11

System Identification

Lennart Ljung¹

ABSTRACT
In this contribution we give an overview and discussion of the basic steps of System Identification. The four main ingredients of the process that takes us from observed data to a validated model are: (1) The data itself, (2) The set of candidate models, (3) The criterion of fit and (4) The validation procedure. We discuss how these ingredients can be blended to a useful mix for model-building in practice.

11.1 Introduction

The process of going from observed data to a mathematical model is fundamental in science and engineering. In the control area this process has been termed "System Identification" and the objective is then to find dynamical models (difference or differential equations) from observed input and output signals. Its basic features are however common with general model building processes in statistics and other sciences.

System Identification has been an active research area for more than thirty years. It has matured and many of the techniques have become standard tools in control and signal processing engineering. The "mainstream approach" is described e.g. in [8] and [19]. Over the past few years there has been a significantly renewed interest in the area with topics like "unknown-but-bounded" disturbances, [17, 11], set membership techniques [4, 14] subspace techniques [15], $H_{\infty}$-identification [16, 6], worst case analysis [5, 10], as well as how to deal with unmodelled dynamics [13].

The procedure is characterized by four basic ingredients:

1. The observed data
2. A set of candidate models
3. A criterion of fit
4. Validation

We shall in the sequel discuss these items more closely.

¹Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden, E-mail: Ljung@isy.liu.se
FIGURE 11.1. Results from test flights of the new Swedish aircraft JAS-Gripen, developed by SAAB Military Aircraft AB, Sweden. From the top a) Pitch rate. b) Elevator angle. c) Canard angle. d) Leading edge flap

11.2 The Data

The area of system identification begins and ends with real data. Data required to build models and to validate models. The result of the modelling process can be no better than what corresponds to the information contents in the data.

Let us take a look at two data sets:

Example 1 An unstable aircraft. Figure 11.1 shows some results from test flights of the new Swedish aircraft JAS-Gripen, developed by SAAB Military Aircraft AB, Sweden. The problem is to use the information in these data to determine the dynamical properties of the aircraft for fine-tuning regulators, for simulations, and so on. Of particular interest are the aerodynamical derivatives.

Example 2 Vessel dynamics. See Figure 11.2. The problem is to determine the residence time in the buffer vessel. The pulp spends about 48 hours total in the process, and knowing the residence time in the different vessels is important in order to associate various portions of the pulp with the different chemical actions that have taken place in the vessel at different times. (The $K$-number is a quality property that in this context can be seen as a marker allowing us to trace the pulp.)

11.3 The Set of Models: Model Structures

The single most important step in the identification process is to decide upon a model structure, i.e., a set of candidate models. In practice typically a whole lot of them are tried out and the process of identification really becomes the process of evaluating and choosing between the resulting models in these different structures.