

University of Alberta Computer Process Control Group

Data Driven Performance Analysis

Limited Trial Version

Written by: CPC Control Group, University of Alberta

Version 1.0

Table of Contents

Introduction	1
System Requirements	1
Data Generation	1
Quick Start	2
References	7

List of Figures

Figure 1: The Main GUI	3
Figure 2: Screen shot after loading the data	4
Figure 3: Screen shot after clicking "Run"	5
Figure 4: In-Depth View of the Individual p_i Curves	6

Introduction

The data driven performance analysis toolbox was developed by the Computer Process Control Group at the University of Alberta to analysis the performance of a closed-loop process using prediction error and subspace analysis. The results are based on those presented in (Huang, Ding, & Thornhill, 2006).

A “Quick Start” approach to using this toolbox will be presented.

System Requirements

In order to run this programme properly, the following software is required:

- 1) MATLAB 2006a (MATLAB 7.1) or above. It should be noted that this toolbox was optimised for use on MATLAB (MATLAB 2008a). However, there are no issues with using the toolbox on MATLAB 2006a or above.
- 2) The SYSTEM IDENTIFICATION TOOLBOX from MATLAB is required.

Data Generation

The reference data should be stored as the variable “Base_Data”, which contains N rows and M columns, where N is the number of samples and M is the number of output variables. The file name of the .m file that contains the variable is irrelevant. Similarly, the reference data is stored in the variable “Comp_Data”, which is the same size as the reference data. Both variables can be stored in the same file, but they must be loaded separately. For example, if it is desired to load the following data: Reference data is a 5-by-2 matrix with entries

$$R = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \\ 11 & 12 \\ 14 & 15 \end{bmatrix}$$

and the evaluation data is a 5-by-2 matrix with entries

$$E = \begin{bmatrix} 3 & 4 \\ 6 & 7 \\ 9 & 10 \\ 13 & 14 \\ 16 & 17 \end{bmatrix}$$

then, the following would be performed in MATLAB:

```
>>Base_Data=[1 2; 4 5; 7 8;11 12; 14 15];
>>save filename.m Base_Data
>>Comp_Data=[3 4; 6 7; 9 10;13 14; 16 17];
>>save filename2.m Comp_Data
>>%If it is desired to save everything into a single file
>> save filenameAll.m Base_Data Comp_Data
```

Quick Start

For quickly using the toolbox, the following steps should be followed:

- 1) Unzip the files to the desired location.
- 2) Start MATLAB, and point the current directory to the location of the unzipped files.
- 3) At the command prompt, type “>> **DataDriven**” to start the toolbox. The GUI shown in Figure 1 should appear. There are 3 main sections:
 - a. **Main Menu**, from which it is possible to load the required data and run the programme.
 - b. **User Input**, which shows the data that the user has loaded into the programme. It is possible to select time sections by clicking the “Select Time Sections” button. For the reference data, only a single time section is permitted, while for the evaluation data, any number of time sections is permitted. The time sections are entered as matrices, using the format: “[t_{i1} , t_{f1} ; t_{i2} , t_{f2}]”, where t_i refers to the start time, t_f refers to the end time, and a semicolon separates the different time sections.
 - c. **Output**, which contains a plot of the p_i curves for both the overall process and the individual output variables. It is possible to pop-out the graphs, by clicking the “Expand...” buttons located underneath the graphs. As well, the i value, which represents the number of points displayed for the curve, can be changed by clicking the “Change i-value” button.

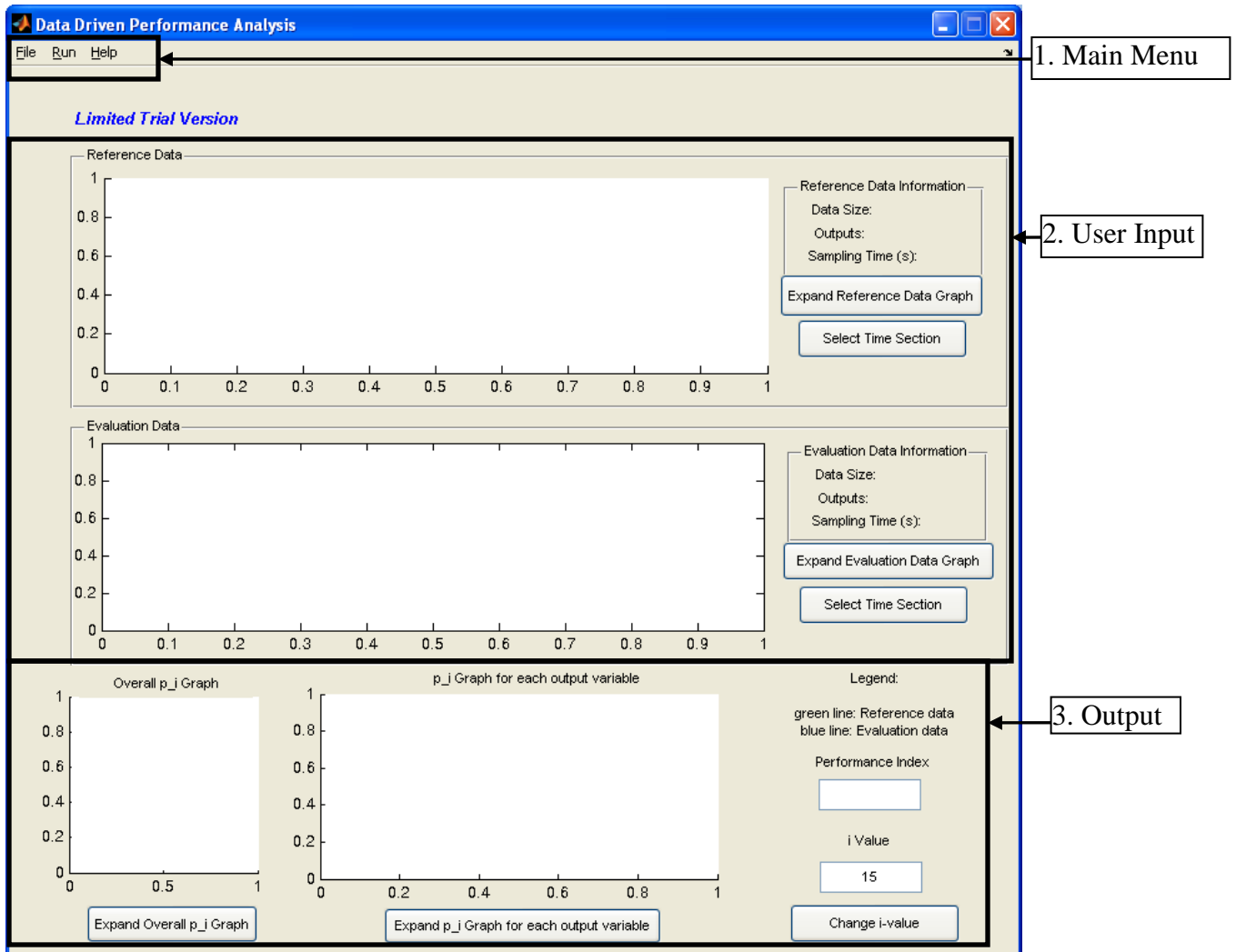


Figure 1: The Main GUI

- 4) For the purpose of this quick start, we will be using the sample data provided in the zip file. Go to the “File” Menu located in the Main Menu area, which is denoted as 1 in Figure 1. Navigate to Load → Data→Reference Data to load the reference data. Alternatively, press Ctrl+D to perform the same operation. Select the file “DataFile.mat” and click “Open”. Any file that contains stores the reference data in the matrix called “Base_Data” will work. Next, the evaluation data must be selected. Go to File →Load →Data→Evaluation Data. Alternatively, press Ctrl+E. Select the file “DataEval.mat” and click “Open”. Both data files contain the closed-loop, operating data for the process. Once the data has been loaded, the GUI should look similar to that shown in Figure 2.

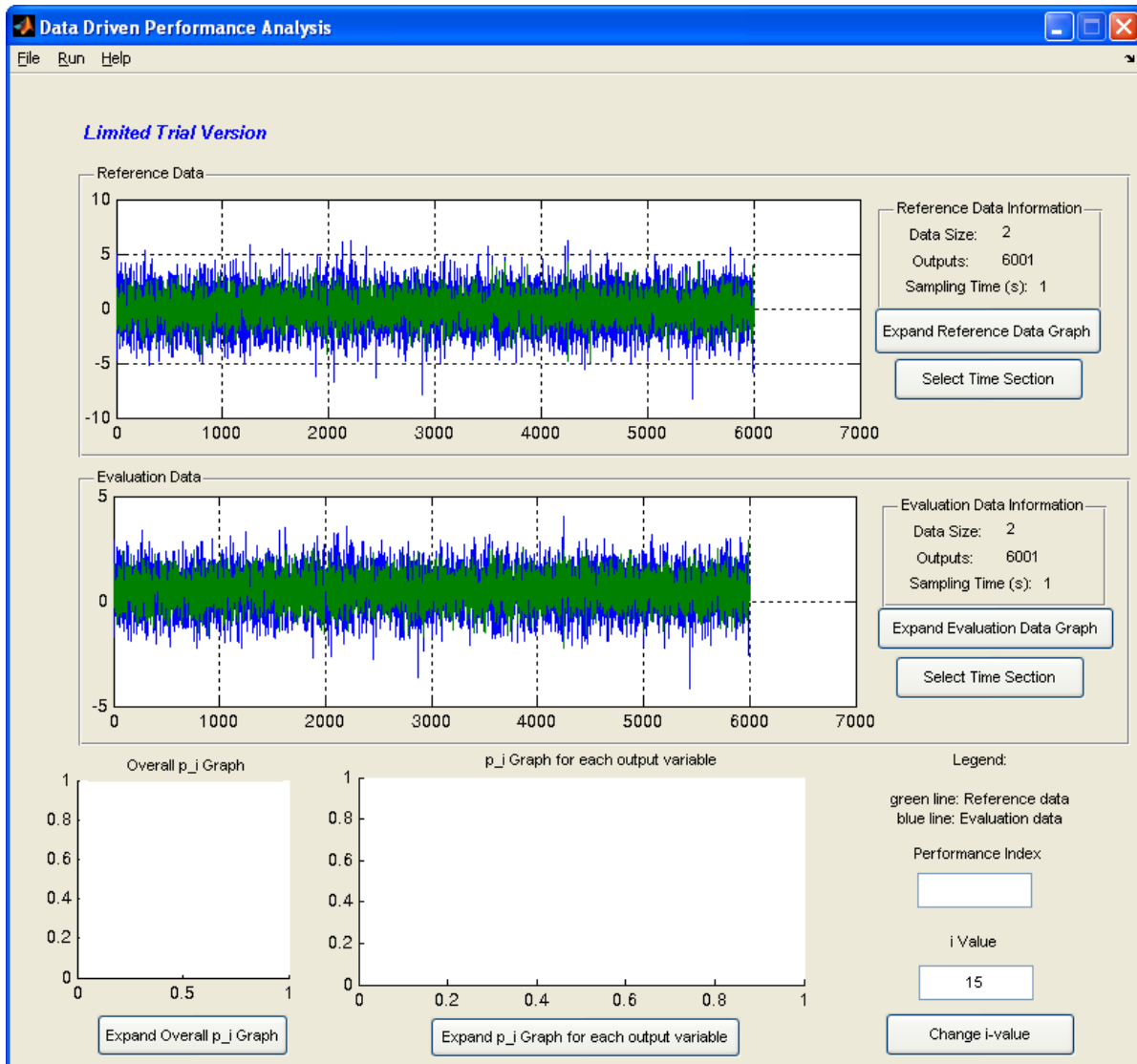


Figure 2: Screen shot after loading the data

- 5) We will select 2 time sections for the evaluation data. Press the “Select Time Section” button located in the Evaluation Data box. In the box that appears, enter “[1,6001;200,4000]”. This suggests that we wish to consider both all of the data (1 to 6001 time samples) and a section of the data from time sample 200 to time sample 4,000.
- 6) Once this has been entered, click the “Run” menu and select “Run” or press Ctrl+R to start processing the data.
- 7) After a few seconds, the GUI should become similar to that shown in Figure 3.

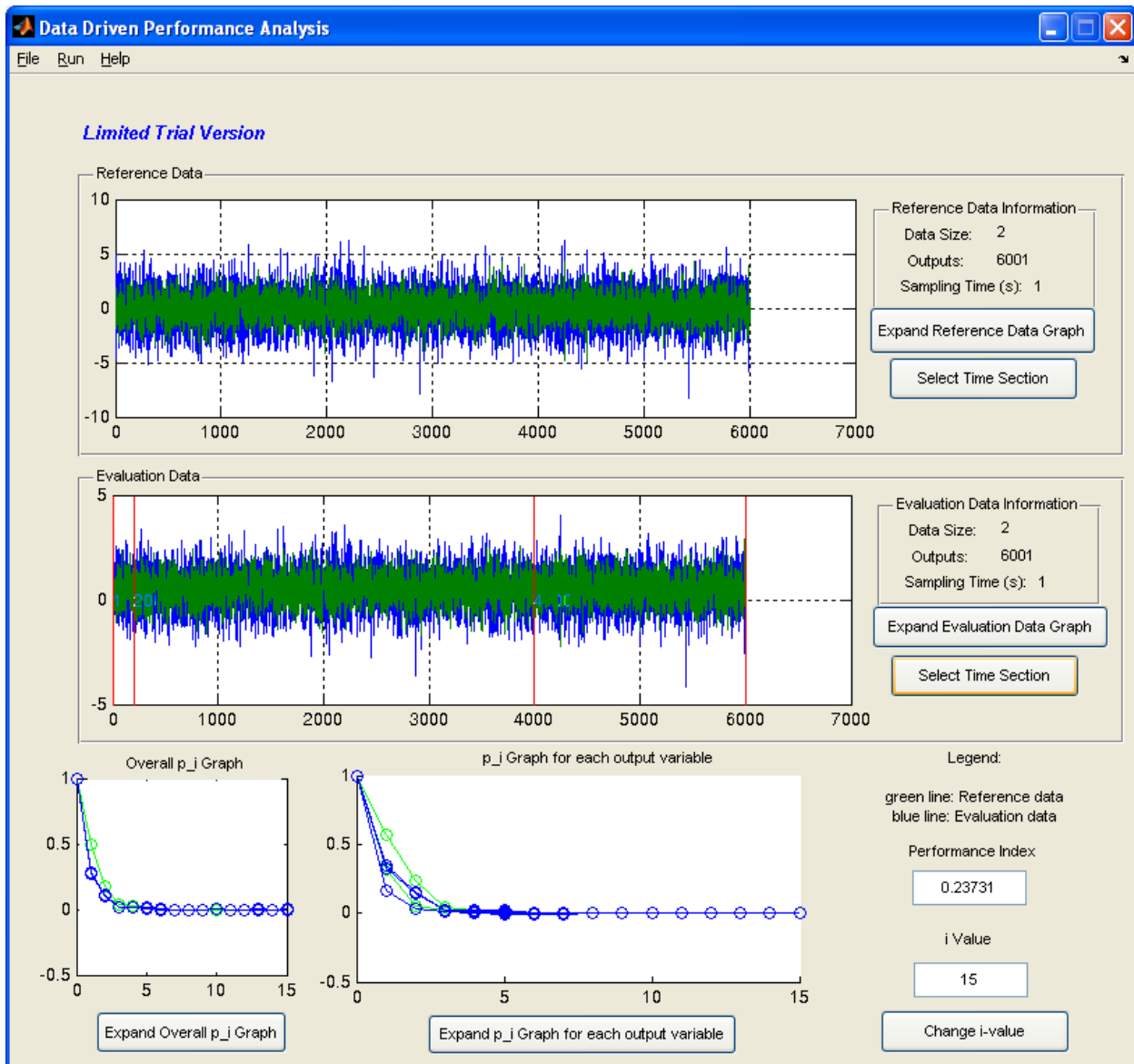


Figure 3: Screen shot after clicking "Run"

- 8) Figure 3 shows that in general, the reference data (green line) is always located above the evaluation data (blue line). This suggests that significant improvement can be obtained for the process. In order to see the data better, press the "Expand p_i Graph for each output variable". The Figure shown in Figure 4 should appear. This figure accentuates the above results.

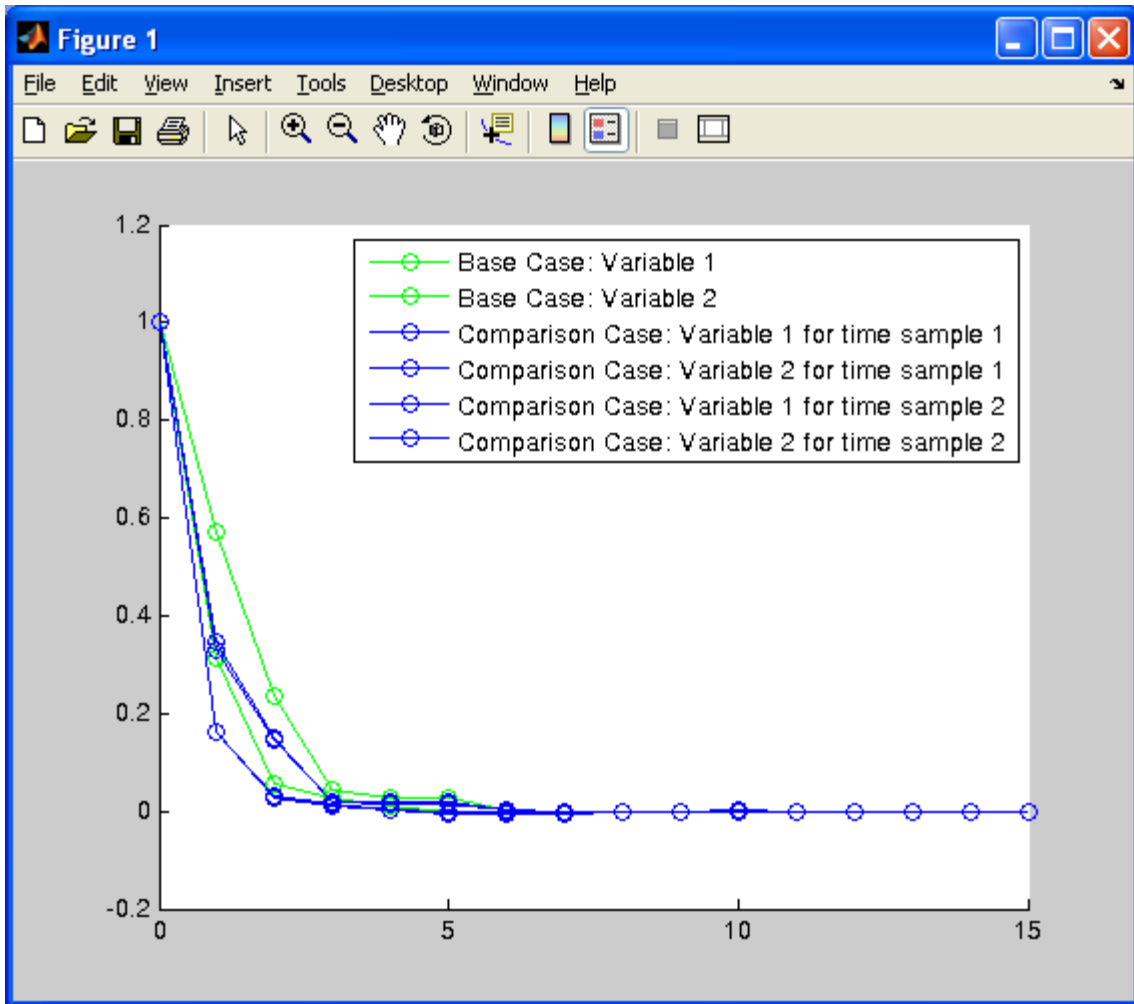


Figure 4: In-Depth View of the Individual p_i Curves

9) To save the current model, go to File → Save. The data will be saved as a .mat file. This saved file can be loaded using the File→Load→Saved File (Ctrl+L) option.

10) To exit the toolbox, click the File→Exit.

References

Huang, B., Ding, S. X., & Thornhill, N. (2006). Alternative Solutions to Multivariate Feedback Control Performance Assessment Problems. *Journal of Process Control*, 16, 457-471.