INVESTIGATION ON EFFECT OF ERROR ORDER SELECTION IN SYSTEM IDENTIFICATION

MOHAMAD SYAHMI BIN MOHD BAKRI

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

SUPERVISOR DECLARATION

I hereby declare that I have read this project report and in my opinion this report is sufficient in terms of scope and quality for the award of the degree of Bachelor of Mechanical Engineering (Structure & Materials).

Signature :. ....

Name of Supervisor : DR. MD FAHMI BIN ABD SAMAD @ MAHMOOD

Date : ....

INVESTIGATION ON EFFECT OF ERROR ORDER SELECTION IN SYSTEM IDENTIFICATION

MOHAMAD SYAHMI BIN MOHD BAKRI

This report is submitted in fulfillment of the requirement for the degree of Bachelor of Mechanical Engineering (Structure & Material)

Faculty of Mechanical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

MAY 2017

DECLARATION

I declare that this project report entitled "Investigation On Effect Of Error Order Selection In System Identification" is the result of my own work except as cited in the references

Signature	:	....
Name	:	MOHAMAD SYAHMI BIN MOHD BAKRI
Date	:	....

DEDICATION

To my beloved mother and father

ACKNOWLEDGEMENT

First and above all, I praise Allah the almighty for providing me this opportunity and granting me the capability to proceed successfully. This project appears in its current form due to the assistance and guidance of several people. Special appreciation goes to my supervisor, Dr. Md Fahmi bin Abd Samad @ Mahmood, for his relentless support, thoughtful guidance and correction of the project. I attribute the level of my bachelor degree to his encouragement and effort and without him, this project would not have been completed or written.

My deepest gratitude goes to my beloved parents, Mohamad Bakri Bin Abd Razak and Anim Binti Hamdan for their material and spiritual support in all aspects of my life. Thank you both for giving me strength to reach for the stars and chase my dreams. I also would like to thank to all my siblings for their endless love, prayers and encouragement. I am particularly grateful to my sweetheart for his devoted support throughout my study.

Last but not least, thanks to all my close friends Jamry Bin Samut for the joyful gatherings and all his supports. I deeply appreciate to all my friends that helped me stay through these difficult years. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship. To those who indirectly contributed in this research, your kindness means a lot to me. Thank you very much.

ABSTRACT

System identification is a technique or application that aims to develop mathematical models for dynamical systems using measured input and output signals. There are 4 important steps which is the observed data, model, parameter estimation and validation. In this project, model structure selection is important because the objective is to find the effect of the error order selection on system identification. Linear model which is ARX and ARMAX model is used to make the simulation to investigate effect of error order. This project was carried out using the Graphical User Interface (GUI) in MATLAB application. Firstly, 3 equations were randomly created with supervisor guide and data acquisition was run to generate 500 data by using the MATLAB software command and save it into folder for next step use. After that, system identification toolbox is open and imports the data that generate from the equation. Next step is modelling where the data use ARMAX model and different sequence in error order or n_c. This project result is based on 5 performance indicator like Means Square Error (MSE), Model Output, Model Residuals, Akaike's Final Prediction Error (FPE) and Parameter Values. This entire indicator will pop out after click on button indicator or from the data of each model. Based on result that acquire from this investigation, the effect or error order selection have a very small difference compared to the true model. This is maybe because the data was not completely scatter and need to bigger. So the conclusion is, there is small effect of error order selection in system identification.

ABSTRAK

Pengenalanpastian sistem adalah satu teknik atau aplikasi yang bertujuan untuk membangunkan model matematik untuk sistem dinamik menggunakan isyarat input dan output yang diukur. Terdapat 4 langkah penting yang merupakan data yang diperhatikan, model, parameter anggaran dan pengesahan. Dalam projek ini, pemilihan struktur model adalah penting kerana tujuannya adalah untuk mencari kesan pemilihan turutan ralat pada pengenalpastian sistem. Model linear yang merupakan model ARX dan ARMAX adalah digunakan untuk membuat simulasi untuk menyiasat kesan nilai turutan ralat. Projek ini telah dijalankan dengan menggunakan antara muka grafik pengguna (GUI) dalam perisian MATLAB. Pertama, 3 persamaan diciptakan secara rawak dengan panduan penyelia dan perolehan data telah dijalankan untuk menjana 500 data dengan menggunakan arahan perisian MATLAB dan simpan ke dalam folder untuk digunakan langkah seterusnya. Selepas itu, toolbox pengenalpastian sistem dibuka dan mengimport data yang dijana membentuk persamaan. Langkah seterusnya ialah memodelkan data dengan menggunakan model ARMAX dan menggunakan turutan yang berbeza untuk ralat atau n_c*. Hasil projek ini adalah berdasarkan kepada 5 petunjuk prestasi seperti Means Square Error (MSE), keluaran model, sisa model, Ramalan Ralat Akhir Akaike (FPE) dan Parameter. Kesemua penunjuk akan ditunjukkan selepas klik pada penunjuk butang atau dari maklumat setiap model. Berdasarkan keputusan yang diperolehi daripada penyiasatan ini, kesan turutan ralat sangat kecil apabila dibanding dengan model sebenar. Ini masih boleh diterima kerana mungkin data tidak sepenuhnya berselerak dan data perlu lebih besar. Jadi kesimpulannya ialah, terdapat kesan kecil pada pemilihan turutan ralat dalam pengenalpastian sistem.*

CONTENT

LIST OF FIGURES

LIST OF ABBEREVATIONS

ARX	Exogenous Input
ARMAX	Auto Regressive Moving Average Exogenous Input
MATLAB	Matrix Laboratory
GUI	Graphical User Interface
RELS	Recursive Extended Least Square
PEM	Prediction Error Minimization
y(t)	Output at time t.
n_a	Number of poles.
n_b	Number of zeroes plus 1.
n_k	Number of input samples that occur before the input affects the output, also called the dead time in the system.
e(t)	White-noise disturbance value.
n_c	Number of C coefficients.

CHAPTER 1

INTRODUCTION

1.1 Background

In recent years, system identification has been discussed and gets attention from many universities. This is because system identification was start use in some industries. In the other word, system identification can be categories as the basic necessity in regions, for example, control, correspondence, power system and instrumentation for acquiring a model of a system of interest or another system to be produced.

System identification is the art and science of building mathematical models of dynamic systems from observed input–output data. It can be seen as the interface between the real world of applications and the mathematical world of control theory and model abstractions. As such, it is a ubiquitous necessity for successful applications (Ljung, 2010). Actually system identification is a one term in control area that refers as a technique constructing model from observed data.
There are four basic ingredients in system identification:

1) The observed data
2) Model structure
3) A criterion of fit
4) Validation

When formulating and solving an identification problem it is important to have the purpose of the identification in mind. In control problems the final goal is often to design control strategies for a particular system. There are, however, also situations where the primary interest is to analyse the properties of a system. The purpose of the identification is to design a control system. The character of the problem might vary widely depending on the nature of the control problem (Astrom, 1971). For the example, design a stable regulator or design a control program for optimal transition from one state to another.

In this project, the main objective is to investigate the effect of error order selection in system identification. The error order that only occur in ARMAX model will be discuss to. So to make this experiment happen, this experiment will be carry out using MATLAB's system identification graphical user interface (GUI) named _IDENT' found in System Identification Toolbox.

1.2 Problem Statement

There are many aspects that we can discuss in system identification such as type of model, the techniques and many more. System identification deals with the problem of building mathematical models of dynamic system based on observed data from the system. The subject is thus part of basic scientific methodology and since dynamical system is abundant in our environment, the techniques of system identification have a wide application area (Ljung., 1987). Besides, the most important in system identification is to achieve model from system data and from the selection of error order will cause an accurate model. This experiment will discuss about vary the value of error order that happen to the accurate model. Since the presence of measurement error order only occur in ARMAX model, the discussion about why wrong selection of error order will cause the developed ARMAX model unable to represent system behaviour.

ARMAX model are widely used in identification and are a standard tool in control engineering for both system description and control design. These models, however, can be non-realistic in many practical contexts because of the presence of measurement errors that play an important role in applications like fault diagnosis and optimal filtering. ARMAX models can be enhanced by introducing also additive error terms on the input and output observations. How much the error order effects the output between the actual result and simulation result in ARMAX model. Besides that, we want to know what the different output result when the wrong selection of error order is estimate in the system.

1.3 OBJECTIVE

The objectives of this project are as follows:

1. To investigate the effect of error order on performance of identification.
2. To perform system identification using linear difference equation which is auto regressive moving average exogenous input (ARMAX) model.

1.4 Scope of Project

The scopes of this project are:

1. Identification will be performed using GUI (Graphical User Interfaces) 'ident' on MATLAB (matrix laboratory) software.
2. Performance comparison will be made from various perspectives such as final prediction error and established loss function.

CHAPTER 2

LITERATURE REVIEW

2.1 System Identification

System identification is the art and science of building mathematical model of dynamic systems from observed input-output data. It can be seen as the interface between the real world of application and the mathematical world of control theory and model (Ljung, 2010). Or in other way, system identification is the process of developing or improving a quantitative numerical model from a set of input and output data that represents the response of a dynamic system. It is necessary to use model to describe the relationships among the system variables. The developed model has the trademark performance similar like the unknown system.

The parameter estimation step decides inside of the arrangement of models, the model that is the best guess or gives the best clarification of the observed data. The estimation of the model parameters relates to the minimization of the chosen criterion. The decision of basis relies on upon the accessible data about and the motivation behind the model. Model validation is conceivably the most vital stride in the model building arrangement. It is likewise a standout amongst the most disregarded. Regularly the approval of a model appears to comprise of simply citing the measurement from the fit (which measures the part of the aggregate variability in the reaction that is represented by the model).

As mention before, there are four steps in system identification which is the observed data, model structure, a criterion of fit and validation and the loop of this all four step can be determine as Figure 2.1 and will be discuss more detail in next.

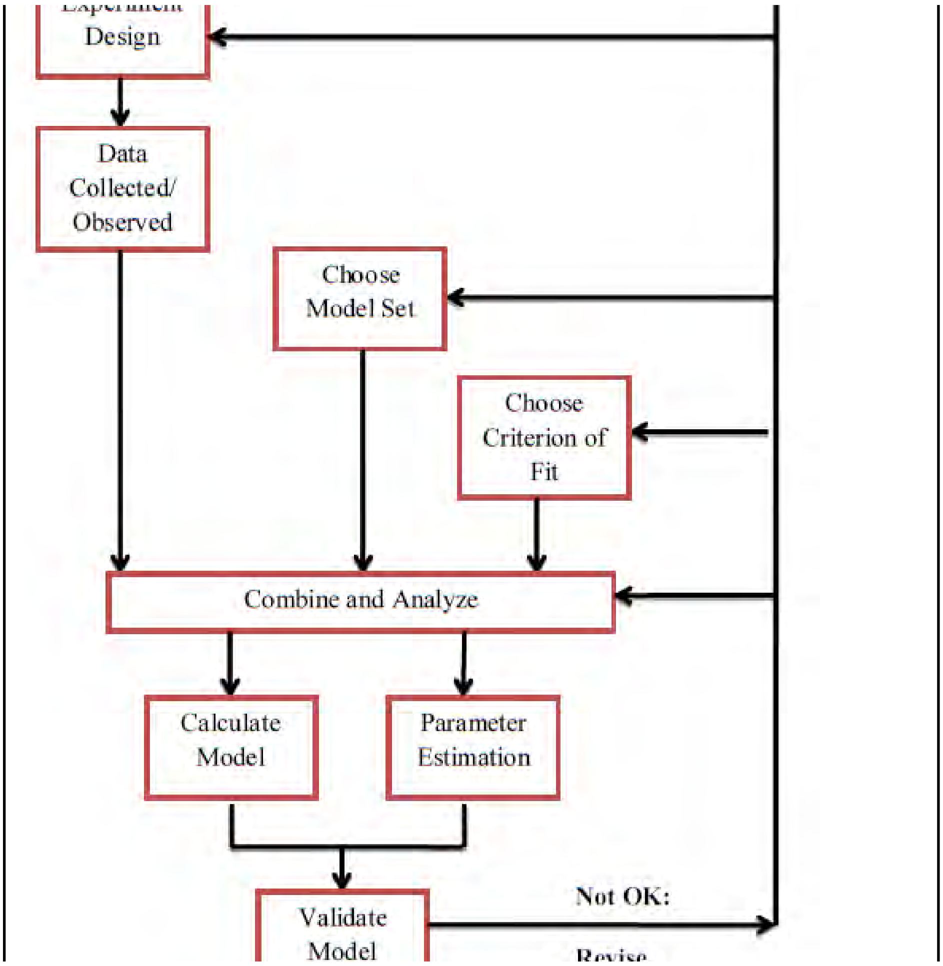

Figure 2.1: Loop of System Identification Step

2.1.1 The Observed Data

If we want to build a model for a system, we should get data or information about it. This can be done by just watching the natural fluctuation (e.g., vibration analysis of a bridge that is excited by normal traffic), but most often efficient to set up dedicated experiments that is actively excites the system (e.g., controlled excitation that optimize his own goal (for example, minimum cost, minimum time, or minimum power consumption for a given measurement accuracy) within the operator constraints (e.g., the excitation should remain below a maximum allowable level). The quality of the final result can depend heavily on the choices that are made

2.1.2 Model Structure

A choice should be made within all the possible mathematical models that can be used to represent the system. Again a wide variety of possibilities exist, such as below;

- Parametric versus non-parametric models

In a parametric model, the system is described using a limited number of characteristic quantities called the parameter of the model, whereas in a non-parametric model the system is characterized by measurements of a system function at a large number of points. Example of parametric models is the transfer function of a filter described by its poles and zeros and the motion equation of a piston. An example of a non-parametric model is the description of a filter by its impulse response at a large number of points.

Usually it is simpler to create a non-parametric model than a parametric one because the modeller needs less knowledge about the system itself in the former case. However, physical insight and concentration of information are more substantial for parametric models than for non-parametric ones.

- White box models versus black box models

In the construction of a model, physical laws whose availability and applicability depend on the insight and skills of the experimenter can be used (Kirchhoff's laws, Newton's laws, etc.) Specialized knowledge related to different scientific fields may be brought into this phase of the identification process. The modelling of a loud speaker, for example, requires extensive understanding of mechanical, electrical and acoustical phenomena. The result may be a physical model, based on comprehensive knowledge of the internal functioning of the system. Such a model is called a white box model.

The choice between the different methods depends on the aim of the study: the white box approach is better for gaining insight into the working principles of a system, but black box model may be sufficient if the model will be used only for prediction of the output.

Although, as a rule of thumb, it is advisable to include as much prior knowledge as possible during the modelling process, it is not simple to express this information if the polynomial coefficients are used as parameters.

- Linear models versus non-linear models.

In real life, almost every system is non-linear. It is because mostly approximated by linear models, assuming that in the operation region the behaviour can be linearized. This kind of approximation makes it possible to use simple models without jeopardizing properties that are of importance to the modeller. This choice depends strongly on the intended use of the model. For example, a non-linear model is needed to describe the distortion of an amplifier but a linear model will be sufficient to represent its transfer characteristics if the linear behaviour is dominant and is the only interest.

- Linear in the parameter versus non-linear in the parameters

A model is called linear in the parameter if there exists a linear relation between these parameters and the error that is minimized. This does not imply that the system itself is linear. For example $\varepsilon = y - (a_1 u + a_2 u^2)$ is linear in the parameter a_1 and a_2 but describes a non-linear system.

$$\mathcal{E}(j\omega) = \mathrm{Y}(j\omega) - \frac{a_0 + a_1 j\omega}{b_0 + b_1 j\omega} U(j\omega) \tag{2.1}$$

Equation 2.1 describe a linear system but the model is non-linear in the b_1 and b_2 parameters. Linearity in the parameters is a very important aspect of models because it has a strong impact on the complexity of the estimators if a (weighted) least squares cost function is used. In that case, the problem can be solved analytically for models that are linear in the parameters so that an iterative optimization problem is avoided.

2.1.3 Criterion Of Fit

Once a model structure is chosen (e.g., a parametric function model), it should be matched as well as possible with the available information about the system. Mostly, this is done by minimizing a criterion that measures a goodness of the fits. The choice of this criterion is extremely important because it determines the stochastic properties of the final estimator. For the example like resistance, many choice are possible and each of them can lead to a different estimator with its owns properties. Usually, the cost function defines a distance between the experiment data and the model.

2.1.4 Validation

Finally, the validity of the selected model should be tested: does this model describe the available data properly or are there still indications that some of the data are not well modelled, indicating remaining model errors? In practice, the best model (the smallest errors) is always preferred. Some tools that guide the user through this process by separating the remaining errors into different classes, for example, modelled linear dynamics and non-linear distortions. From this information, further improvement of the model can be proposed, if necessary.

During the validation tests it is always important to keep the application in mind. The model should be tested under the same conditions as it will be used later. Extrapolation should be avoided as much as possible. The application also determines what properties are critical.

2.1.5 System Identification Step Overview

This brief overview of the identification process shows that it is a complex task with a number of interacting choices. It is important to pay attention to all aspects of this procedure, from the experiment design to the model validation, in order to get the best results. The reader should be aware that besides these actions, other aspects are also important. A short inspection of the measurement setup can reveal important shortcomings that can jeopardize a lot of information.

Good understanding of the intended applications helps to set up good experiments, and is very important to make the proper simplifications during the model building process. Many times, choices are made that are not based on complicated theories but are dictated by the practical circumstances. In these cases a good theoretical understanding of the applied methods will help the user to be aware of the sensitive aspect of his techniques. This will enable him to put all his effort on the most critical decisions. Moreover, he will become aware of the weak points of the final model.

2.2 Type of Model

Model is a relationship between observed quantities. In loose terms, a model allows for prediction of properties or behaviours of the object. Typically the relationship is a mathematical expression, but it could also be a table or a graph.