
-2
Jan-81
Nov-81
Sep-82
Jul-83
May-84
Mar-85
Jan-86
Nov-86
Sep-87
Jul-88
May-89
Mar-90
Jan-91
Nov-91
Sep-92
Jul-93
May-94
Mar-95
Jan-96
Nov-96
Sep-97
Jul-98
May-99
Mar-00
Jan-01
Nov-01
Sep-02
Jul-03
May-04
Mar-05
Jan-06
Nov-06
Sep-07
Jul-08
May-09
Mar-10
Jan-11
Nov-11
Sep-12
Jul-13
May-14
Mar-15
Jan-16
Nov-16
log(AAPL.Adjusted)
Traning
Testing

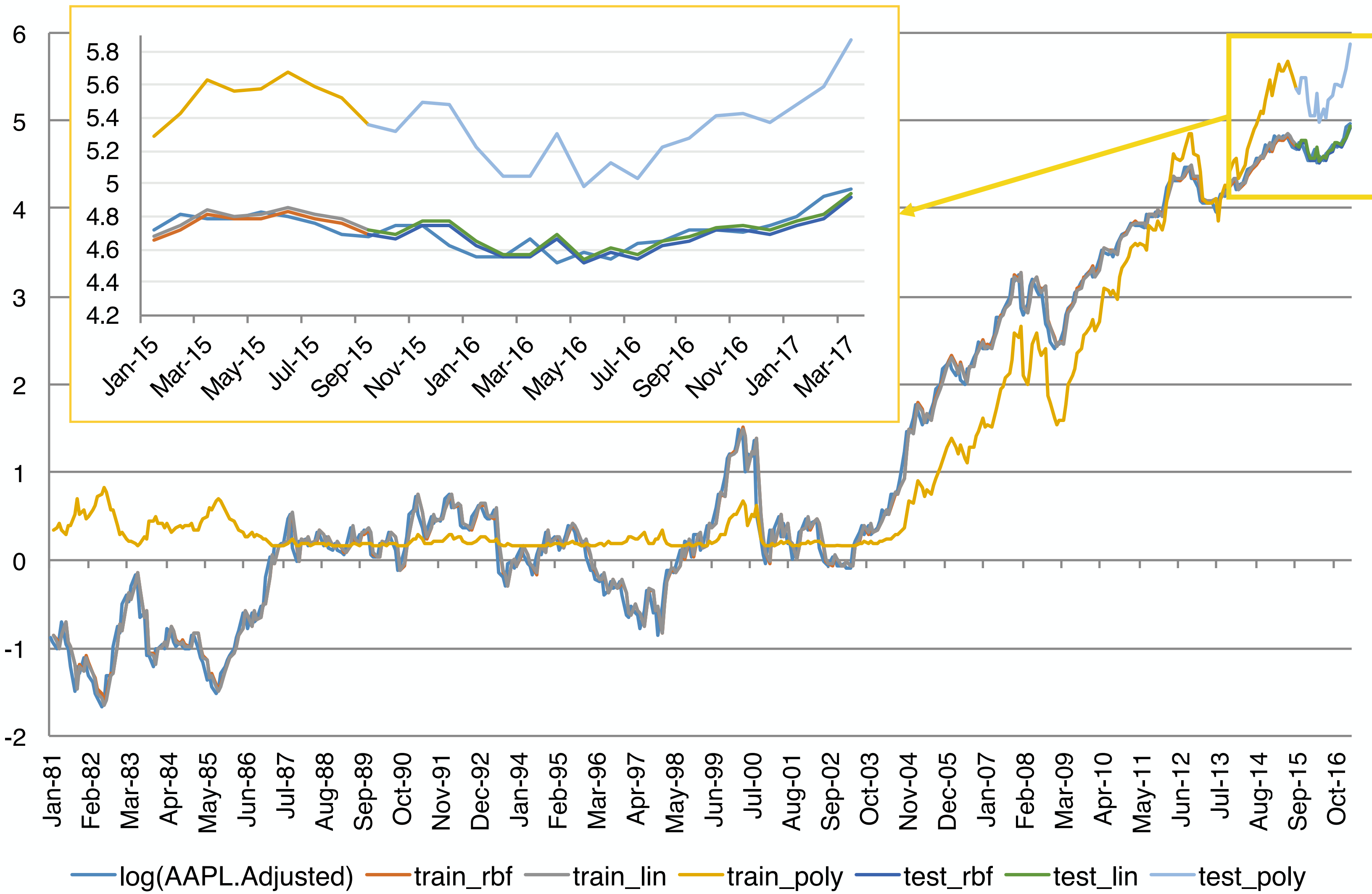
SVM

- Rationale
  - Constructs linear model to implement nonlinear class boundaries through the transforming the inputs into the high-dimensional feature space
- Model
  - $y = b + \sum \alpha_i \, y_i K(\boldsymbol{x}(i), \boldsymbol{x})$
- Our SVM time series model(lag=1)
  - Set y_t as our dependent variable
  - Set y_{t-1} as our independent variable

SVM

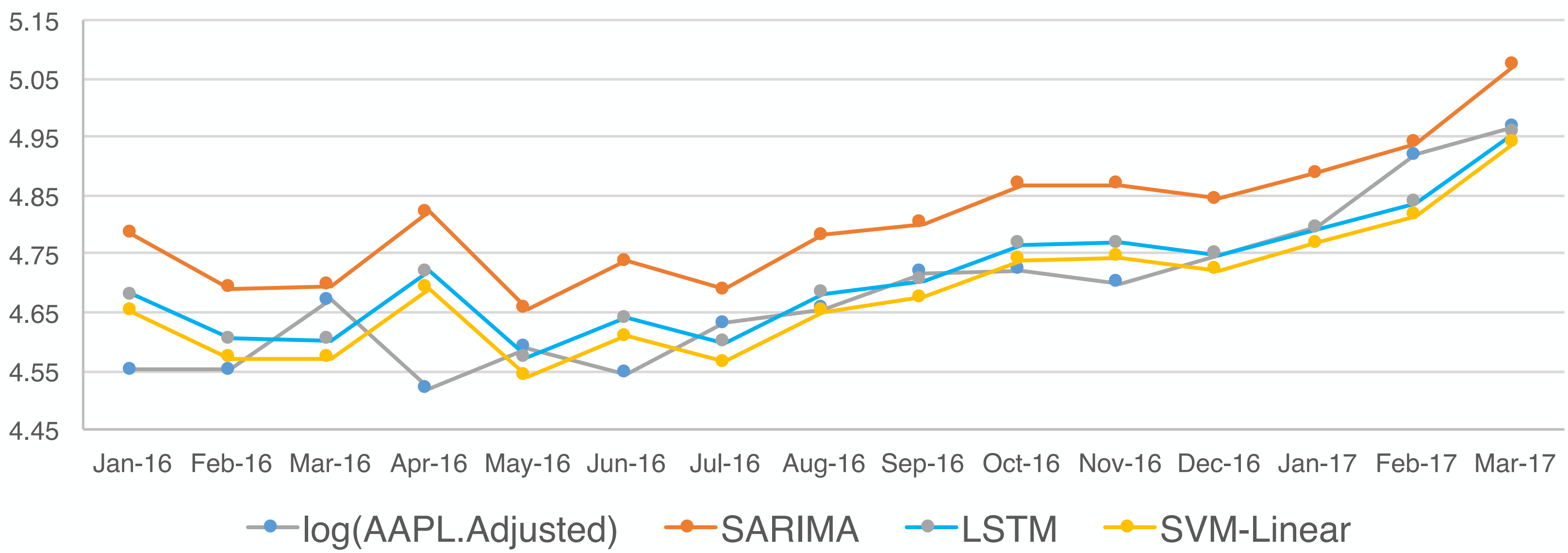
- Kernel methods
  - Map the data into higher dimensional spaces
  - Hope the data could become more easily separated or better structured
- Types
  - Linear Kernel $\mathrm{K}(x, y) = x^T y + c$
  - Polynomial Kernel $\mathrm{K}(x, y) = (\alpha x^T y + c)^{d}$
  - Gaussian Kernel $\mathrm{K}(x, y) = \exp(-\gamma \|x - y\|^2)$

SVM



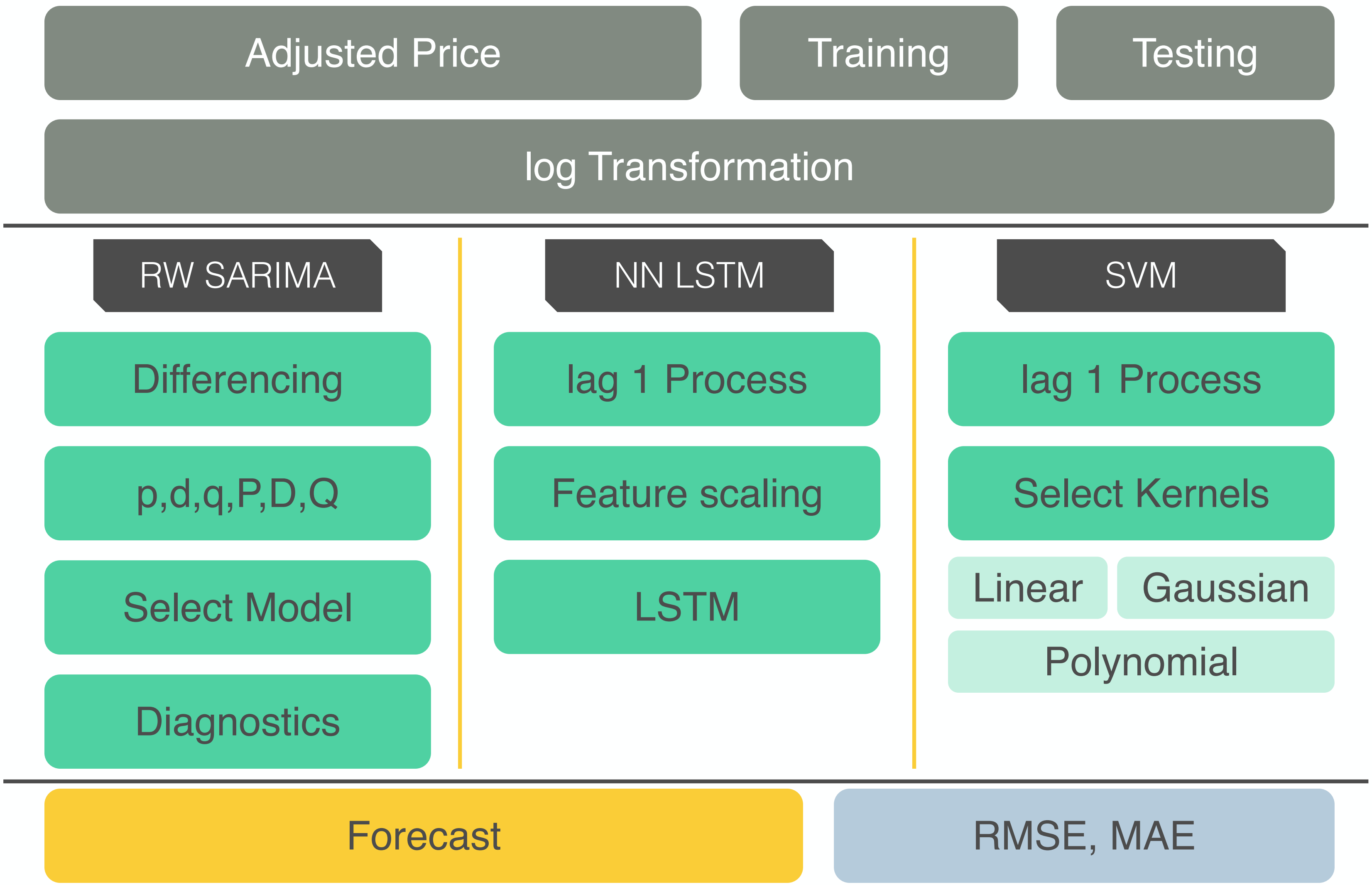
6
5
4
3
2
1
0
-1
-2
5.8
5.6
5.4
5.2
5
4.8
4.6
4.4
4.2
Jan-15 Mar-15 May-15 Jul-15 Sep-15 Nov-15 Jan-16 Mar-16 May-16 Jul-16 Sep-16 Nov-16 Jan-17 Mar-17
Jan-81 Feb-82 Mar-83 Apr-84 May-85 Jun-86 Jul-87 Aug-88 Sep-89 Oct-90 Nov-91 Dec-92 Jan-94 Feb-95 Mar-96 Apr-97 May-98 Jun-99 Jul-00 Aug-01 Sep-02 Oct-03 Nov-04 Dec-05 Jan-07 Feb-08 Mar-09 Apr-10 May-11 Jun-12 Jul-13 Aug-14 Sep-15 Oct-16
log(AAPL.Adjusted) train_rbf train_lin train_poly test_rbf test_lin test_poly

Comparison



	Method	SARIMA	LSTM	SVM: Linear
RMSE	Train	0.1881	0.1520	0.1387
	Test	0.1558	0.0837	0.0733
MAE	Train	0.1505	0.1153	0.1037
	Test	0.1307	0.0654	0.0592

Summary


Adjusted Price
Training
Testing
log Transformation
RW SARIMA
Differencing
p,d,q,P,D,Q
Select Model
Diagnostics
NN LSTM
lag 1 Process
Feature scaling
LSTM
SVM
lag 1 Process
Select Kernels
Linear
Gaussian
Polynomial
Forecast
RMSE, MAE

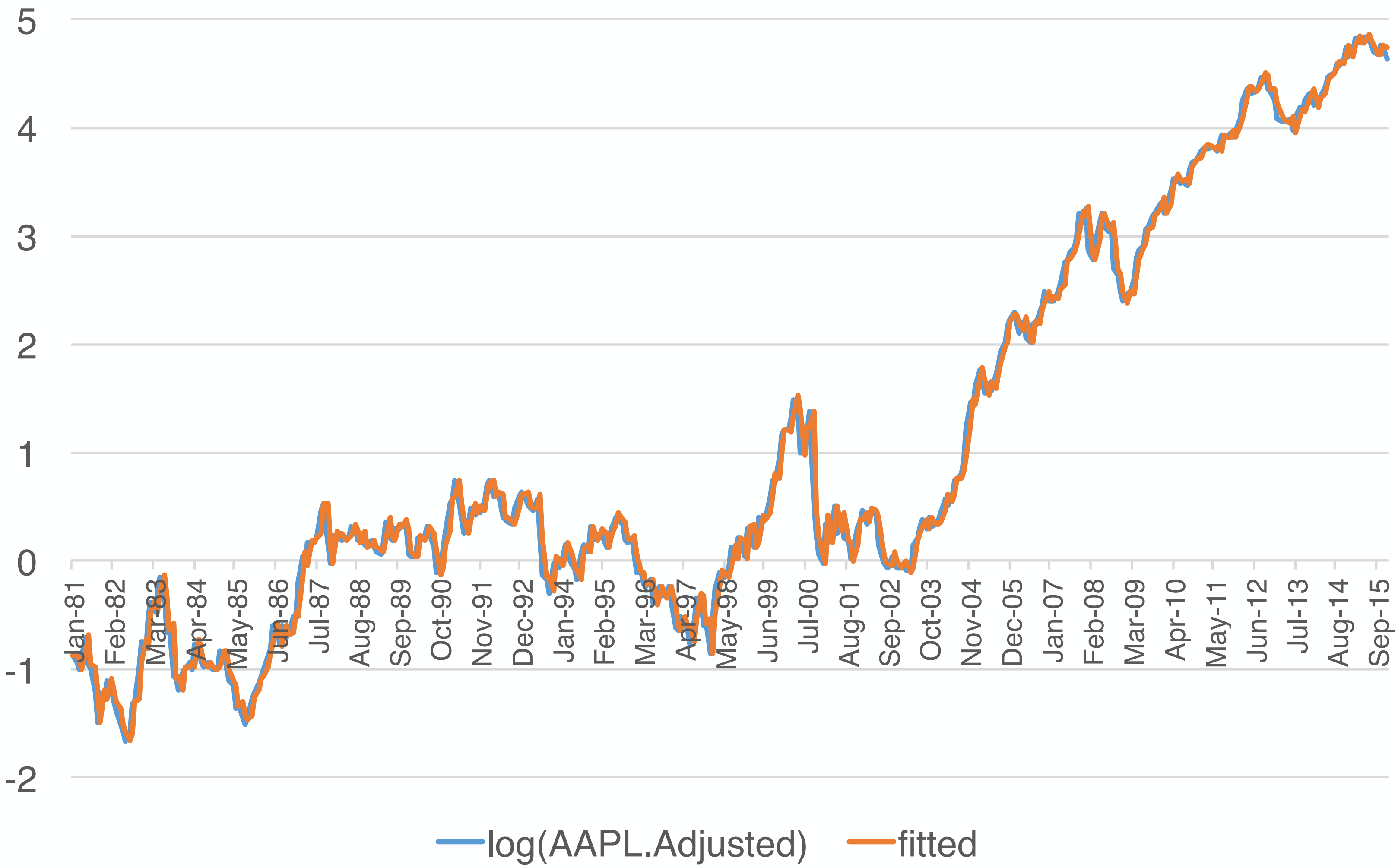
Conclusions

- SVM(linear kernel) has the best performance and SARIMA fitted values are generally higher than the true values, which is the worst among the three
- Neural Networks are easier to perform compared with SARIMA model in their steps
- SARIMA model's coefficients are easier to interpret than Neural Networks.
- In next step we are going to improve our first RW method by considering the ARMA+GARCH model and make comparisons with SVM

Thank You!

Appendix

SARIMA


5
4
3
2
1
0
-1
-2
Jan-81
Feb-82
Mar-83
Apr-84
May-85
Jun-86
Jul-87
Aug-88
Sep-89
Oct-90
Nov-91
Dec-92
Jan-94
Feb-95
Mar-96
Apr-97
May-98
Jun-99
Jul-00
Aug-01
Sep-02
Oct-03
Nov-04
Dec-05
Jan-07
Feb-08
Mar-09
Apr-10
May-11
Jun-12
Jul-13
Aug-14
Sep-15
log(AAPL.Adjusted)
fitted

LSTM

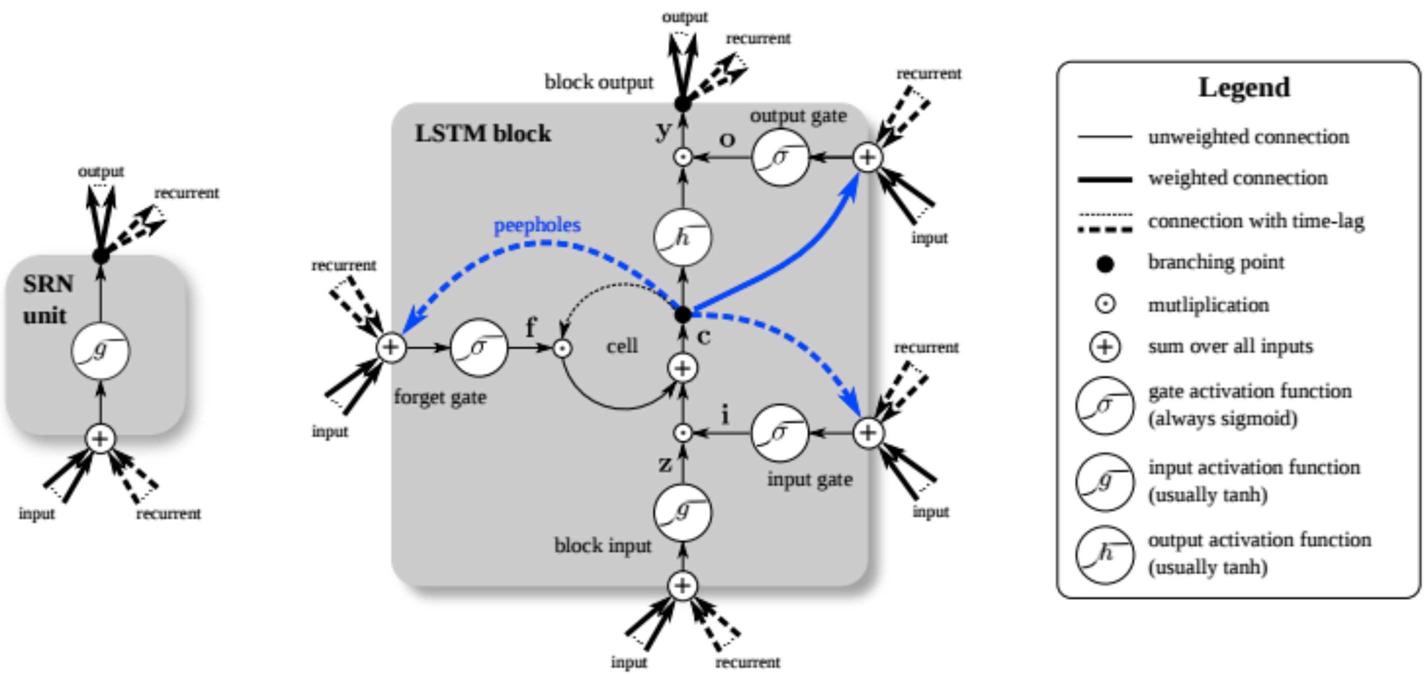


Figure 1. Detailed schematic of the Simple Recurrent Network (SRN) unit (left) and a Long Short-Term Memory block (right) as used in the hidden layers of a recurrent neural network.

SVM

- Advantages:
  - Effective in high dimensional spaces
  - Always avoid the overfitting problem because the maximum hyperplane is relatively stable
  - Versatile: A variety of kernel functions to choose from
- Disadvantages:
  - If the number of features is much greater than the number of samples, the method is likely to give poor performances(not enough data)
  - Can’t interpret parameters, the weighting matrix