Vehicle Dynamics and Simulation

Parameter Tuning

Dr Byron Mason

Loughborough University

Overview

- Motivation
- How good is the model
  - Plots
  - Metrics
- Optimisation
- Parameter optimisation in MATLAB

Motivation

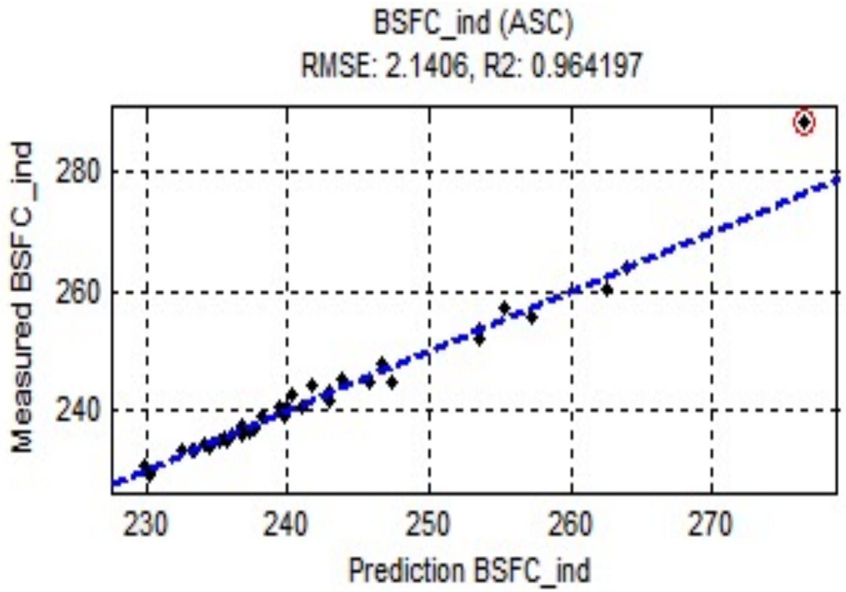
- Physical models have parameters that physically represent chacteristics of the system;
  - Length
  - Weight
  - Volume
  - Density
- To best represent behaviour of the system using the model it is sometimes necessary to change the value of model parameters.
- It is (almost) impossible to perfectly represent the modelled system by comparison with data since measurement is corrupted by noise;

$$y = f(x)$$
$$y' = f(x) + e$$

- y is the real system, y' is what we observe i.e. system + measurement noise

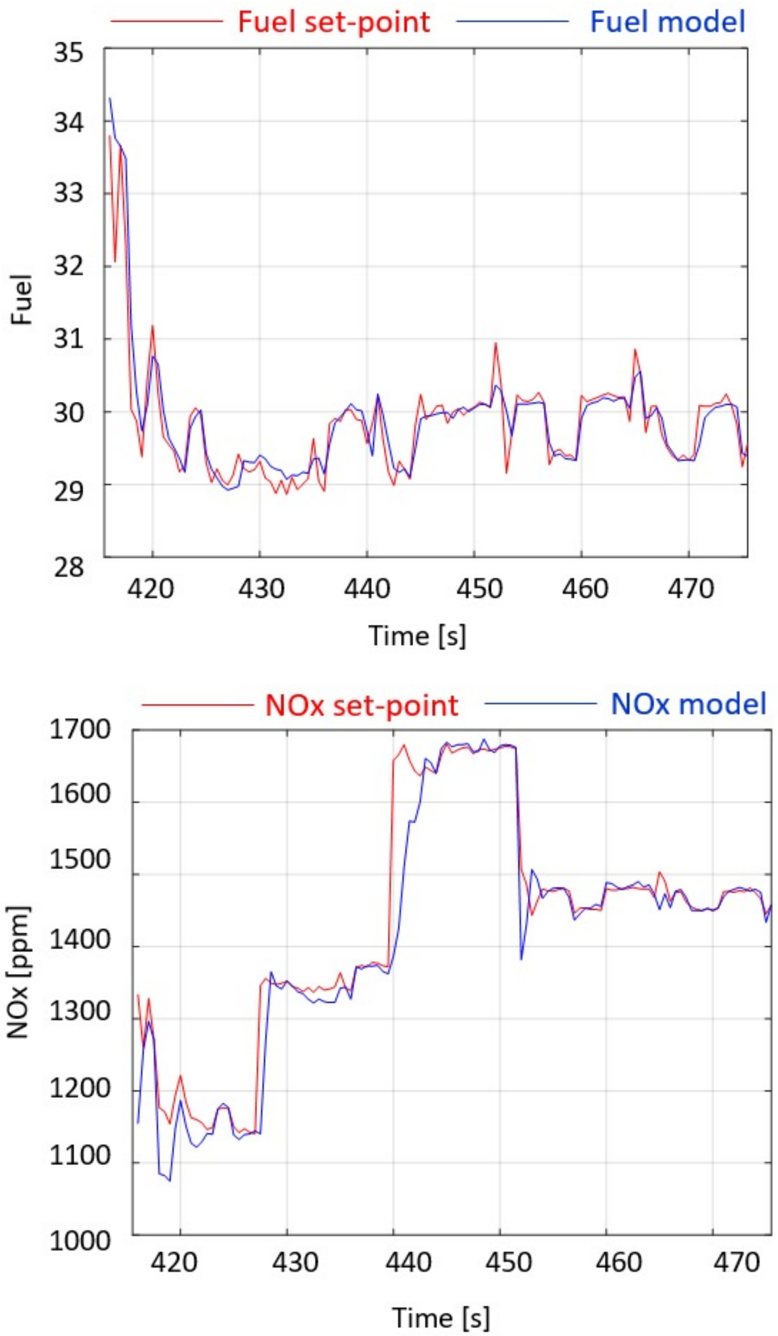
Prediction vs Data

- The difference between a model and the real system that it describes can be shown in a number of ways;
  - Time series plot
  - Correlation plot
  - Residuals plot
- Each of these gives a slightly different view of the models ability to represent the real system



Timeseries plot

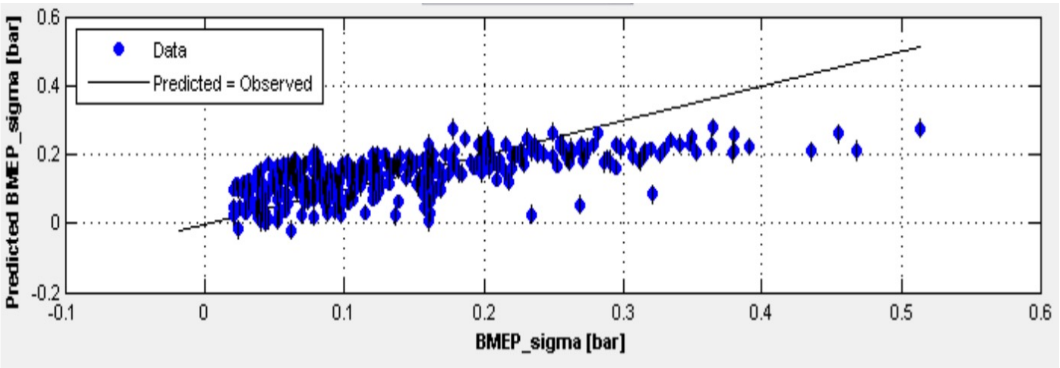
- Perhaps the most intuitive and common approach to evaluating the difference between model and reality.
- Easy to see where (in time) the model is struggling to explain the system behaviour.
- It can be misleading since many small differences can sum to large cumulative difference. Whether this is important or not depends on the model use.

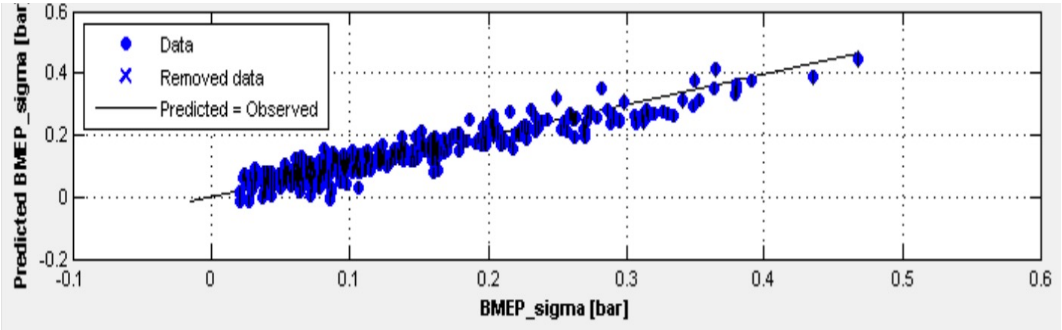


ough

Correlation plot

- Bad correlation. High model error. For high BMEP_sigma the predicted is lower than measured BMEP_sigma
- Good correlation between model and measured data.

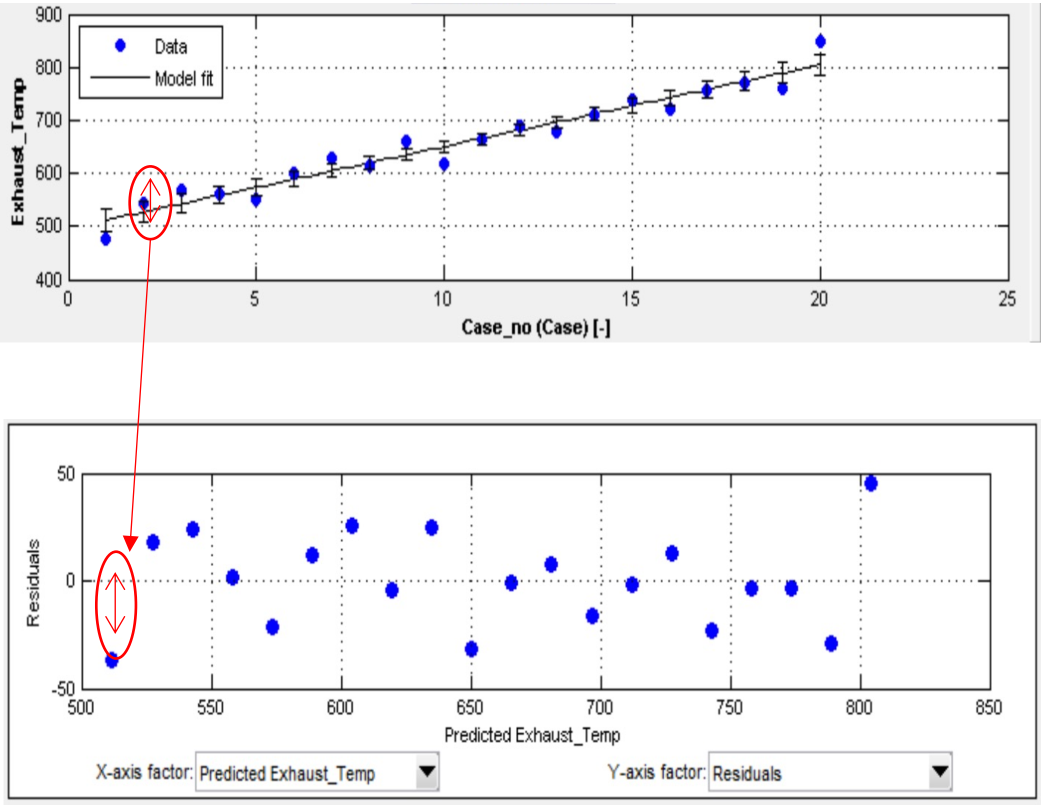




Loughborough University

Residuals plot

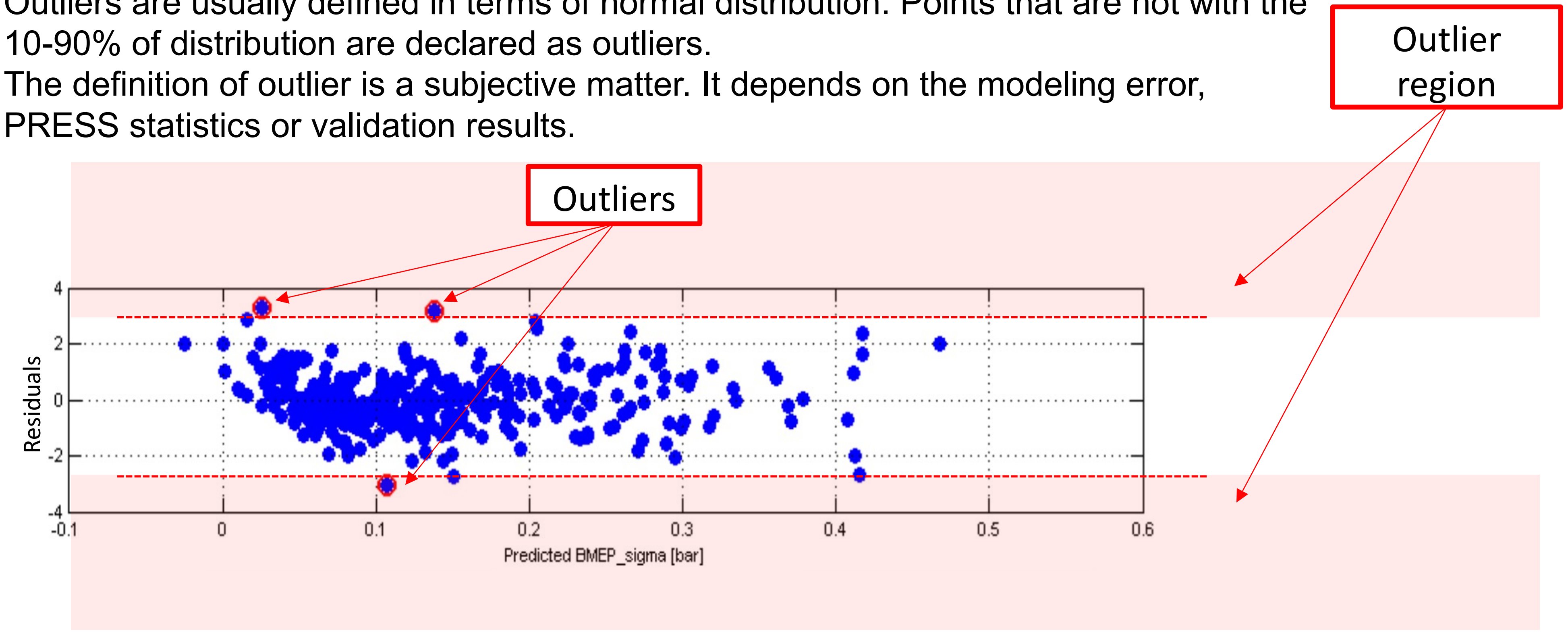
- A correlation plot between the measured dataset and model fit is plotted.
- Residuals are simply what is left after the model has been fitted - the unexplained variation
- **Residual = data – fit(model)**
- Any structure in the residuals is indicative of a model that isn't fully explaining the data.



Loughborough University

Outlier plot

Outliers are usually defined in terms of normal distribution. Points that are not with the 10-90% of distribution are declared as outliers.
The definition of outlier is a subjective matter. It depends on the modeling error, PRESS statistics or validation results.



Metrics

- In addition to the obvious visual inspection of model performance a plethora of metrics exist to quantify the performance of the model.
- Each has a different purpose.
- The most common are;

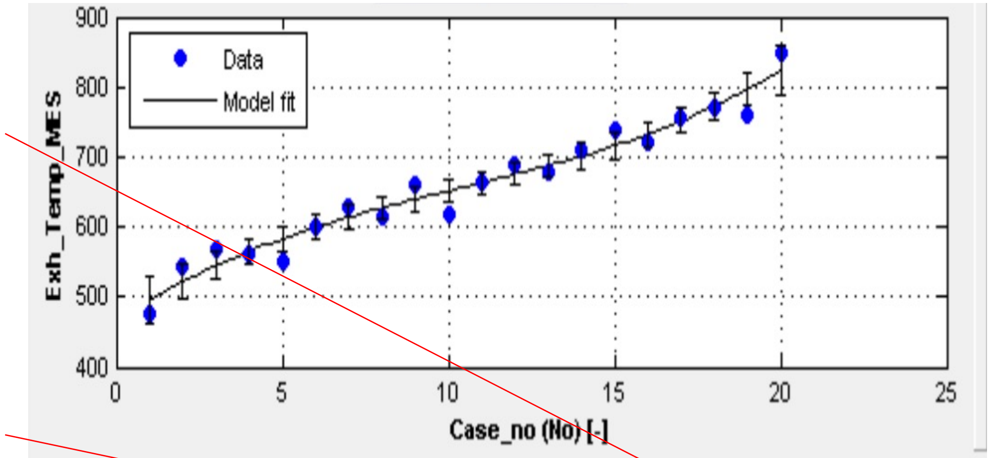
Model evaluation - Error - RMSE

- RMSE is the root square of the mean(norm) of all squares of error

- Model with smallest RMSE is usually selected. But RMSE alone doesn't guarantee model fitness because that is training RMSE.

- Typically, RMSE (between model and validation data) or PRESS RMSE are used as an indicator for model selection.

For example, an exhaust temp dataset was modeled using 4 types of model: linear, quadratic, cubic and 4th order polynomial. The table below shows the difference of RMSE between the models. The best model is selected based on RMSE.



Name	Observations	Parameters	Box-Cox	PRESS RMSE	RMSE
Linear	20	2	1	24.668	22.532
Quadratic	20	3	1	28.661	23.169
Cubic	20	4	1	27.745	21.58
Poly_4	20	5	1	35.341	22.278

Loughborough University

Model evaluation - Error - R^2

- R^2 (coefficient of determination) gives a measure of the goodness of fit by comparing the explained component of the data with the unexplained.

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}}$$

Unexplained variation

Total variation

Model evaluation - Error – PRESS RMSE

- The PRESS statistic gives a good indication of the predictive power of your model
- PRESS statistics is a method to validate a model without having to use validation data.
- The same training data is use to validate the model.
- Iteratively, a fraction of training data use as validation data and the rest use as training data. The cycle is iterated until all fraction of training data have been once used for validation data. The RMSE are calculated for each iteration. Finally, the average of RMSE is calculated.
- PRESS R^2 can be calculated using the same procedure

PRESS RMSE = SUM(RMSE)/k

Method

- Change to parameters can be done in several ways
  - Trial and error approaches (for simple models)
  - Direct optimisation (for low parameter numbers and computationally efficient models)
  - Through a meta-model (high order computationally demanding model)

Optimisation

- Optimisation is the process of minimising a cost function
- In an optimisation process a cost function is formulated and minimised.
- The Cost function contains quantities to be minimised.

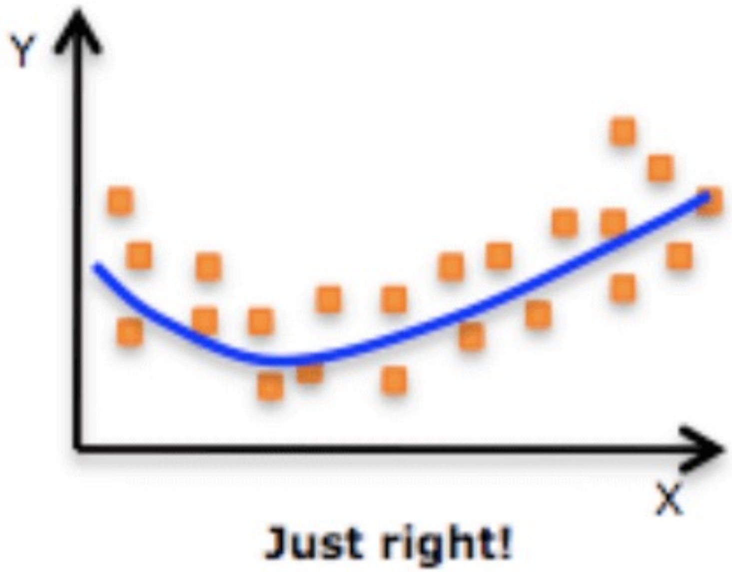
$$\min_{p \in P} J(p) \text{ with } J = \sum_i \alpha_i f_i(p)$$

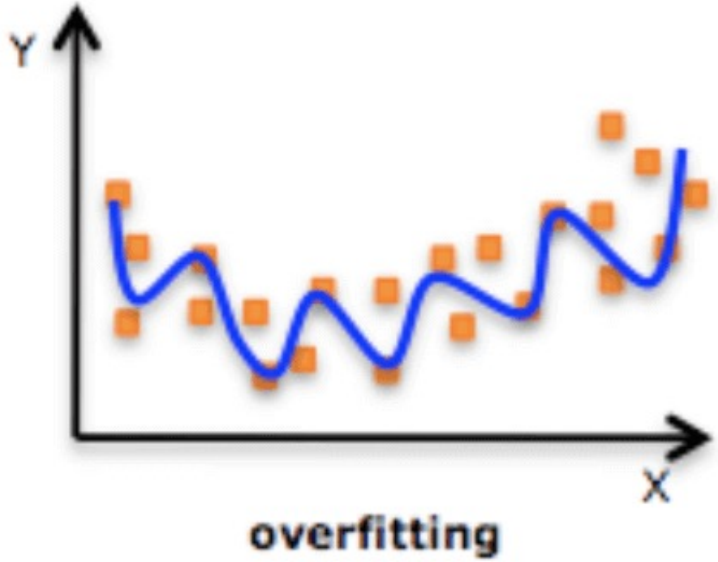
Optimisation

- The cost function can be changed to emphasise particular types of error
  - Sum of error penalises residuals based on linear distance
  - sum of squared error penalises outliers more
- Within a parameter optimisation activity you would choose a fit metric as the cost function e.g. r^2 or RMSE, etc.
- Care should be take in selecting the right metric i.e. what is the modelling objective.

Optimisation

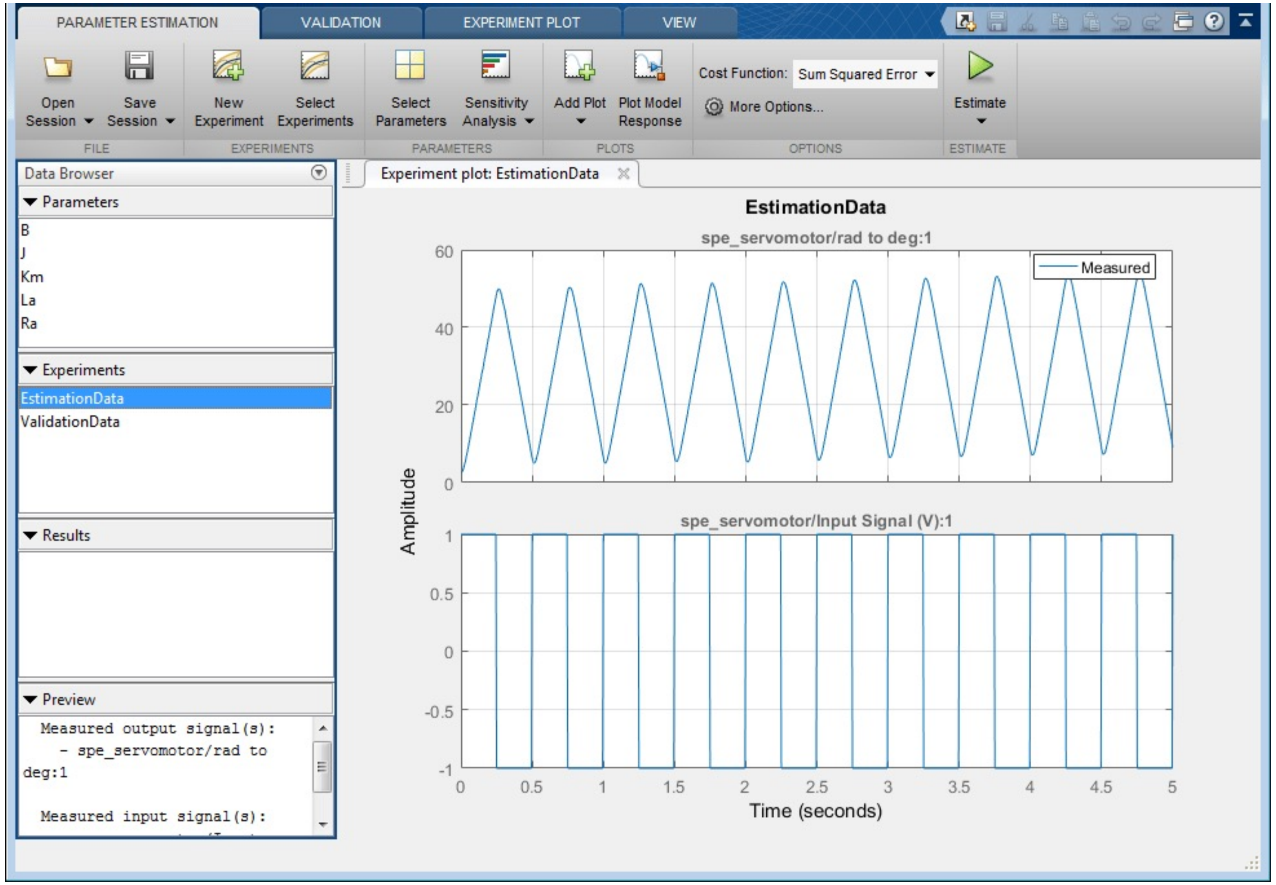
- Care should be taken to avoid overfitting.
  - This is not really an issue with physical models.





Parameter Optimisation in MATLAB

- MATLAB/Simulink has a model parameter tuning GUI.
- Leverages MATLAB optimisation algorithm to minimise a cost function.
- The cost function and optimiser used can be changed through the GUI



Loughborough University