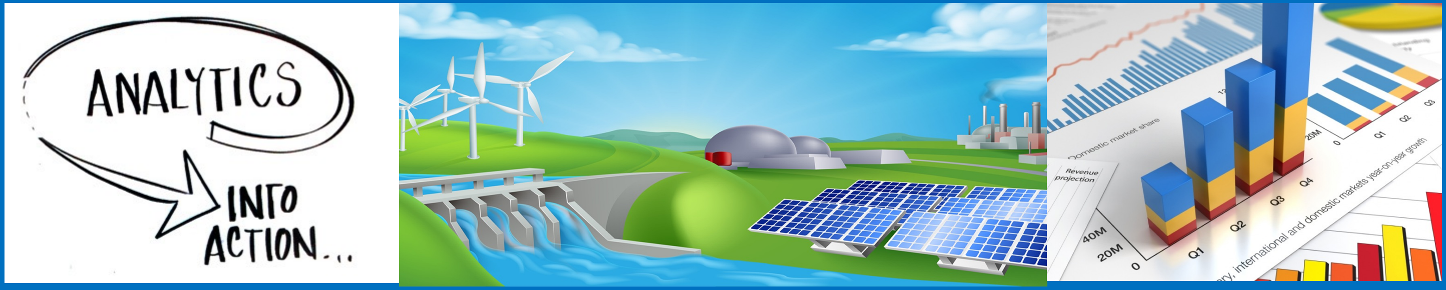


TIME SERIES ANALYSIS FOR ENERGY DATA

M8 - Model Diagnostics, Selection and Performance

Prof. Luana Medeiros Marangon Lima, Ph.D.

Nicholas School of the Environment - Duke University

Learning Goals

- Forecast fit vs forecast error
- Model Selection
  - Residual Analysis
  - AIC, AICc and BIC
- Performance measures
  - MAD, MSE, MAPE

Forecast fit vs forecast error

- Forecast fit
  - Backward-looking assessment
  - Residual Analysis: describes the difference between actual historical data and the fitted values generated by a statistical model
  - How well the model represents historical data
  - Help choose the model that will be further used to forecast unobserved values (Model Selection/Diagnostics)

- Forecast error
  - Forward-looking assessment
  - Difference between actual and forecasted values

Model Selection/Diagnostics

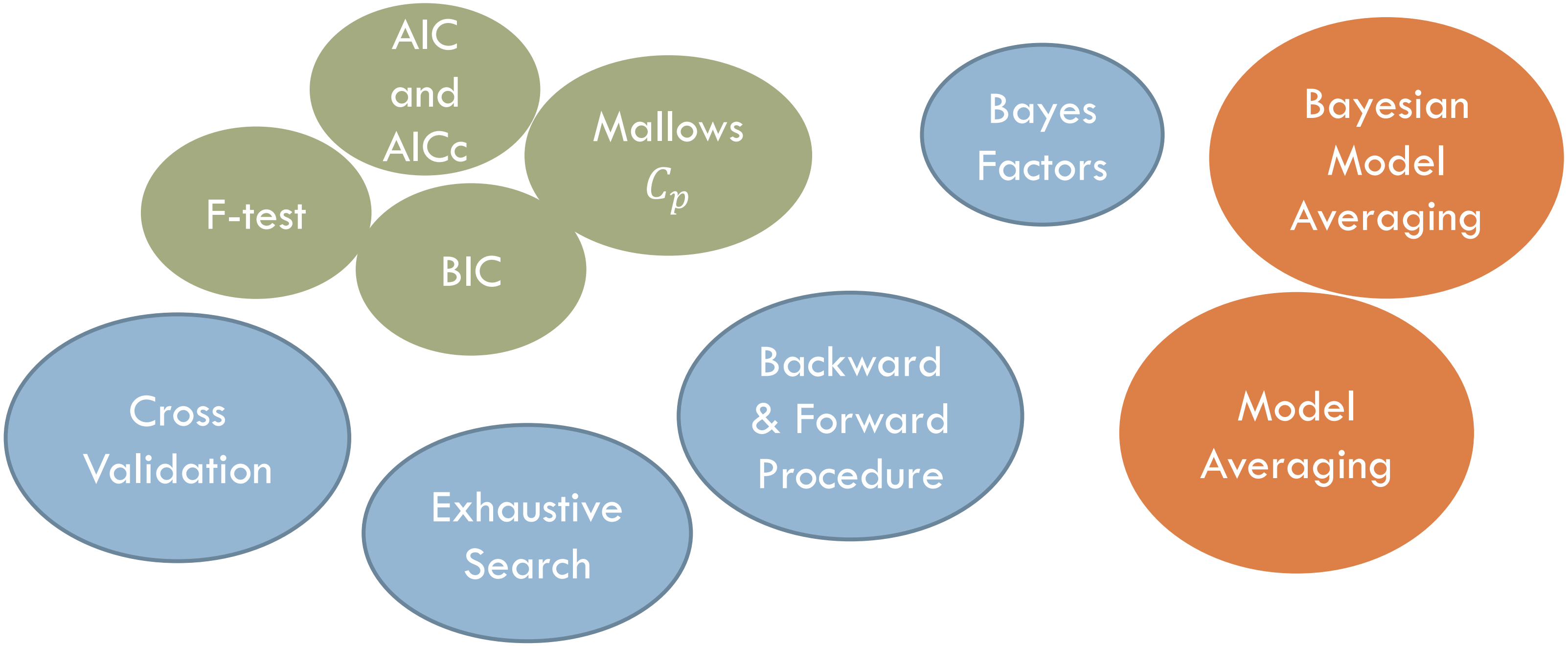
Model Selection

"Unsolved" problem in statistics: there are no magic procedures to get you the "best model" (Kadane and Lazar)

- With a limited number of predictors, it is possible to search all possible models
- But when we have many predictors, it can be difficult to find a good model (many possibilities)
- How do we select models?
  - We need a criteria or benchmark to compare two models
  - We need a search strategy

Model Selection Criteria

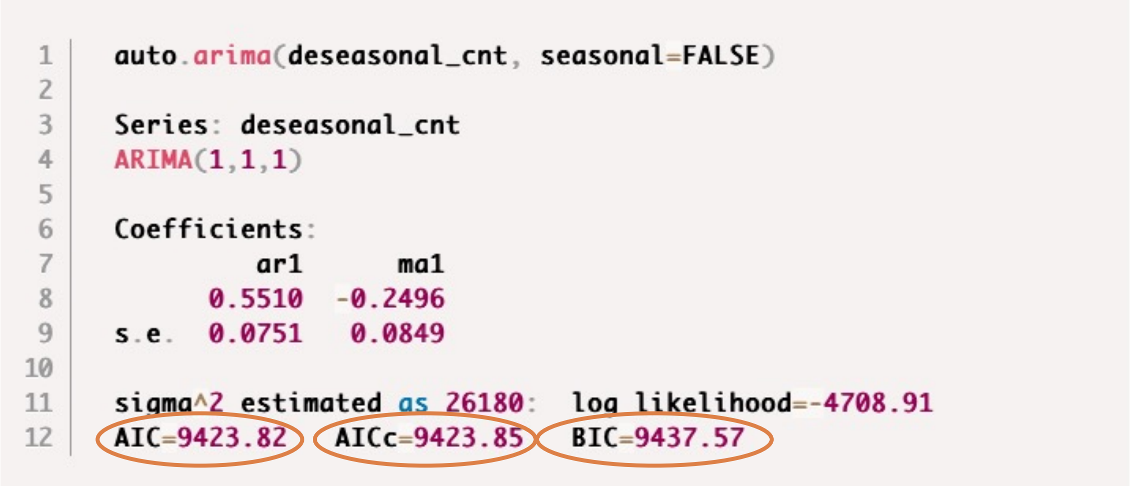
- Some popular and well-known methods



- Some criteria work well for some types of data, others for different data

Model Selection Criteria (cont’d)

- We will focus on the ones that R prints after fitting an ARIMA model



```
auto.arima(deseasonal_cnt, seasonal=FALSE)

Series: deseasonal_cnt
ARIMA(1,1,1)

Coefficients:
         ar1      ma1
      0.5510  -0.2496
s.e.  0.0751   0.0849

sigma^2 estimated as 26180:  log likelihood=-4708.91
AIC=9423.82   AICc=9423.85   BIC=9437.57
```

- And the residual analysis

Akaike Information Criterion (AIC)

- Estimator of the **quality of statistical models**
- Select the model with **lowest AIC**
- Let k be the number of estimated parameters and $\hat{L}$ be the maximum value of the likelihood function

$$AIC = 2k - 2\ln(\hat{L})$$

Penalty for increasing number of parameters

Reward based on the likelihood

- Trade-off between the **goodness of fit** and the **simplicity** of the model
- The AICc is used when sample size (n) is small

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1}$$

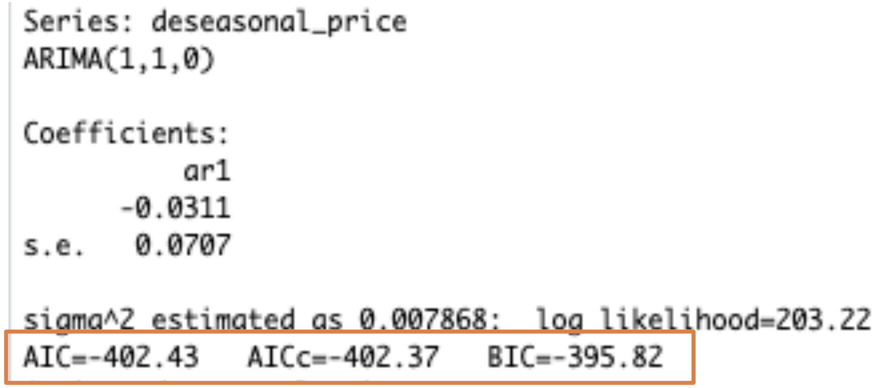
Bayesian Information Criterion (BIC)

- Closely Related to AIC
- Also an estimator of **quality of model**
- Select the model with **lowest BIC**
- Let k be the number of estimated parameters, $\hat{L}$ be the maximum value of the likelihood function and n the number of observations (sample size)

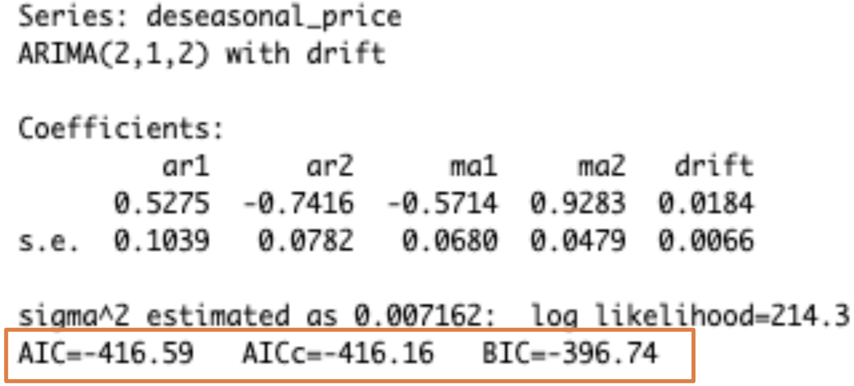
$$BIC = k * \ln(n) - 2\ln(\hat{L})$$

- Sample size should be much larger than number of parameters

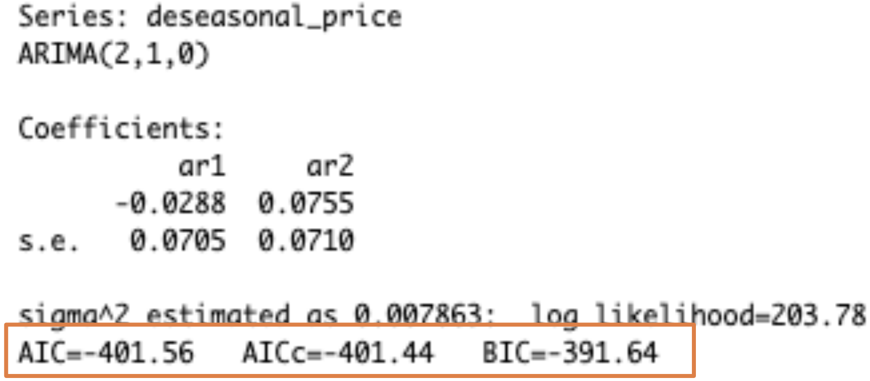
Recall Electricity Prices Example



```
Series: deseasonal_price
ARIMA(1,1,0)

Coefficients:
          ar1
      -0.0311
s.e.   0.0707

sigma^2 estimated as 0.007868:  log likelihood=203.22
AIC=-402.43   AICc=-402.37   BIC=-395.82
```

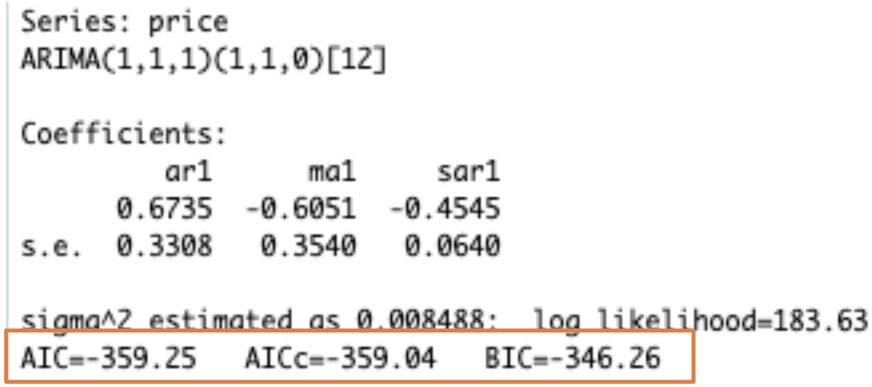


```
Series: deseasonal_price
ARIMA(2,1,2) with drift

Coefficients:
         ar1      ar2      ma1     ma2   drift
      0.5275  -0.7416  -0.5714  0.9283  0.0184
s.e.  0.1039   0.0782   0.0680  0.0479  0.0066

sigma^2 estimated as 0.007162:  log likelihood=214.3
AIC=-416.59   AICc=-416.16   BIC=-396.74
```

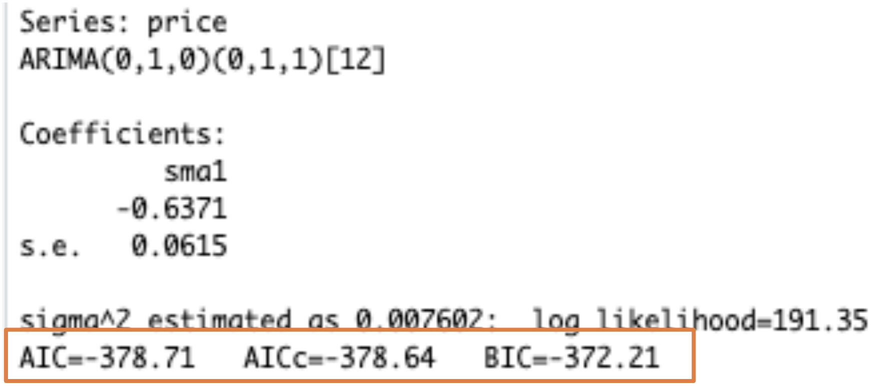


```
Series: deseasonal_price
ARIMA(2,1,0)

Coefficients:
          ar1     ar2
      -0.0288  0.0755
s.e.   0.0705  0.0710

sigma^2 estimated as 0.007863:  log likelihood=203.78
AIC=-401.56   AICc=-401.44   BIC=-391.64
```

Recall Electricity Prices Example



```
Series: price
ARIMA(1,1,1)(1,1,0)[12]

Coefficients:
         ar1      ma1     sar1
      0.6735  -0.6051  -0.4545
s.e.  0.3308   0.3540   0.0640

sigma^2 estimated as 0.008488:  log likelihood=183.63
AIC=-359.25   AICc=-359.04   BIC=-346.26
```



```
Series: price
ARIMA(0,1,0)(0,1,1)[12]

Coefficients:
         sma1
      -0.6371
s.e.   0.0615

sigma^2 estimated as 0.007602:  log likelihood=191.35
AIC=-378.71   AICc=-378.64   BIC=-372.21
```

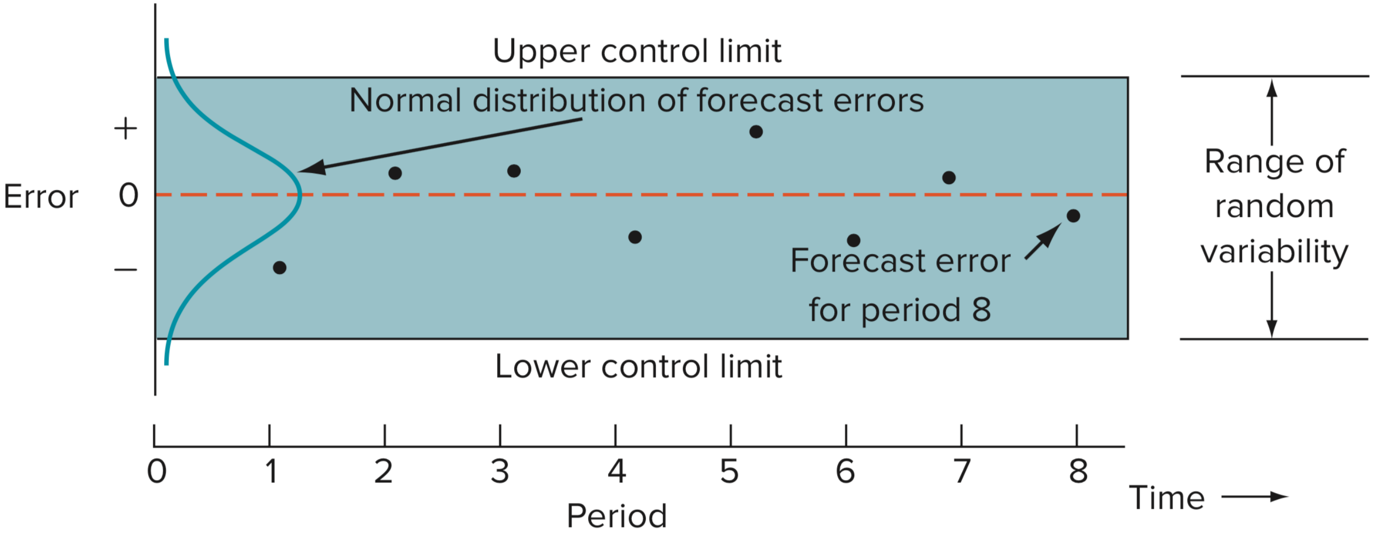
Residual Analysis

Monitoring the Forecast

- **Tracking forecast errors** and analyzing them can provide useful insight into whether forecasts are performing satisfactorily
- **Sources of forecast errors**
  - The model may be inadequate
  - Irregular variations may have occurred
  - The forecasting technique has been incorrectly applied
  - Random variation
- **Residual analysis are useful for identifying the presence of non-random error in forecasts**

Residuals Analysis

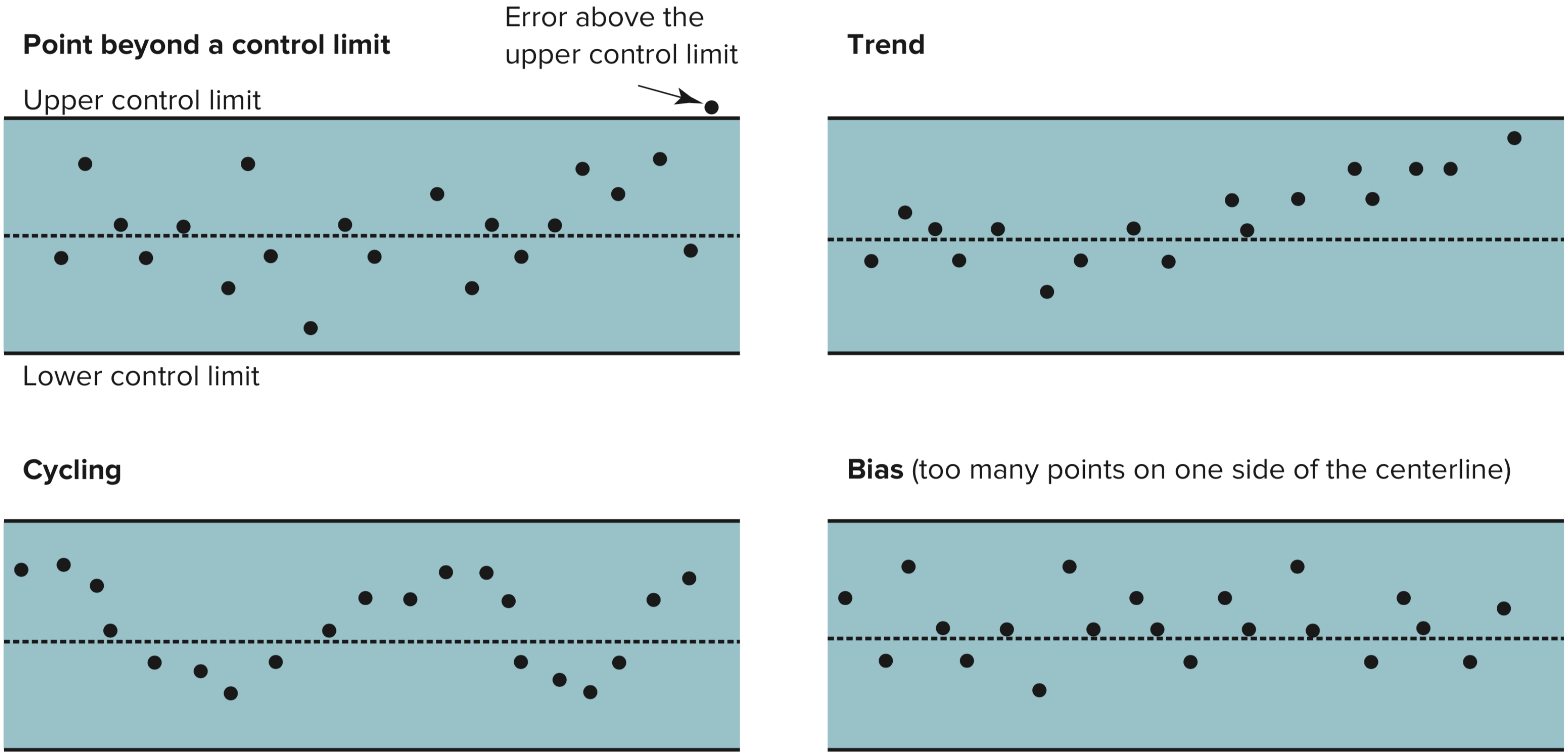
- Errors are plotted on a chart in the order that they occur



- Forecasts are in control when:
  - All errors within control limits
  - No patterns are present (e.g. seasonality, cycles, non-centered data)

Examples of Nonrandomness

FIGURE 3.12 Examples of nonrandomness



Constructing a Control Chart

- Compute the mean square error (MSE)
- The **square root of the MSE** is used in practice as an estimate of the **standard deviation** of the distribution of errors → $s = \sqrt{\text{MSE}}$
- **Errors are random**, therefore, they will be distributed according to a **normal distribution** around a mean of zero
- For a normal distribution:
  - +/- 95.5 % of the values (errors in this case) can be expected to fall within limits of 0 ± 2S (i.e., 0 ± 2 standard deviations)
  - +/- 99.7 % of the values can be expected to fall within ±3s of zero
- Compute the limits as: →

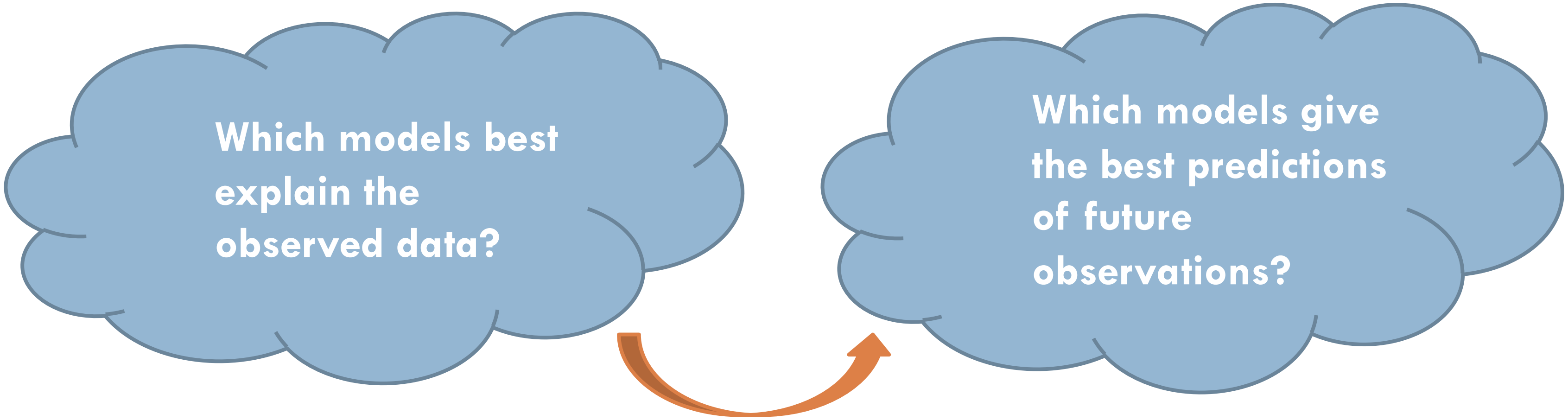
$$\text{UCL: } 0 + z\sqrt{\text{MSE}}$$
$$\text{LCL: } 0 - z\sqrt{\text{MSE}}$$

Number of standard deviations

Model Evaluation/Performance

Model Performance

- ***Keep in mind that these criteria are not measures of predictive power, they just represent how good the model fit the observed data***
- It's possible to look at the predictions from the various models
- In this case we shift the question



Which models best explain the observed data?

Which models give the best predictions of future observations?

Model Performance (cc’ed)

- Model Performance measures the **forecast accuracy**
- Forecasters want to **minimize forecast errors**
  - It is nearly **impossible to correctly forecast real-world variable** values on a regular basis
  - So, it is important to **provide an indication of the extent to which the forecast might deviate** from the value of the variable that actually occurs
- **Forecast accuracy** should be an important forecasting technique selection criterion
  - Error = Actual – Forecast

    Observed value
  - If errors fall beyond acceptable bounds, corrective action may be necessary

Common Performance Measures

- Mean Error (ME)
- **Mean Squared Error (MSE)**
- **Root Mean Squared Error (RMSE) or** Standard Error (SE)
- Coefficient of Determination or R-Squared (R2)
- **Mean Absolute Deviation (MAD)** or Mean Absolute Error (MAE)
- **Mean Absolute Percentage Error (MAPE)**

Forecast Accuracy Metrics

Mean-absolute Deviation

$$\text{MAD} = \frac{\sum \left|\text{Actual}_t - \text{Forecast}_t\right|}{n}$$

MAD weights all errors evenly

Mean-squared Error

$$MSE = \frac{\sum (Actual_t - Forecast_t)^2}{n}$$

MSE weights errors according to their squared values

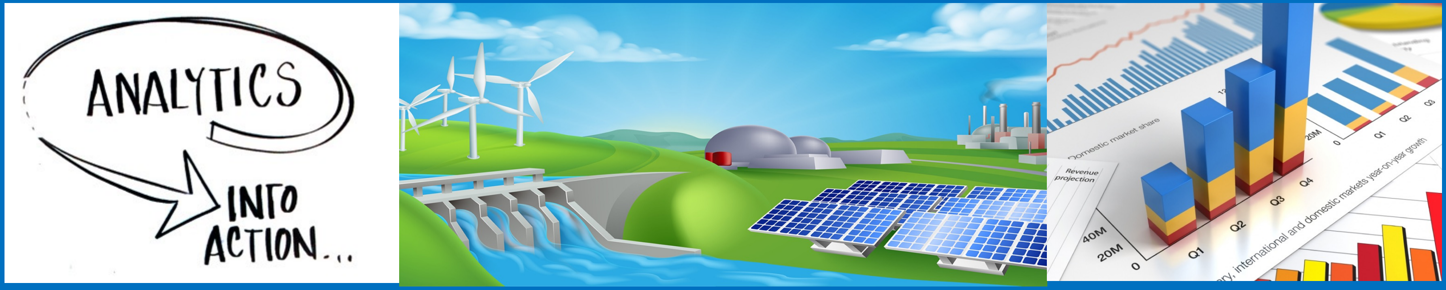
Mean-absolute Percent Error

$$\text{MAPE} = \frac{\sum \frac{\left|\text{Actual}_t - \text{Forecast}_t\right|}{\text{Actual}_t} \times 100}{n}$$

MAPE weights errors according to relative error

Forecast Error Calculation

<table>
<tr><th>Period</th><th>Actual (A)</th><th>Forecast (F)</th><th>(A-F) Error</th><th>|Error|</th><th>Error^2</th><th>[|Error|/Actual]x100</th></tr>
<tr><td>1</td><td>107</td><td>110</td><td>-3</td><td>3</td><td>9</td><td>2.80%</td></tr>
<tr><td>2</td><td>125</td><td>121</td><td>4</td><td>4</td><td>16</td><td>3.20%</td></tr>
<tr><td>3</td><td>115</td><td>112</td><td>3</td><td>3</td><td>9</td><td>2.61%</td></tr>
<tr><td>4</td><td>118</td><td>120</td><td>-2</td><td>2</td><td>4</td><td>1.69%</td></tr>
<tr><td>5</td><td>108</td><td>109</td><td>1</td><td>1</td><td>1</td><td>0.93%</td></tr>
<tr><td></td><td></td><td></td><td>Sum</td><td>13</td><td>39</td><td>11.23%</td></tr>
<tr><td></td><td></td><td></td><td></td><td>$n = 5$</td><td>$n = 5$</td><td>$n = 5$</td></tr>
<tr><td></td><td></td><td></td><td></td><td>MAD</td><td>MSE</td><td>MAPE</td></tr>
<tr><td></td><td></td><td></td><td></td><td>= 2.6</td><td>= 7.8</td><td>= 2.25%</td></tr>
</table>



THANK YOU !

luana.marangon.lima@duke.edu