UNIVERSITY OF TORONTO

Department of Chemical Engineering

CHE1148 – Process data analytics

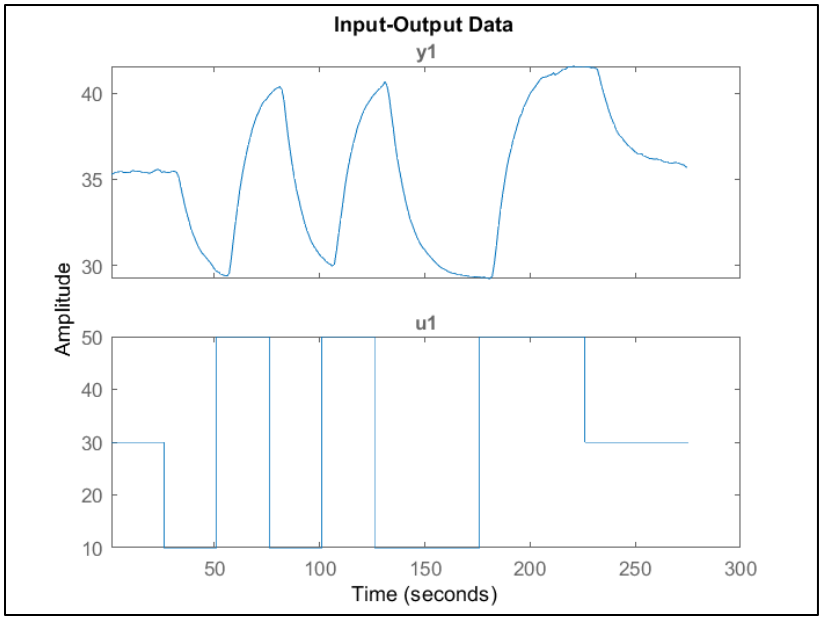
Assignment 3: Estimation of a dynamic model from process data

Student name: Ashutosh Desai
Student Number: 1004774041

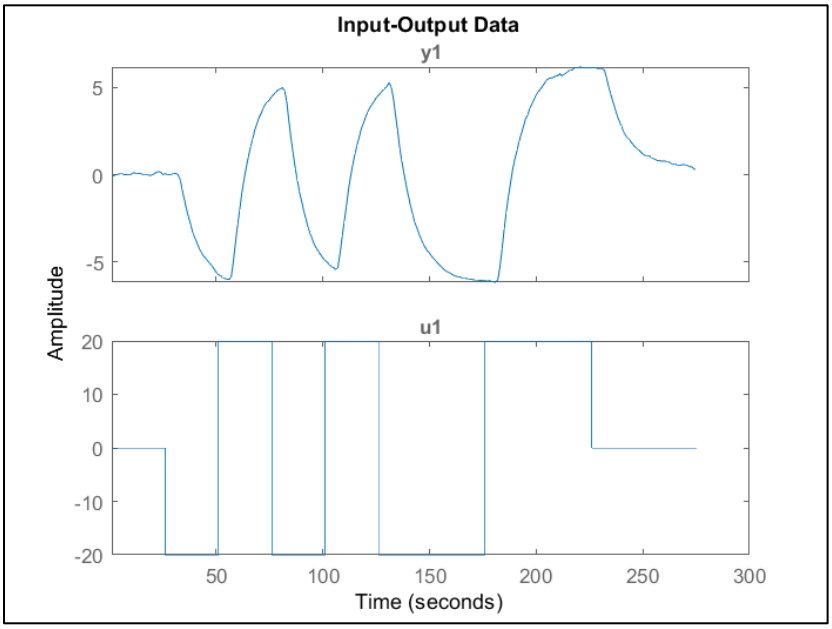
- MATLAB CODE:

```
load STHID-data.txt
cd=detrend(STHID_data, 'constant');  %cd variable stores the zero-mean centered data
figure;idplot(cd)
figure;
[ir, r, cl]=cra(cd);  %this function returns the impulse response along with the confidence line
nn=struc(2,2,4:8);
v1=arxstruc(cd,cd,nn);  %attempt 1 to evaluate the best parameters
v1
nn=struc(1:3,1:3,6);
v2=arxstruc(cd,cd,nn);  %attempt 2 to evaluate the best parameters
v2
for i=1:9
   parameter_plot(i)=v2(2,i)+v2(3,i);
end
figure;
plot(parameter_plot(1:9),v2(1,1:9),'+');
mod=arx(cd,[2,3,6]);  %mod is the ARX model used based on the finalised parameters
present(mod)
figure;
resid(cd,mod,'corr')  %this is for residual correlation analysis

comp_data=compare(cd,mod,inf);
figure;
plot(cd(:,1));
hold;
plot(comp_data,'r');   %superimposed plot
figure; plot(cd(:,1),comp_data,'*');   %direct comparison plot
hold;
plot([-8 8],[-8 8],'k');
```

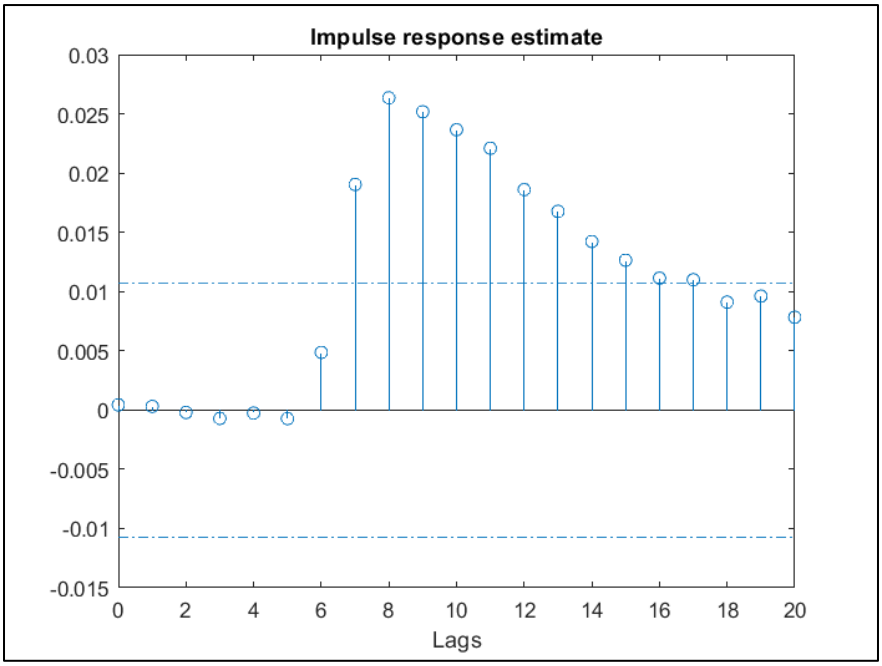
a.) Plotting the input-output data



b.) Plotting the zero-mean centered data



- The output of the ‘cra’ function is shown below



The image shown above shows that the order of A and B is 2 or higher. It also shows that the impulse response above the confidence line is after lag 6. This is one of the steps to estimate the order for a model.

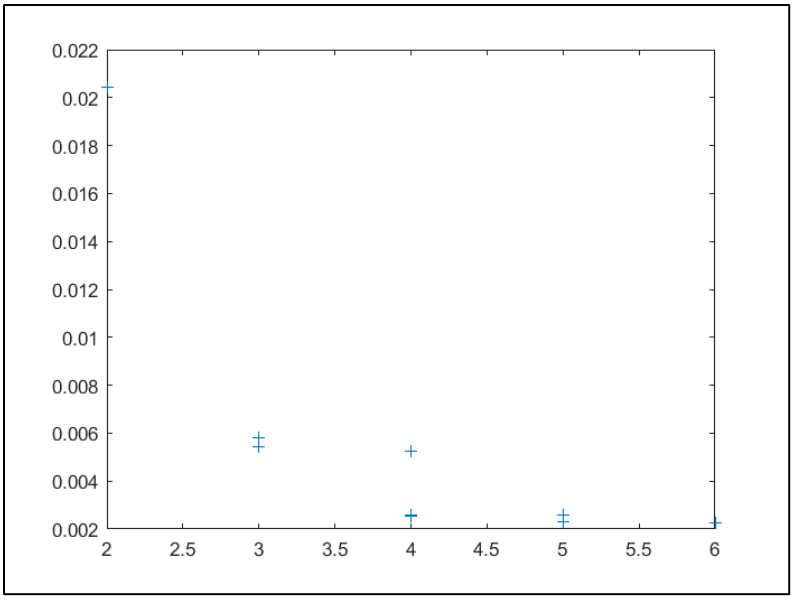
- The output shown below helps to estimate the ‘order’ and the ‘time delay’

Row values for both the tables shown below:

Row 1 – Loss function
Row 2 – Order of A
Row 3 – Order of B
Row 4 - Delay

```
v1 =

    0.0086    0.0054    0.0026    0.0033    0.0111  275.0000
    2.0000    2.0000    2.0000    2.0000    2.0000   15.3290
    2.0000    2.0000    2.0000    2.0000    2.0000         0
    4.0000    5.0000    6.0000    7.0000    8.0000         0
```

```
v2 =

    0.0204    0.0058    0.0026    0.0054    0.0026    0.0023    0.0052    0.0026    0.0022  275.0000
    1.0000    1.0000    1.0000    2.0000    2.0000    2.0000    3.0000    3.0000    3.0000   15.3289
    1.0000    2.0000    3.0000    1.0000    2.0000    3.0000    1.0000    2.0000    3.0000         0
    6.0000    6.0000    6.0000    6.0000    6.0000    6.0000    6.0000    6.0000    6.0000         0
```



The image shown above illustrates the best sum of orders of A and B, extracted from table v2, which correspond to the lowest loss function. Our objective is to minimize the loss function and select an order of the model that is not overfitting. Hence, we go ahead with,

Order of A = 2; Order of B = 3; Delay = 6

- The results after the ARX model was used based on the parameters finalised above.

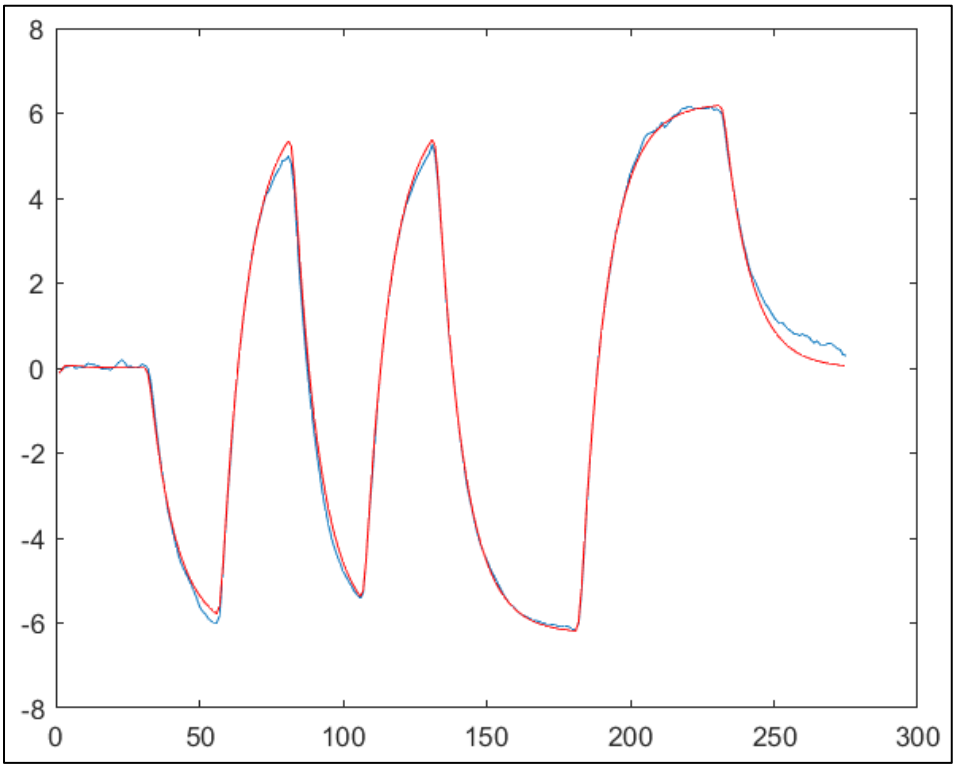
```
mod =
Discrete-time ARX model: A(z)y(t) = B(z)u(t) + e(t)
  A(z) = 1 - 1.112 (+/- 0.03848) z^-1 + 0.1941 (+/- 0.03498) z^-2


  B(z) = 0.00569 (+/- 0.0005298) z^-6 + 0.01349 (+/- 0.0007634) z^-7 + 0.006311 (+/- 0.001072) z^-8

Sample time: 1 seconds

Parameterization:
   Polynomial orders:   na=2   nb=3   nk=6
   Number of free coefficients: 5
   Use "polydata", "getpvec", "getcov" for parameters and their uncertainties.

Status:
Estimated using ARX on time domain data.
Fit to estimation data: 98.78% (prediction focus)
FPE: 0.002428, MSE: 0.00229
More information in model's "Report" property.
Current plot held
Current plot held
```

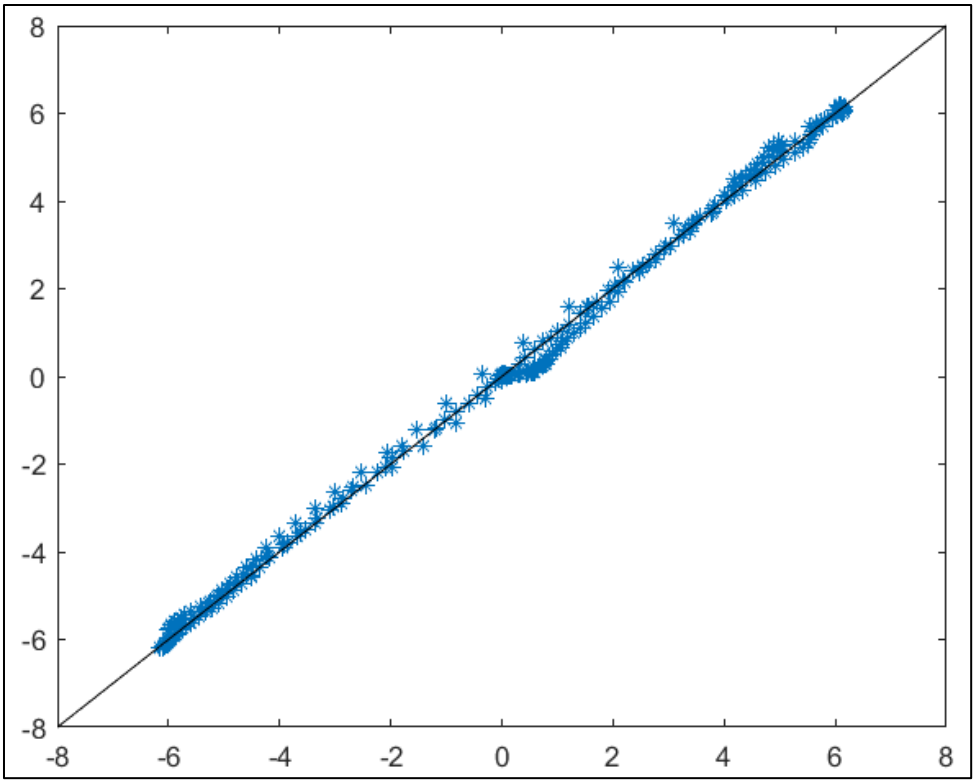
Accuracy = 98.78%

Final prediction error = 0.002428

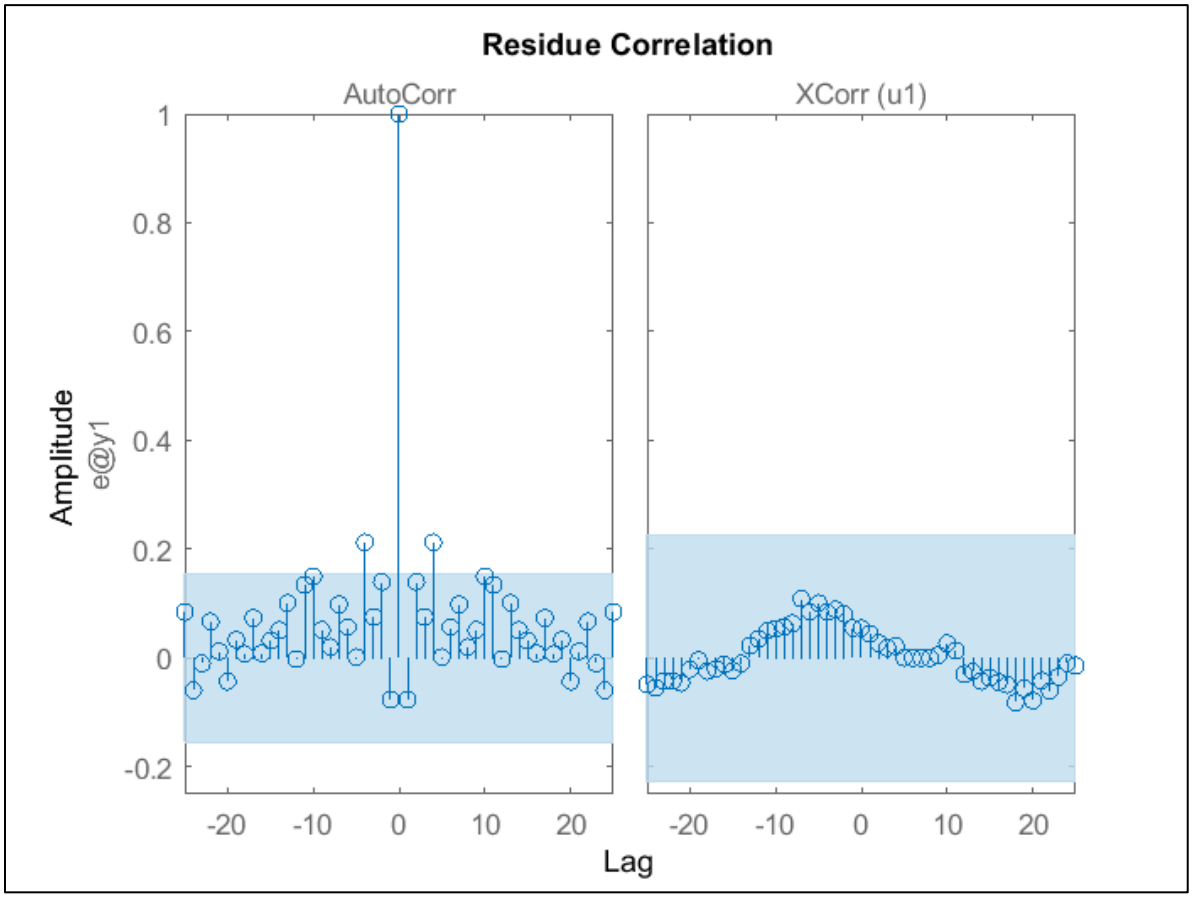
Mean-squared error = 0.00229

c.) Plotting the measured temperature versus the estimated temperature to assess the quality of the model.





- Residual correlation



We assume for this process that it contains White Noise with mean zero. The residual function is used to validate the quality of the model used in this process.

1. Whiteness Test: The image on the left shows the autocorrelation of the residual data. It can be concluded that for lag zero, the amplitude is 1 which is good. Only 2 points go beyond the confidence bound for the non-zero tags and not a significant way.
2. Independence Test: The image on the right shows the correlation between the input data and the residual data. It can be concluded that these data do not show any correlation.

Thus, I am justified in using the ARX structure for this data.

d.) The superimposed predicted output over the measured output, the plot of measured output vs the predicted output and the residual correlation suggest that the ARX was a pretty model for this system. The accuracy, mean-squared error, final prediction error and the residual correlation analysis suggest the same.