Time Series Analysis

Lecture 4

Mixed Autoregressive Moving Average (ARMA) Models

Autoregressive Integrated Moving Average (ARIMA) Models

Seasonal ARIMA (SARIMA) Models

datascience@berkeley

ARMA Models and MA Models

Modeling Using the British Pound–New Zealand Dollar Exchange Rate: Part 1

The Data: Basic Structure

- This is the British pound–New Zealand dollar exchange rate series provided by the book. The series can be obtained from the authors' website.

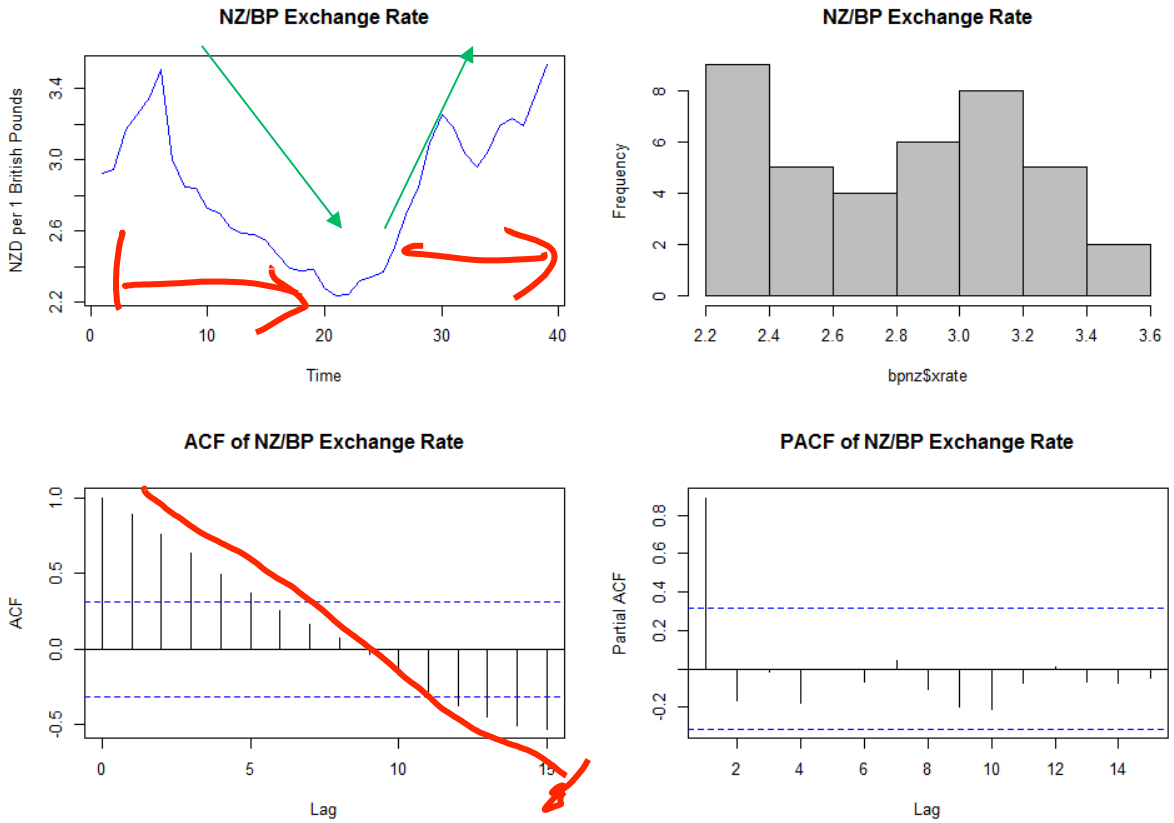
- This is an annual series with 39 observations; the data contains only the series itself (i.e., one variable).

- The summary statistics are displayed below.

```
'data.frame':  39 obs. of  1 variable:
 $ xrate: num  2.92 2.94 3.17 3.25 3.35 ...
> summary(bpnz)
     xrate
 Min.   :2.2
 1st Qu.:2.5
 Median :2.8
 Mean   :2.8
 3rd Qu.:3.2
 Max.   :3.5
```

Descriptive Statistics and Data Visualization

- Similar to the USD–NZD monthly exchange rate series, this series is very persistent, and it cannot be seen using only the histogram and descriptive statistics.
- It does not appear to be captured well by a MA model, although an ARMA model may provide a better model.



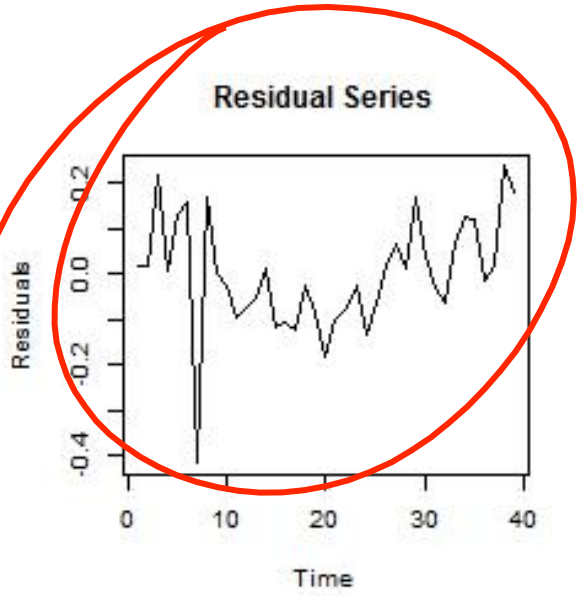
Estimation: MA(5) Model

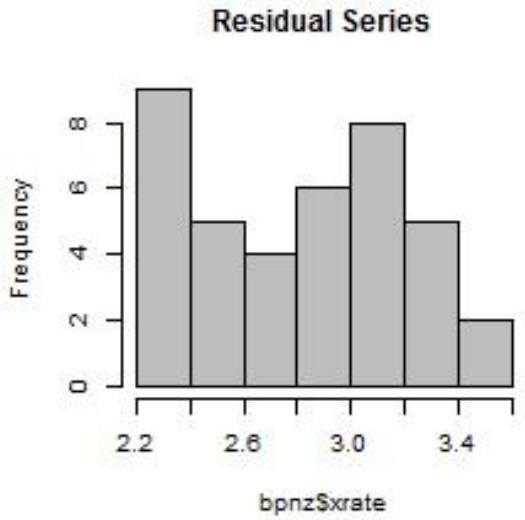
- We will estimate a MA model for comparison.
- The first four MA parameters and the intercept are all significant.
- Note that the last MA parameter is not significant.
- At this point, it is hard to judge how good the estimation is. We will have to examine the residuals, visualizing the in-sample fit and out-of-sample forecast.

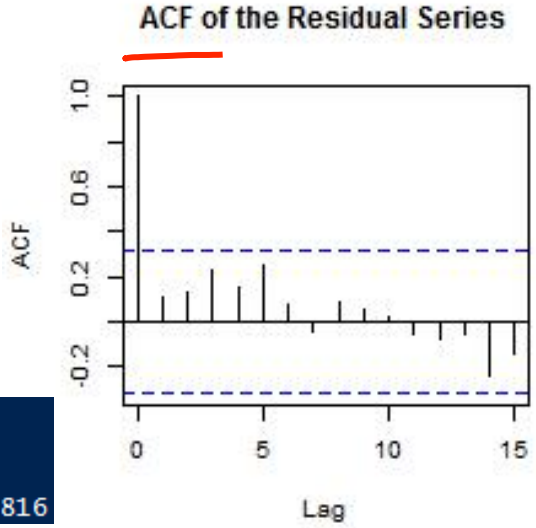
```
Series: bpnz
ARIMA(0,0,5) with non-zero mean

Coefficients:
          ma1   ma2    ma3    ma4    ma5  intercept
      1.5  1.3  1.17  0.64  0.27       2.87
s.e.  0.2  0.3  0.28  0.26  0.17       0.11

sigma^2 estimated as 0.015:  log likelihood=24
AIC=-34   AICc=-31   BIC=-23
```

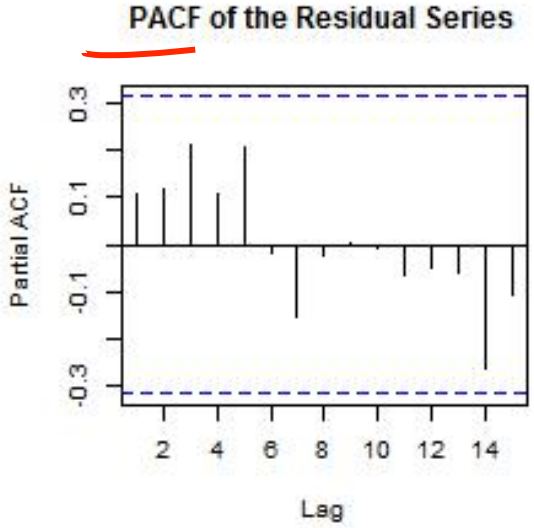
Model Diagnostics

- The time series plot of the residuals does not suggest a white noise sample path.
- However, both the ACF and PACF do not show any statistical significant autocorrelations.
- The null hypothesis that the series is not correlated cannot be rejected based on the Ljung-Box statistic.





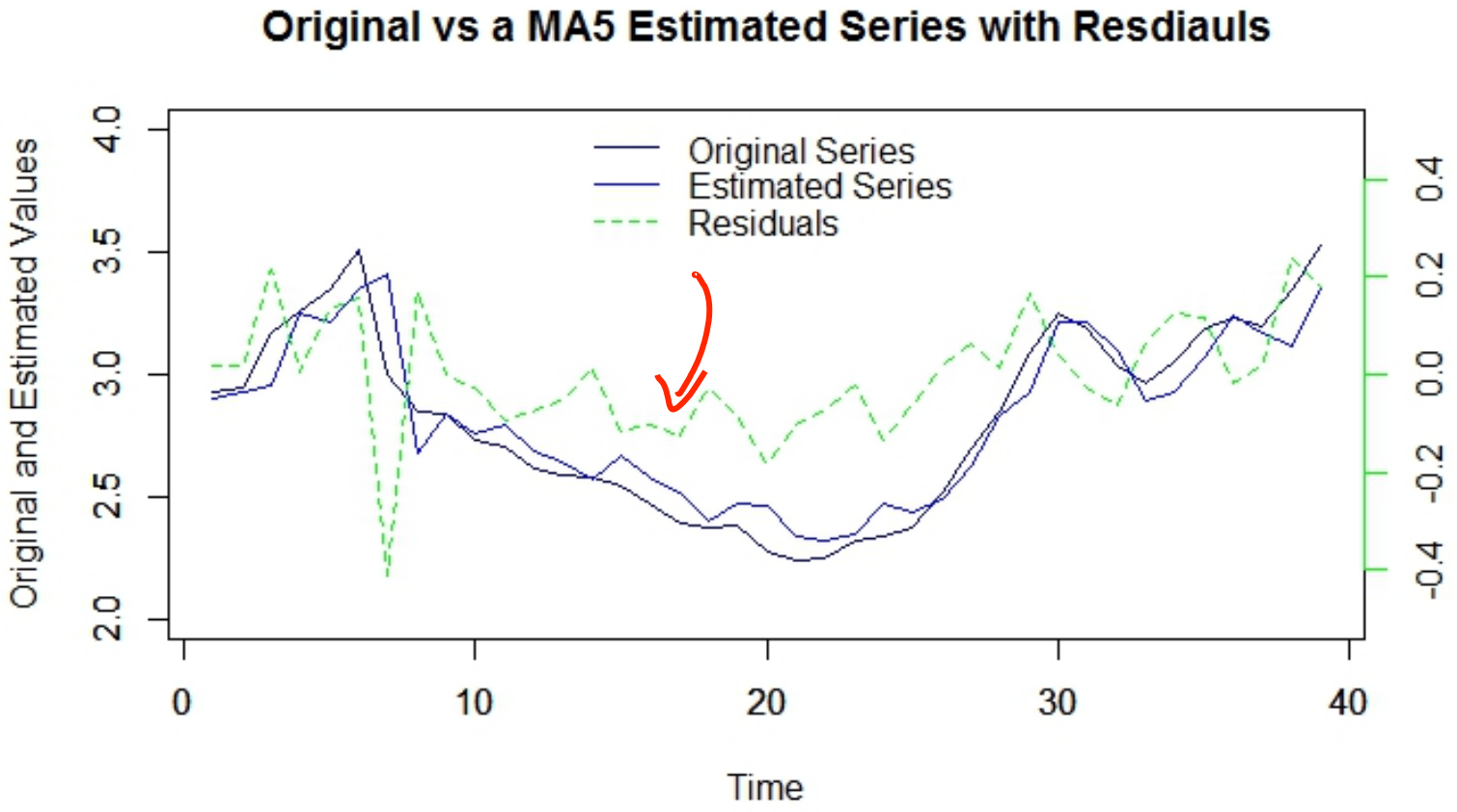





```
 Box-Ljung test

data:  ma5.bpnzfit$resid
X-squared = 0.5, df = 1, p-value = 0.4816
```

Model Performance Evaluation: In-Sample Fit

- The following time series plot displays the original series, which has 39 observations, the estimated values (in blue), overlaid with the residual series (in green and right y-axis).
- In-sample fit "looks reasonable."

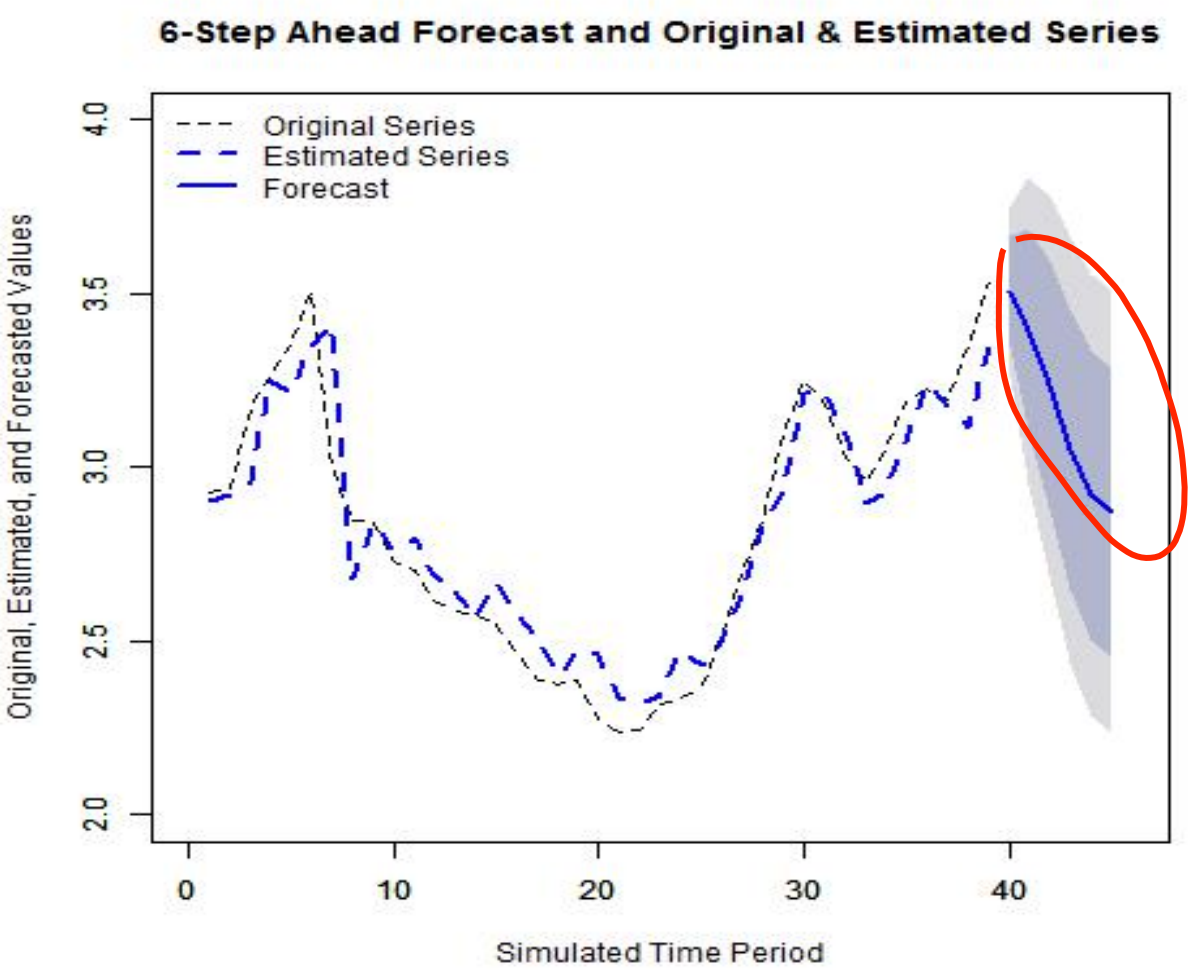


Forecasting

The following plot displays the original series, the estimated values, and a six-step-ahead forecast as well as forecast intervals.

Forecasted Values

	value
1	3.5
2	3.4
3	3.2
4	3.1
5	2.9
6	2.9



Back-Testing and Out-of-Sample Forecasting

- An alternative way to evaluate the model (or the class of model under considerations) is to use out-of-sample forecasting (or back-testing) by leaving out a subsample in the estimation.

- Leaving out a subsample from a time series has to be done very carefully, because we cannot just "randomly" drop observations from the series, as it will break the dependency structure embedded in the series.

- In this example, I excluded the last six observation (i.e., only 33 observations left) from the sample and re-estimate the model using a MA(5).

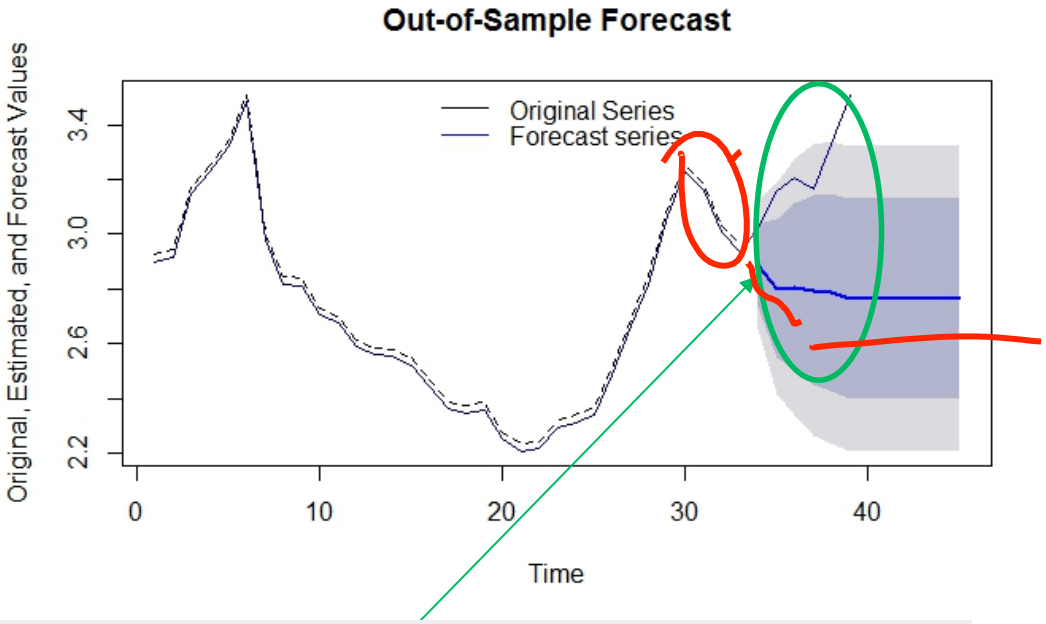
- Then, I produced a 12-step-ahead forecast. Because the last six observations are actually observed, we can compare the forecasts with the actual values. This practice is call back-testing.

Back-Testing and Out-of-Sample Forecasting

- This shows a very dramatic difference from the forecast produced using the entire series.
- The forecast continues with the downward trend observed from the series, although the confidence intervals, with the first few forecasts, do include the observed values.
- The forecast after the first five periods stay flat.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
34	2.9	2.7	3.0	2.7	3.1
35	2.8	2.5	3.1	2.4	3.2
36	2.8	2.5	3.1	2.3	3.3
37	2.8	2.4	3.1	2.3	3.3
38	2.8	2.4	3.1	2.2	3.3
39	2.8	2.4	3.1	2.2	3.3
40	2.8	2.4	3.1	2.2	3.3
41	2.8	2.4	3.1	2.2	3.3
42	2.8	2.4	3.1	2.2	3.3
43	2.8	2.4	3.1	2.2	3.3
44	2.8	2.4	3.1	2.2	3.3
45	2.8	2.4	3.1	2.2	3.3

Note that the slight difference between the forecast values in the graph and in the list is due to rounding.



Divergence between the actual and forecasted values, though the 95% confidence intervals, do include most of the actual values.

Berkeley SCHOOL OF INFORMATION