Identification of Stochastic Process in MATLAB

Ojonuigwa Adukwu

Department of Industrial and Production Engineering, School of Engineering and Engineering Technology

The Federal University of Technology Akure

Akure, Nigeria

Email: oaualaku [AT] yahoo.co.uk

_____

ABSTRACT---- *The system identification toolbox in MATLAB has been successfully used to compare model identification of a first order system subjected to high and low disturbances. The model structures used are FIR, ARX, AMX, OE and BJ. The obtained Model was validated using data generated from the actual process. It shows that the more the variance of the noise input into the system, the more difficult it is for the model identified to reproduce that validation data obtained from process response. Also when the measurement noise has zero mean and low variance, the effect on the steady state gain and other process parameters is negligible.*

Kyewords---- System Identification, model structures, process noise
_____

1. INTRODUCTION

System identification is used to obtain dynamic models of processes. One basic tools used for building these dynamic models is the system identification tool box in MATLAB [1]. The tool box uses both graphical approach for model building and command in the command line interface. Both the graphical user interface and the command line interface are used to model and display the system models graphically and also extract model parameters [2, 3]. Models are used for various purposes such as state estimation, prediction and fault detection [4].

A general input-output linear model for a single-output system with input u and output y can be written as:

$$A(q)y(t) = \sum_{i=1}^{nu}[\frac{B_i(q)}{F_i(q)}]u_i(t - nk_i) + [\frac{C(q)}{D(q)}]e(t) \qquad 1.$$

Here u_i denotes input number i, and $A, B_i, C, D \ and \ F_i$ are polynomials in the shift operator (q). The general structure is defined by giving the time-delays nk and the orders of these polynomials (i.e., the number of poles and zeros of the dynamic model from u to y, as well as of the disturbance model from e to y).Based on the polynomial used, the given process can be classified as;

FIR; if B only is used

ARX: if A and B only are used

ARMAX; if A, B and C only are used

OE (Output-Error); if B and F only are used

BJ (Box-Jenkins); if B, F, C and D only are used. [5, 6]

In this research, both the graphical user interface and the command line interface of the system identification tool box are used to model a first order system subjected to high and low disturbances through first and second order filters respectively using a step input. Thereafter, two input process, the second input being measurement noise were modeled. The main aim of the research is to model the system under different disturbance condition and compare the performance of the various model structures obtained with the original data and the predicted models.

2. DISTURBANCE INJECTION

The system is simulated for 600s with step time at 275s. Three process conditions were simulated; the clean process devoid of measurement noise and disturbances, the V1 process having disturbance of zero mean and 0.001 variance added through a first order filter and V2 process having disturbance of zero mean and 0.1 variance added through a second order filter.

Figure 1 shows the responses of clean, low disturbance V1 and high disturbance V2 process.
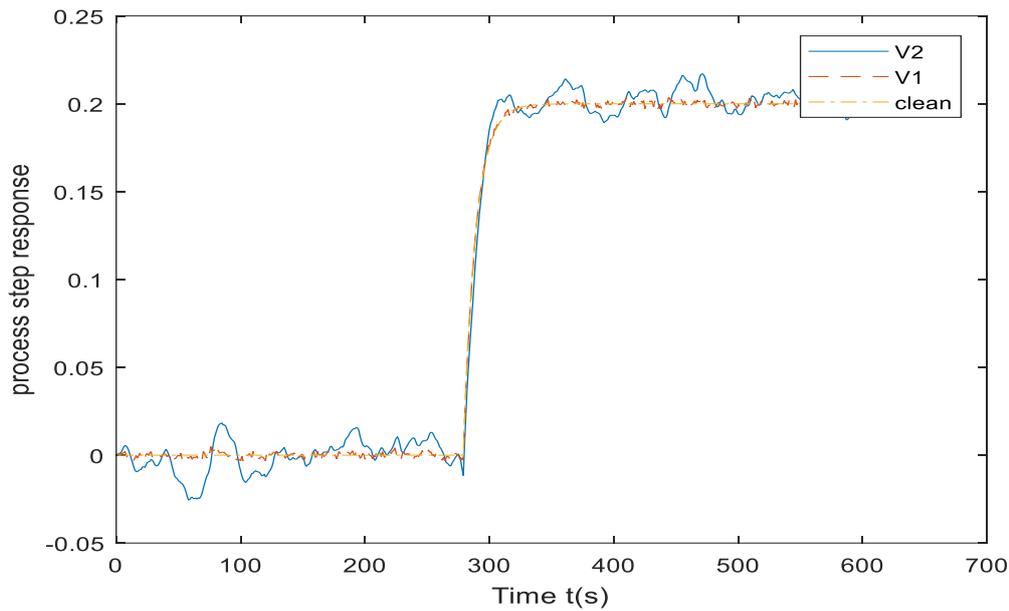
Figure 1`: Process step response with clean, V1 and V2 disturbances.

From figure 1, the clean process and the V1 process responses are almost the same while the disturbance effect from the V2 process makes V2 process to differ from the other two. This is because the high variance of the noise input to the second order filter makes disturbance V2 to have more effect on the process. This is so despite higher order filters having better effect on disturbance attenuation than lower order filter.

3. PROCESS MODEL ESTIMATION

The clean process in item section 2 above was estimated using various model structures in MATLAB system Identification Toolbox. The FIR was estimated using time span of -5s to 50s since the approximate settling time of the process is 50s. The optimal ARX model was searched automatically by the toolbox and the best model order is ARX141. The toolbox has no means of automatically searching for the right model order for the AMX, OE and BJ. These were searched manually by first varying the output coefficient (na) till the best fit is obtained. If the best fit value starts declining before reaching 100%, input coefficient, nb is then searched till the highest attainable is obtained. Table 1 shows the results for ARMAX, OE and BJ structure. From table 1, the best fits for AMX, OE and BJ are all 100% at AMX2421, OE421 and BJ42221 respectively, hence these model orders are selected. Figure 2 shows the model output for the original data (process response data), impulse response, ARX141, AMX2421, OE421 and BJ42221. From the figure it is seen that all the models matched the data except the impulse response. The only parameter change for the impulse response models is the steady state gain which is approximately 1.9 (0.19/0.1).

Table 1: Model orders with best fits (%)

AMX		OE		BJ	
Order	Best fit (%)	Order	Bets fit (%)	Order	Best Fit (%)
AMX1221	96.97	OE121	98.74	BJ12221	98.73
AMX2221	99.39	OE221	99.48	Bj22221	99.45
AMX3221	99.63	OE321	99.69	BJ32221	99.62
AMX2321	99.66	OE421	100	BJ42221	100
AMX2421	100				

Figure 3 shows step response to the clean process and model when a step input of 0.1 amplitude is applied. From the figure, it is seen that both the process response and the models responses are exactly alike except for the finite impulse response (FIR). The difference arises from the fact that the FIR best fit to model data was poor (98%) as shown in figure 2.
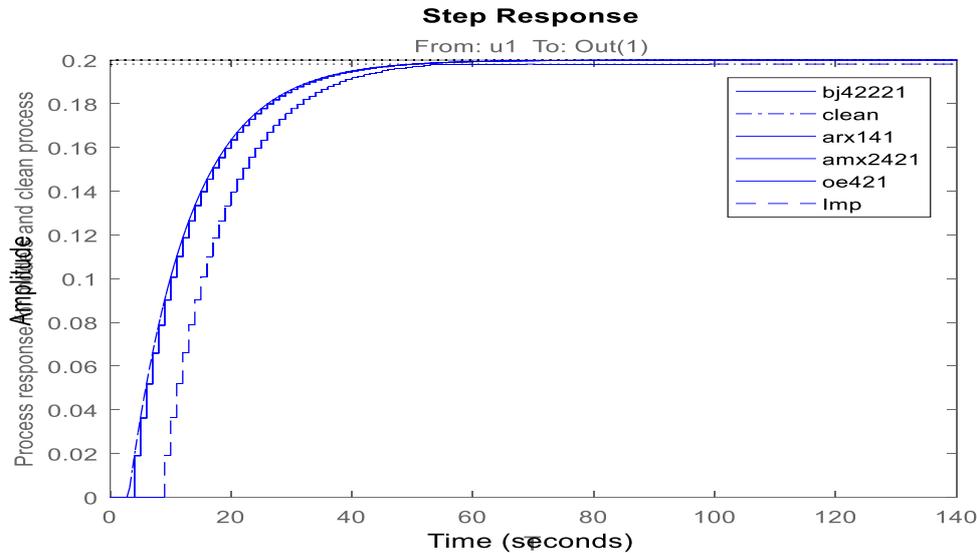

Figure 2: Clean Process step response with various model structures.
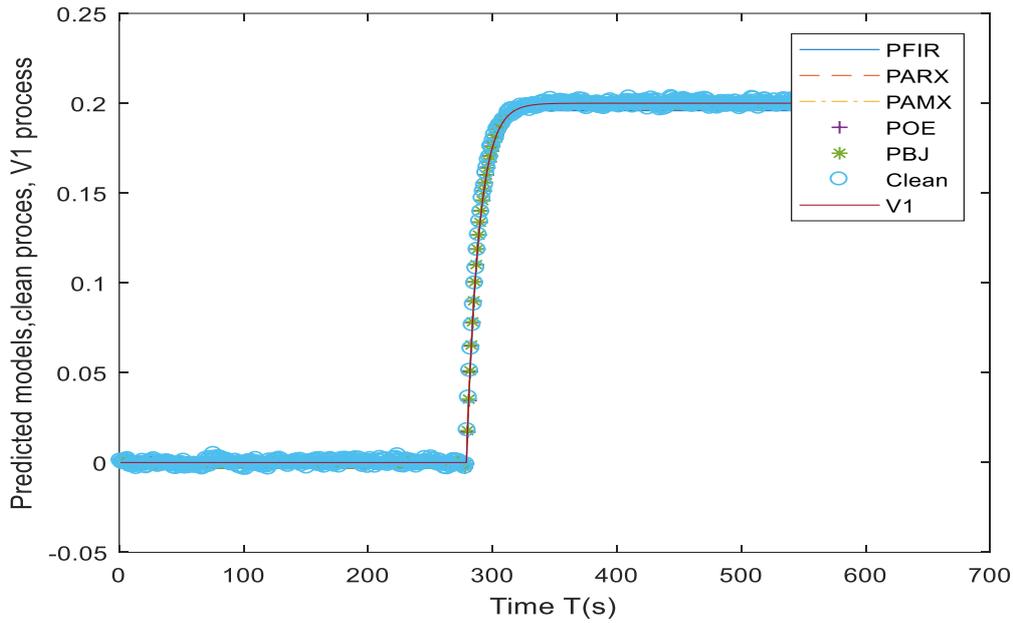
4. NOISE INJECTON

Noise of zero mean and variance 0.001 was injected into the system through a first order filter to obtain process with disturbance V1 (low disturbance). The V1 process as in section 3 above was estimated using various model structures in MATLAB system Identification Toolbox. The FIR was estimated using time span of -5s to 50s since the approximate settling time of the process is 50s. The optimal ARX model was searched automatically by the toolbox and the best model order is ARX1041. The other models were searched manually. Table 2 shows the results for the optimal model structure and the percentage of fit .

Table 2: Model orders with best fits (%) for second experiment

Models	FIR	ARX1041	AMX3421	OE421	BJ43221
Best Fit (%)	92.2	92.08	91.13	92.12	91.92

From table 2, the best structure obtained are AMX3421 (98.61%), OE421 (98.61%) and BJ42221 (98.62%). Compared to results from table 1, none of the model structures have best fit of 100%. This is because despite the validation data being the same as the working data, the model cannot reproduce itself exactly due to the additive disturbance that is random in nature.

Figure 3 compares predicted process with low disturbance (V1) models with V1 measured data and the clean process step response with amplitude 0.1. The figure shows that no much differences between the predicted models and the measured data for the low disturbance model. There exist slight differences between the predicted models and the clean data. This is indicated by the noisy pattern on the predicted models graphs. Compare to figure 2, the graph look identical except the presence of noise envelop created by the predicted graph and the V1 measured data.

Figure 3: Predicted low disturbance process models, clean data and V1 data.

Noise of zero mean and variance 0.1 was injected into the process through a second order filter to obtain process with disturbance V2 (high disturbance). The V2 process as in section 3 above was estimated using various model structures in MATLAB system Identification Toolbox as described earlier for V1. Figure 4 shows the model output for high disturbance process along with the measured data. Compared to figure 2 for clean process where the measured data was reproduced exactly, the models in figure 4 were far from the measured data. Also compared to figure 3, the estimated models step responses are farther from the measured signal. This is due to the fact that with high disturbance input, the process output becomes more stochastic and the variance of the process becomes bigger.

It can also be seen from figure 4 that the predicted model response nearly matched the clean process. Closer look shows that the clean process response lies in the middle of the predicted process. This shows that no matter the disturbance effect on a clean process, the mean response of the repeated experiment on disturbed process approaches the true (clean in this case) process response.
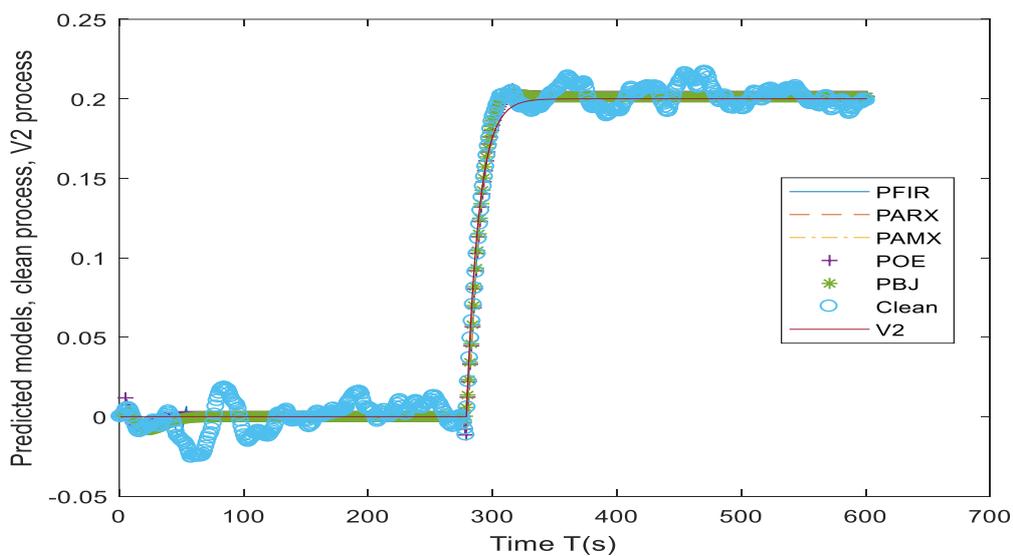
Figure 4: Predicted high disturbance process models, clean data and V2 data.

5. TWO INPUT PROCESS

The procedure in section 4 was repeated but this time using two inputs to the process, the normal step input of 0.1 amplitude and a random input of zero mean and 0.000001 variance. We start with low disturbance in this two-input process. Figure 5 shows the outputs while figure 6 shows arx1041 for this two-input process, clean and the two-input data responses. Figure 7 shows arx1041 for this two-input process, V1 and V2 responses. Figure 5 shows that the model tracks the response very closely. The presence of additive disturbance is also visible from random response. From figure 6, the models for ARX, clean and the response data also match very closely. But considering figure 7, the ARX model only approaches V1 process. This shows that the noise input through the system has negligible effect on process output due to the noise variance being too low.
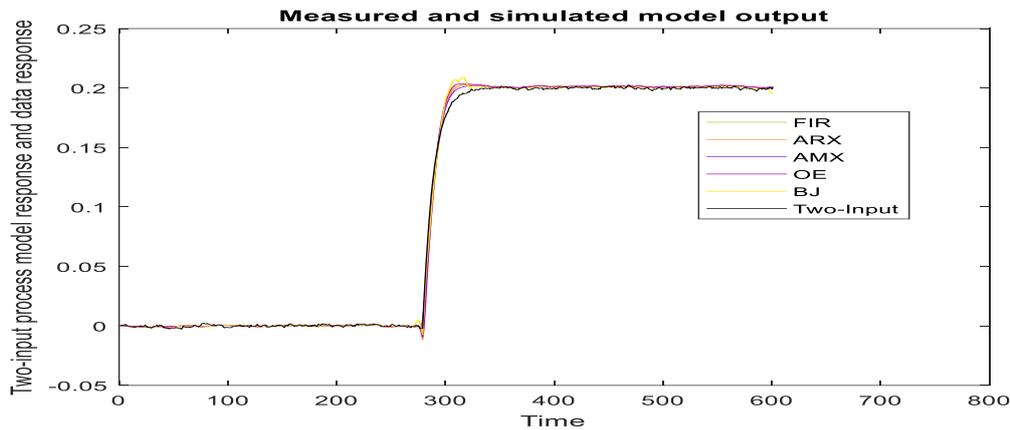

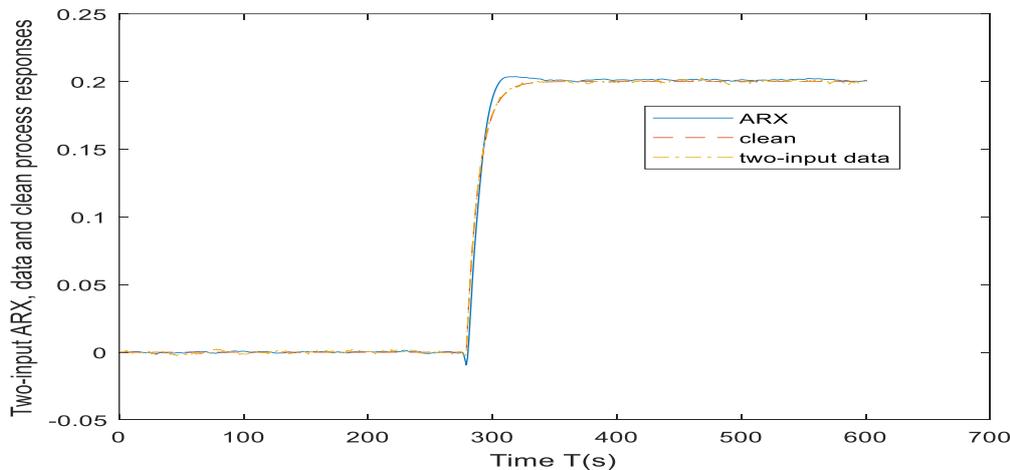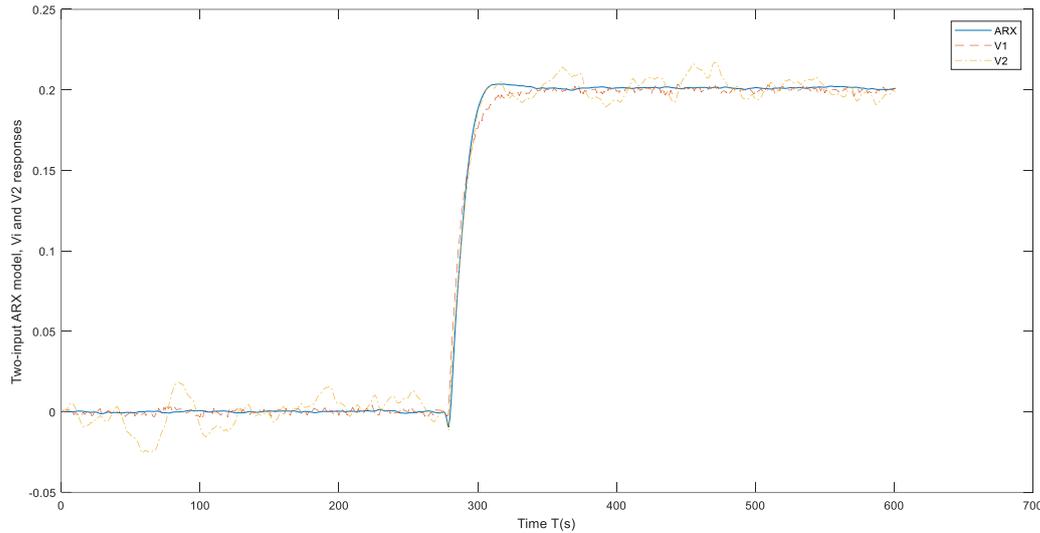*Figure 5: Two-input models output and process response.*

Figure 6: Two-input ARX model, data and clean process responses

Figure 7: Two-input ARX model, V1 and V2 process responses.

It can be seen from figures 5-7 that the steady state gain still remains the same for the disturbed and the clean process. The steady state gain is therefore not affected by the two-input process.

6. CONCLUSION

The system identification toolbox in MATLAB has been successfully used to model a first order system subjected to high and low disturbances. The model structures used are FIR, ARX, AMX, OE and BJ. The more the variance of the noise input into the system, the more difficult it is for the model identified to reproduce that validation data which is the data obtained from process response. Also when the measurement noise has zero mean and low variance, the effect on the steady state gain and other process parameters is negligible. The stationary gain of the models is basically independent of the level of disturbance input as long as the right model structure is selected. It is also independent on the number of input when the second input is noise of low variance and zero mean.

7. REFERENCES

[1] Kollar I., Pintelon R., Schoulkens J. (1991), Frequency Domain Identification ToolBox for MATLAB, *IFAC Proceedings*, 24(3), pp 1243-1247.

[2] Jimenez M.J., Madsen H., Anderson K.K., (2008), Identification of the Main Characteristics of Building Components Using MATLAB, *Journal of Building and Environment*, 43(2), pp. 170-180.

[3] Yuen K.V., Mu H.Q. (2015)., Real-Time Identification: An Algorithm for Simultaneous Model Class Selection and Parametric Identification, *Journal of Computer-Aided Civil and Infrastructure Engineering*, pp 1-17.

[4] Adukwu O., Oku D. E. (2018), Modeling Analysis and Monitoring DC Motor Using Proportional-Integral Controller and Kalman Filter, *International Journal of Science and Research*, 7(8), pp 454-458.

[5] Ljung L. (1987), System Identification: Theory for User, Prentice-Hall, Englewood Cliffs, NJ, USA.

[6] Ljung L. (1998), System Identification, Birkhäuser, Boston, MA, USA.