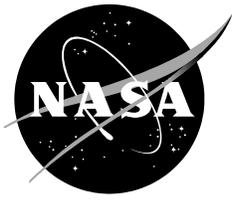


NASA/TM—2011–215966 (Corrected Copy)



# **Data Driven Model Development for the SuperSonic SemiSpan Transport (S<sup>4</sup>T)**

*Sunil L. Kukreja*

*Dryden Flight Research Center, Edwards, California*

This document was incorrectly issued as report number 216432; the correct number is 215966.

---

**March 2011**

## NASA STI Program ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

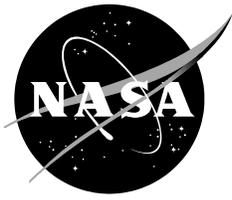
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing help desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to [help@sti.nasa.gov](mailto:help@sti.nasa.gov)
- Fax your question to the NASA STI Help Desk at 443-757-5803
- Phone the NASA STI Help Desk at 443-757-5802
- Write to:  
NASA STI Help Desk  
NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320

NASA/TM—2011–215966 (Corrected Copy)



# Data Driven Model Development for the SuperSonic SemiSpan Transport (S<sup>4</sup>T)

*Sunil L. Kukreja*

*Dryden Flight Research Center, Edwards, California*

This document was incorrectly issued as report number 216432; the correct number is 215966.

National Aeronautics and  
Space Administration

*Dryden Flight Research Center  
Edwards, CA 93523-0273*

---

**March 2011**

Available from:

NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320  
443-757-5802

## Abstract

In this report, we will investigate two common approaches to model development for robust control synthesis in the aerospace community; namely, reduced order aeroservoelastic modelling based on structural finite-element and computational fluid dynamics based aerodynamic models, and a data-driven system identification procedure. It is shown via analysis of experimental SuperSonic SemiSpan Transport (S<sup>4</sup>T) wind-tunnel data that by using a system identification approach it is possible to estimate a model at a fixed Mach, which is parsimonious and robust across varying dynamic pressures.

## Nomenclature

$a$	parameter of the output
$A(q)$	polynomial in $a$
AIC	Akaike's information criterion
ASE	aeroservoelastic
ARMA	AutoRegressive, Moving Average
ARMAX	AutoRegressive, Moving Average eXogenous
ARX	AutoRegressive eXogenous input
$b$	parameter of the input
$B(q)$	polynomial in $b$
BJ	Box Jenkins
$c$	parameter of the noise model numerator
$C(q)$	polynomial in $c$
$D(q)$	polynomial in $d$
$e(n)$	an unobservable white-noise disturbance
$F(q)$	polynomial in $f$
FE	finite element
FIR	finite impulse response function
FPE	final prediction error
$G(q)$	$B(q)/A(q)$
$H(q)$	$1/A(q)$
$H_\infty$	H-infinity
MDL	minimal description length
$n_a$	previous values of the output
$n_b$	previous values of the input
$N$	number of data points
$p$	number of model parameters
PEI	prediction error identification
POD	proper orthogonal decomposition
$Q$	dynamic pressure
ROM	reduced order model
S <sup>4</sup> T	SuperSonic SemiSpan Transport
TPWL	trajectory piecewise linearization
$u(n)$	an accessible input
$V$	the prediction error or the residual sum of squares
$y$	measured system output
$\hat{y}$	predicted model output
$y(n)$	measured output
%QF	percent quality of fit

# 1. Introduction

Undesirable aeroservoelastic (ASE) interactions are a major concern in modern aircraft design. ASE interactions between aircraft structure, aerodynamics, and flight control systems can lead to divergent oscillations resulting in catastrophic failure (ref. 1). As such, analytical model development is an important step in the design and certification of aircraft. Accurate models allow for robust control design, which is critical for aircraft safety, gust-load alleviation, ride quality, et cetera.

Finite element (FE) based models are used in the design process of aircraft to aid in the description of complex elastic and structural components (refs. 2–3). FE based models that accurately characterize the aerodynamic and structural components are of very high order (for example, thousands of degrees of freedom) and computationally intensive. To reduce the ASE model order, a modal approach is used. This approach can reduce the state-space model order to several tens of states. The robustness of this model type highly depends on the accuracy of the FE structural and aerodynamic models, and on the number of states applied to the modelling. Models of high complexity inhibit their use for control synthesis because their real-time implementation is difficult or not possible (ref. 4). This difficulty has led to considerable activity in the areas of model and controller reduction techniques in the last decade. The literature is rich with many reduced order model (ROM) techniques, which would require a review paper to properly discuss them. As such, in this report, we limit ourselves to a handful of approaches.

A leading strategy to FE based model reduction is the proper orthogonal decomposition (POD), which is also known as the Karhunen-Loeve procedure (refs. 5). The so-called POD technique is well known in the statistical literature as principal-component analysis (ref. 6). The reduced basis method was first proposed in references 7 and 8 for structural analysis, and it has been used for structural problems in references 9 through 11. This technique uses synthetic data from a high fidelity FE based model to capture the dominant characteristic information utilising an orthogonalisation process. This allows the POD approach to accurately describe a system using a few basis terms, which gives it an advantage compared to other numerical procedures (ref. 12). These reasons have allowed POD to become a popular technique for the implementation of real-time control (ref. 13). Moreover, POD has been successfully used in a variety of fields including signal analysis and pattern recognition (ref. 14), fluid dynamics and coherent structures (refs. 15–17), control theory (refs. 18–20), civil engineering (ref. 21) and inverse problems (ref. 22).

More recently the ROM techniques utilising POD were developed for aeroelastic systems analysis (ref. 23). This methodology was introduced to the aerospace community for the reduction of aeroelastic equations (ref. 24). As a follow-on to this technique, frequency-domain approaches were developed which efficiently compute POD basis functions for linearised aeroelastic systems (refs. 25–26). Due to its popularity and utility, POD methodology has been proposed and implemented for static and dynamic continuous-time nonlinear aeroelastic problems. Subsequent developments lead to extensions to encompass discrete Euler equations (refs 27–28). Successful application of this approach has been demonstrated for the analysis of limit-cycle oscillation of an airfoil with a nonlinear structural coupling in the transonic regime (ref. 29).

Nevertheless, it has been observed that standard POD procedures are less robust for nonlinear problems and typically require more basis functions as the function complexity increases (refs. 23 and 30). To address this limitation, a linearisation strategy, the trajectory piecewise linearisation (TPWL), was proposed (refs. 31–32). The TPWL technique combines reduced-order modelling with linearisation of the governing equations as a solution to this problem (ref. 23).

Although POD offers a significant reduction of the full FE based model, it is often too large to lend itself to efficient control design. An alternative approach is to use data driven techniques to let the data dictate what the optimal model should be. Such a procedure is commonly known as system identification.

This area, as with ROM methods, also has an extensive base of literature for both linear and nonlinear system identification techniques and is too large to give a proper review here. As such we refer the reader to an often-cited authoritative treatise in the area, which provides an excellent overview and references (ref. 33). Below, we provide a brief and incomplete introduction.

There are two broad classes of techniques that can be pursued to accomplish the task of system identification: (1) nonparametric and (2) parametric methods. The finite impulse response function (FIR) has been widely used for modelling linear time-invariant systems. This type of system description is known as nonparametric because it is a numeric representation of the system's impulse response or kernel (refs. 33–35). Although nonparametric methods can be used to represent many classes of systems, they do so at the expense of introducing an excessive number of unknown coefficients, which must be estimated. Most expansions map the past inputs into the present output and so require a very large number of coefficients to characterize the process. Moreover, the parameters are not readily linked to the underlying system, except in special cases where significant a priori knowledge of the system has been assumed.

Due to this shortcoming, parametric identification methods have been developed for use in the design of better control systems. Parametric models have some advantages in applications. They: (1) are easier to understand and interpret, (2) can simplify forecasts (e.g., obtaining forecast intervals), and (3) model comparison in a parametric context (i.e., parameter estimates, model order, and model structure) has been well studied. Hence, the difficulty of model comparison encountered using nonparametric tools can be avoided (ref. 36).

A parametric model consists of a set of differential or difference equations describing the system dynamics. Such equations usually contain a “small” number of parameters, which can be varied to alter the behavior of the equation. In this report, we only consider the discrete-time case since in any practical experimental situation the data available to the experimenter is in discrete-time. As such, most systems for identification purposes are represented in discrete-time. In addition, we assume the ROM with which we compare our data-driven model is available and as such we do not discuss its development.

The organization of this report is as follows. In section 2 we formulate the identification problem addressed here. Section 3 describes the experimental SuperSonic SemiSpan Transport (S<sup>4</sup>T) test-bed (ref. 37) and methods used for model development. Section 4 illustrates the results of our study on three flight conditions; namely, model robustness to fixed Mach but varying dynamic pressure. Section 5 provides a discussion of our findings, and section 6 summarises the conclusions of our study.

## **2. Problem Statement**

System identification is the process of developing or improving a mathematical representation of a physical system based on observed data. Often the observed data consists of an external user selected input, used to perturb the system and elicit an output response. In any experimental situation the system output is a sum of the true unknown system output and observation or measurement noise. Given this paradigm there are several model structures that can be explored to model a system's dynamics and noise.

### **2.1 Model Structures**

After preprocessing the recorded data, the first step in system identification is to select a model structure to describe the observations. There are a number of model forms to select from when developing

a data-driven model (ref. 33). However, the problem often dictates the model form(s) that are reasonable to consider. This insight may come from a priori knowledge of the physical process or previous morphological modelling efforts. In this report, we use knowledge gained from previous work to limit the model sets considered; namely, linear, time-invariant processes where both input-output are available to the user (38–39).

### 2.1.1 ARX Model Structure

The simplest input-output polynomial model is the AutoRegressive eXogenous (ARX) input model, represented in equation (1) as (refs. 33, 40, and 41):

$$\begin{aligned} y(n) = & - a_1 y(n-1) - \dots - a_{n_a} y(n-n_a) \\ & + b_1 u(n-1) + b_2 u(n-2) + \dots + b_{n_b} u(n-n_b) \\ & + e(n) \end{aligned} \quad (1)$$

where  $y(n)$  is the measured output,  $u(n)$  is an accessible input, and  $e(n)$  is an unobservable white-noise disturbance. The current output depends on  $n_b$  previous values of the input,  $n_a$  previous values of the output, and the current disturbance.

This structure can be represented more compactly as shown in equations (2), (3), and (4):

$$A(q)y(n) = B(q)u(n) + e(n) \quad \text{or} \quad (2)$$

$$y(n) = G(q)u(n) + H(q)e(n) \quad \text{where} \quad (3)$$

$$G(q) = \frac{B(q)}{A(q)}, H(q) = \frac{1}{A(q)} \quad (4)$$

where  $A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}$ ,  $B(q) = b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}$ ,  $q^{-1}$  is the backward shift operator, and  $a$ 's and  $b$ 's are the parameters of the output and input, respectively.

When there is evidence of significant noise in the system a more flexible noise model may be required to model the dynamics and noise using different polynomials.

### 2.1.2 ARMAX Model Structure

For linear systems, the relationship between input-output and noise can be written as a linear difference equation, and is shown in equation (5):

$$\begin{aligned} y(n) = & - a_1 y(n-1) - \dots - a_{n_a} y(n-n_a) \\ & + b_1 u(n-1) + b_2 u(n-2) + \dots + b_{n_b} u(n-n_b) \\ & + e(n) + c_1 e(n-1) + \dots + c_{n_c} e(n-n_c) \end{aligned} \quad (5)$$

This is known as the AutoRegressive, Moving Average eXogenous (ARMAX) model. In this model structure the current output,  $y(n)$ , depends on an exogenous input,  $u(n)$ , an innovation process,  $e(n)$ , and past values of the output. This structure can be represented more compactly as shown in equation (6):

$$A(q)y(n) = B(q)u(n) + C(q)e(n) \quad (6)$$

Substituting equation (6) into equation (3) yields equation (7):

$$G(q) = \frac{B(q)}{A(q)}, H(q) = \frac{C(q)}{A(q)} \quad (7)$$

where  $C(q) = 1 + c_1q^{-1} + \dots + c_{n_c}q^{-n_c}$ , and the  $c$ 's are parameters of the noise model numerator. The extra polynomial,  $C(q)$ , gives the ARMAX structure additional flexibility to model additive disturbance. When additional complexity is needed to model noise  $H(q)$  can be fully parameterised independent of the system dynamics.

### 2.1.3 Box–Jenkins Model Structure

A natural development of the ARMAX model structure is to parameterise the noise process as an AutoRegressive, Moving Average (ARMA) model as shown in equation (8):

$$y(n) = \frac{B(q)}{F(q)}u(n) + \frac{C(q)}{D(q)}e(n) \quad (8)$$

where  $F(q) = 1 + f_1q^{-1} + \dots + f_{n_f}q^{-n_f}$ , and  $D(q) = 1 + d_1q^{-1} + \dots + d_{n_d}q^{-n_d}$  model the poles of the system and noise separately. Note that,  $n_f = n_a$ . This model structure is known as the Box-Jenkins (BJ) model due to their seminal work proposing this model form (ref. 42).

## 2.2 Model Order Selection

Once a model structure is selected, the dynamic order of the system needs to be determined. Although there are many techniques that offer a solution to this problem, below we describe three commonly used techniques to estimate model order.

### 2.2.1 Final Prediction Error

The Final Prediction Error (FPE) measure estimates the error in model fit when it is used to predict new outputs (ref. 43). The FPE defines an optimal model as one that minimises as shown in equation (9):

$$\text{FPE} = V \left( 1 + \frac{2p}{N - p} \right) \quad (9)$$

where  $N$  is the number of data points,  $V$  is the prediction error, or the residual sum of squares, and  $p$  is the number of model parameters.

### 2.2.2 Akaike's Information Criterion

Akaike's Information Criterion (AIC) is a weighted estimation error based on the unexplained variation of a given time series with a penalty term when exceeding the optimal number of parameters to represent the system (ref. 44). Utilising AIC, an optimal model is defined as one that minimises as shown in equation (10):

$$AIC = V \left( 1 + \frac{2p}{N} \right) \quad (10)$$

According to Akaike's theory, the most accurate model has the smallest prediction error.

### 2.2.3 Minimal Description Length

Rissanen's Minimal Description Length (MDL) approach is based on  $V$  plus a penalty for the number of terms used (ref. 45). With MDL, an optimal model is one that minimises as shown in equation (11):

$$MDL = V \left( 1 + \frac{p \ln N}{N} \right) \quad (11)$$

A model that minimises the MDL allows the shortest description of measured data.

## 2.3 Cross-Validation

Model validation is an important step in developing strategies for robust control. This step is typically preceded by system identification. Model validation is concerned with assessing whether a given nominal model can reproduce data from a plant, collected after some initial experiments to obtain estimation data (ref. 46). The model validation problem is really one of model invalidation since a given model can only be said to be not invalidated with the current evidence. Future evidence may invalidate the model.

We cross-validate the parameter estimates of the model dynamics using a 1-step-ahead predictor (ref. 33). Model goodness is assessed by computing the percent quality of fit (%QF) as shown in equation (12):

$$\%QF = \left( 1 - \frac{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2}{\frac{1}{N} \sum_{n=1}^N (y_n)^2} \right) \times 100 \quad (12)$$

where  $y$  is the measured system output, and  $\hat{y}$  the predicted model output.

In the sequel, we show that by implementing these well-known identification strategies it is possible to develop a model, which is parsimonious and a robust predictor of measured data and, hence, represent the physical process more accurately.

## 3. Experimental S<sup>4</sup>T Wind-Tunnel Data

The modelling and identification techniques were applied to experimental wind-tunnel data from the S<sup>4</sup>T project conducted at NASA Langley Research Center (Ref. 37). Figure 1 shows a scale model of the S<sup>4</sup>T test bed in the wind tunnel.



Figure 1. S<sup>4</sup>T model.

The data analysed for this study used horizontal tail position input and structural accelerometer response output.

### 3.1 Data Collection

Wind-tunnel data was gathered during transonic clearance of the S<sup>4</sup>T. At each flight condition the aircraft model was perturbed with a log sine sweep input which had a frequency content of 0.5–25 Hz, mean value of 3.5 deg, and  $\pm 0.3$  deg amplitude. The inputs were applied to the horizontal tail. The accelerometer output was collected from a sensor located at the nacelle inboard aft position. Wind-tunnel tests were conducted at subsonic, transonic, and supersonic conditions; and varying dynamic pressures. The input-output were antialiasing, filtered by an eighth-order Bessel filter with a cut-off at 200 Hz and recorded at 1,000 Hz. Data was collected for a range of Mach numbers from  $M = 0.6$  to 1.2 and dynamic pressures  $Q = 20$  to 65 psf.

### 3.2 Data Analysis

In this study, the objective was to develop a parsimonious model for control synthesis to improve gust-load alleviation and ride quality. As such, we focused our identification efforts on data collected at Mach 0.80, 0.95, and 1.10; and  $Q = 30, 55, 60,$  and 65 psf. At each Mach we used  $Q = 30$  psf as the estimation data; and  $Q = 55, 60,$  and 65 psf as the cross-validation data. At  $M = 0.95$  data was not collected at  $Q = 65$  psf, and, therefore, the estimated model could only be cross-validated at  $Q = 55$  and 60 psf. This approach of modeling and validation was taken to assess the model's predicative capability and robustness.

Data was preprocessed to remove the linear trend, mean, and outliers. The preprocessing step ensured that all unwanted low-frequency disturbances, offsets, trends, and drifts were removed and allowed for an accurate representation of the system dynamics.

### 3.3 Identification Procedures

The identification process was performed in four stages to assess the full range of models discussed in section 2. In all cases the model order was estimated using the FPE, AIC, and MDL approaches. The optimal model order was deemed as one that produced the lowest prediction error of the aforementioned techniques. Once the model order was fixed, prediction error identification (PEI) was used to estimate the unknown parameters, and its predictive capability was computed for a cross validation set as described previously (see section 2.3).

1. The ARX structure's ability to model wind-tunnel data was assessed because it is the simplest input-output model invoking the principle of parsimony. For the ARX model the order was determined by preselecting a range of model orders to search over. Specifically,  $n_a = 2-5$ ,  $n_b = 1-5$  and  $n_k = 1-10$  was chosen as the search range, where  $n_k$  denotes input delay. This range was selected to allow sufficient model complexity whilst maintaining an efficient system description.
2. Further complexity was added to the ARX model to assess whether the same model structure could account for significantly greater output variance with  $n_a = 2-10$ ,  $n_b = 1-10$  and  $n_k = 1-20$ .
3. Using the best-fit ARX model order to fix  $n_a$  and  $n_b$ , we searched for an ARMAX order, namely,  $n_c = 1-20$  that provided the smallest prediction error as the optimal ARMAX order.
4. The same procedure was followed to assess whether the BJ model structure could account for more of the output variance than the ARMAX model with  $n_d = 1-20$ .

## 4. Results

### 4.1 ARX Model

Two ARX model orders were evaluated. Specifically, a model with order  $n_a = 5$ ,  $n_b = 1$ ,  $n_k = 1$  and  $n_a = 10$ ,  $n_b = 1$ ,  $n_k = 1$  were used for model development. The model order was estimated as discussed in section 2.2. This analysis allowed for us to assess whether adding complexity to the same structure could account for significantly more of the output variance. The models were estimated using wind-tunnel data measured at Mach 1.10,  $Q = 30$  psf which yielded structures of the form shown in equations (13) and (14).

ARX5:

$$\hat{A}(q)y(n) = \hat{B}(q)u(n) + e(n) \quad \text{where}$$

$$\hat{A}(q) = 1 + \hat{a}_1q^{-1} + \hat{a}_2q^{-2} + \hat{a}_3q^{-3} + \hat{a}_4q^{-4} + \hat{a}_5q^{-5} \quad \text{and}$$

$$\hat{B}(q) = \hat{b}_1q^{-1} \tag{13}$$

ARX10:

$$\tilde{A}(q)y(n) = \tilde{B}(q)u(n) + e(n) \quad \text{where}$$

$$\begin{aligned} \tilde{A}(q) = & 1 + \tilde{a}_1q^{-1} + \tilde{a}_2q^{-2} + \tilde{a}_3q^{-3} + \tilde{a}_4q^{-4} + \tilde{a}_5q^{-5} \\ & + \tilde{a}_6q^{-6} + \tilde{a}_6q^{-6} + \tilde{a}_6q^{-6} + \tilde{a}_6q^{-6} + \tilde{a}_{10}q^{-10} \quad \text{and} \end{aligned}$$

$$\tilde{B}(q) = \tilde{b}_1q^{-1} \quad (14)$$

Figures. 2a–2c shows representative results for the ARX structure’s (eqs. (13) and (14)) ability to represent S<sup>4</sup>T wind-tunnel data by evaluating it with cross-validation data at Mach 1.10, and  $Q = 55, 60, 65$  psf.

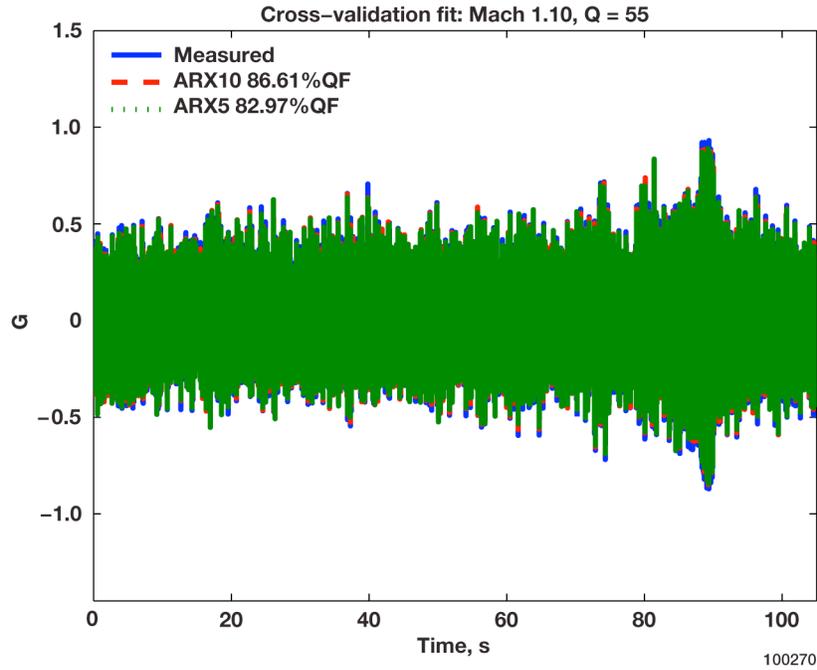


Figure 2(a). Cross-validation fit: Mach 1.10, Q = 55 psf.

Figure 2. Measured and predicted output for ARX5 and ARX10 models at Mach 1.10; (a):  $Q = 55$  psf, (b):  $Q = 60$  psf, and (c):  $Q = 65$  psf. Solid line (“—”) measured output. Dash-dash line (“- -”) predicted ARX10 model output. Dot-dot line (“...”) predicted ARX5 model output.

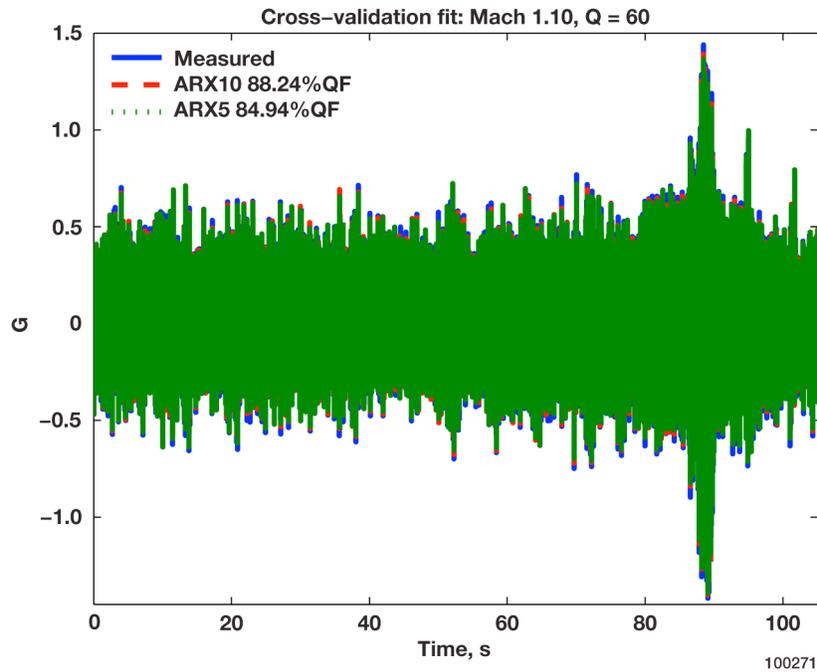


Figure 2(b). Cross-validation fit: Mach 1.10,  $Q = 60$  psf.

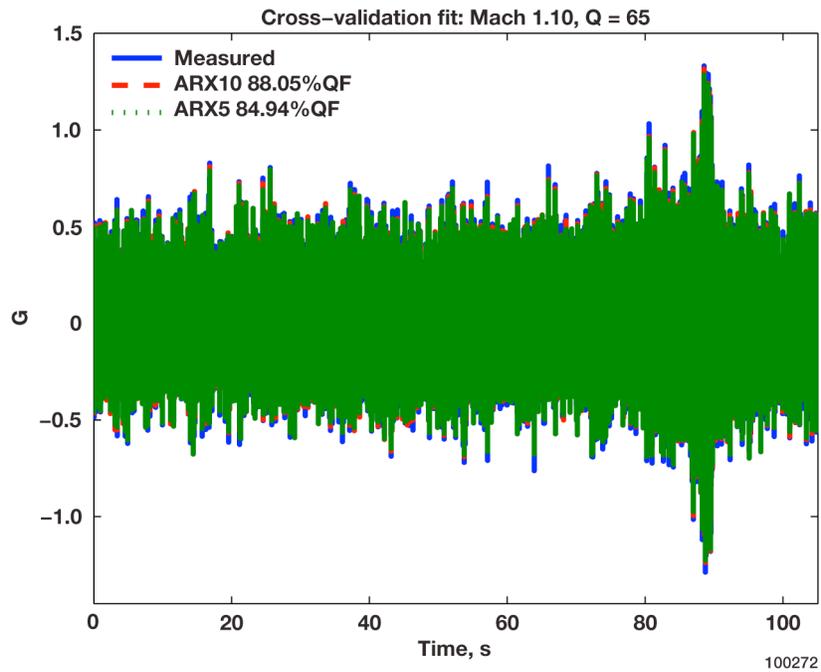


Figure 2(c). Cross-validation fit: Mach 1.10,  $Q = 65$  psf.

Figure 2. Concluded.

This figure compares %QF of the predicted output for the ARX5 and ARX10 models superimposed on top of measured data. The %QF's obtained for the ARX5 model at  $Q = 55, 60, 65$  psf are 82.97%,

84.94% and 84.94%, respectively. For the ARX10 model the QF's percentage are 86.61%, 88.24 and 88.05%.

Figures 2a–2c demonstrate that adding complexity to the ARX model improves the %QF. As such, we deem that the ARX10 model (eq. (14)) is the better fit model due to its predicative capability.

## 4.2 ARMAX Model

Next, we used the ARX10 model order as a starting point to develop a more complex ARMAX model to assess whether it could account for more output variance whilst maintaining an efficient model description. The model was estimated at Mach 1.10,  $Q = 30$  psf and yielded a structure of the form shown in equation (15):

ARMAX:

$$\tilde{A}(q)y(n) = \tilde{B}(q)u(n) + \tilde{C}(q)e(n) \quad \text{where}$$

$$\begin{aligned} \tilde{C}(q) = & 1 + \tilde{c}_1q^{-1} + \tilde{c}_2q^{-2} + \tilde{c}_3q^{-3} + \tilde{c}_4q^{-4} + \tilde{c}_5q^{-5} \\ & + \tilde{c}_6q^{-6} + \tilde{c}_7q^{-7} + \tilde{c}_8q^{-8} + \tilde{c}_9q^{-9} + \tilde{c}_{10}q^{-10} \end{aligned} \quad (15)$$

Figures 3a–3c shows representative results for the ARMAX structure's (eq. (15)) ability to represent S<sup>4</sup>T wind-tunnel data by evaluating it with cross-validation data at Mach 1.10, and  $Q = 55, 60, 65$  psf.

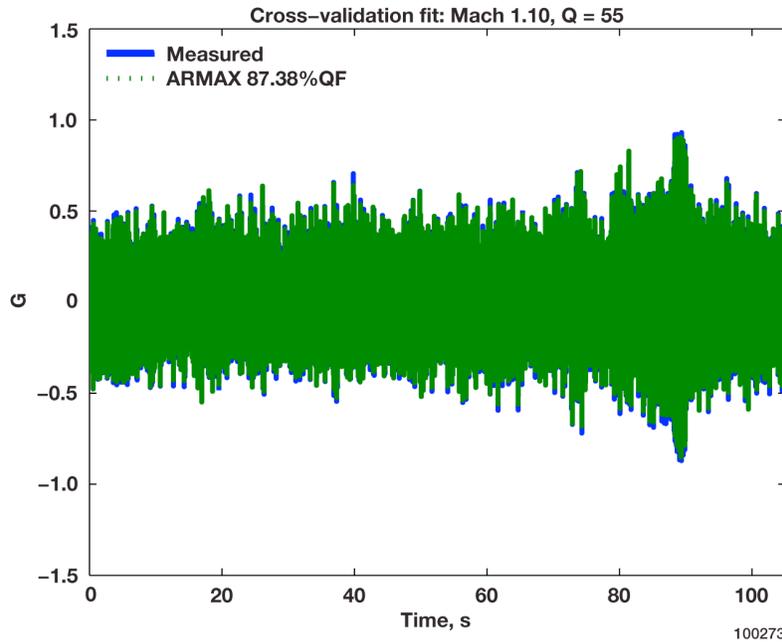


Figure 3(a). Cross-validation fit: Mach 1.10,  $Q = 55$  psf.

Figure 3. Measured and predicted output for ARMAX model at Mach 1.10; (a):  $Q = 55$  psf, (b):  $Q = 60$  psf, and (c):  $Q = 65$  psf. Solid line (“—”) measured output. Dot-dot line (“...”) predicted ARMAX model output.

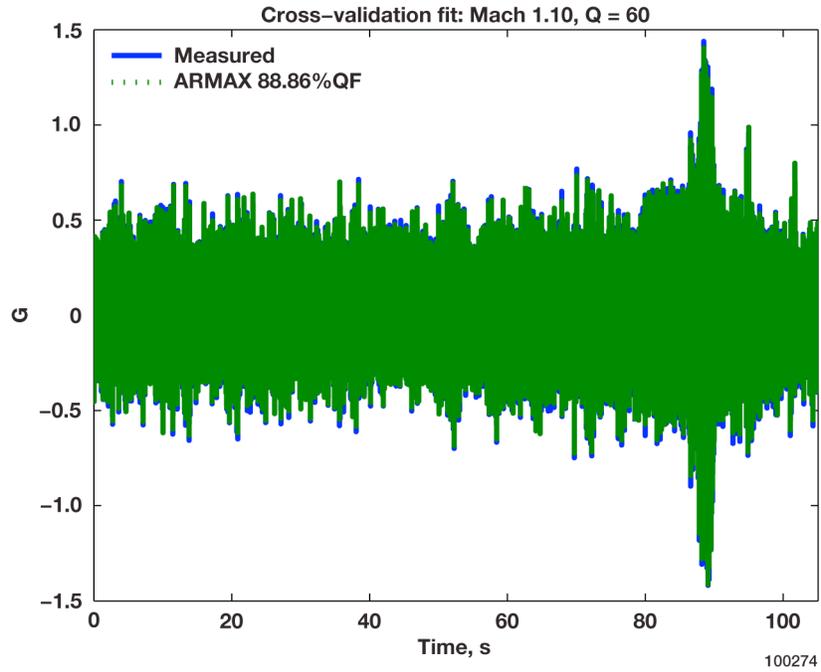


Figure 3(b). Cross-validation fit: Mach 1.10, Q = 60 psf.

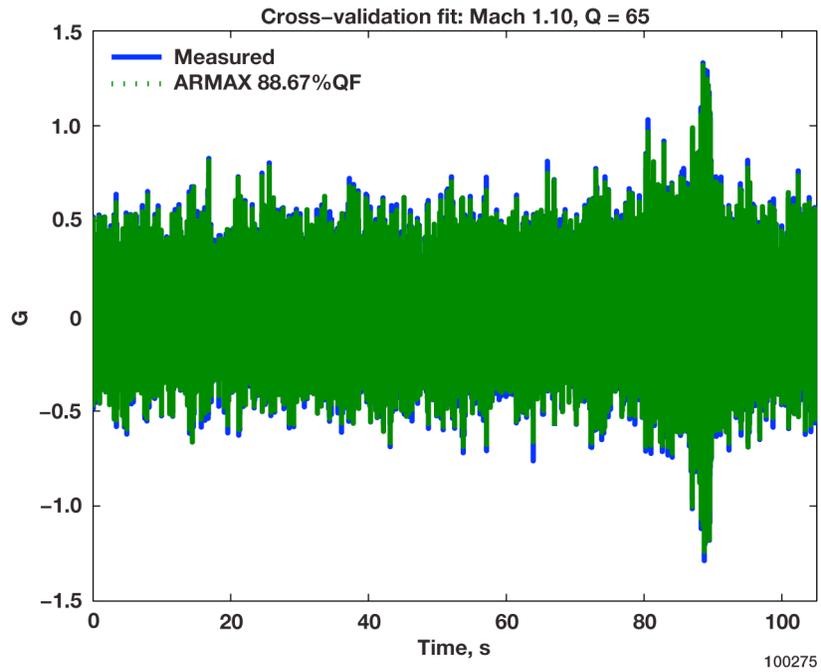


Figure 3(c). Cross-validation fit: Mach 1.10, Q = 65 psf.

Figure 3. Concluded.

Figures 3a–3c illustrate %QF of the predicted output for the ARMAX model superimposed on top of measured data. The %QF’s obtained for the ARMAX model at  $Q = 55, 60, 65$  psf are 87.38%, 88.86%,

and 88.67%, respectively. Notice that although this model structure adds complexity, it accounts for incrementally more of the output variance.

### 4.3 Box-Jenkins Model

Lastly, we used the ARMAX model (eq. (15)) order as a starting point to develop a more complex BJ model to assess whether it could account for more output variance whilst maintaining an efficient model description. The model was estimated at Mach 1.10,  $Q = 30$  psf, and yielded a structure of the form shown in equation (16):

BJ:

$$y(n) = \frac{\tilde{B}(q)}{\tilde{F}(q)}u(n) + \frac{\tilde{C}(q)}{\tilde{D}(q)}e(n) \quad \text{where}$$

$$\tilde{F}(q) \approx \tilde{A}(q) \quad \text{and}$$

$$\begin{aligned} \tilde{D}(q) = & 1 + \tilde{d}_1q^{-1} + \tilde{d}_2q^{-2} + \tilde{d}_3q^{-3} + \tilde{d}_4q^{-4} + \tilde{d}_5q^{-5} \\ & + \tilde{d}_6q^{-6} + \tilde{d}_7q^{-7} + \tilde{d}_8q^{-8} + \tilde{d}_9q^{-9} + \tilde{d}_{10}q^{-10} \end{aligned} \quad (16)$$

Figures 4a–4c shows representative results for the BJ structure's (eq. (16)) ability to represent S<sup>4</sup>T wind-tunnel data by evaluating it with cross-validation data at Mach 1.10, and  $Q = 55, 60, 65$  psf.

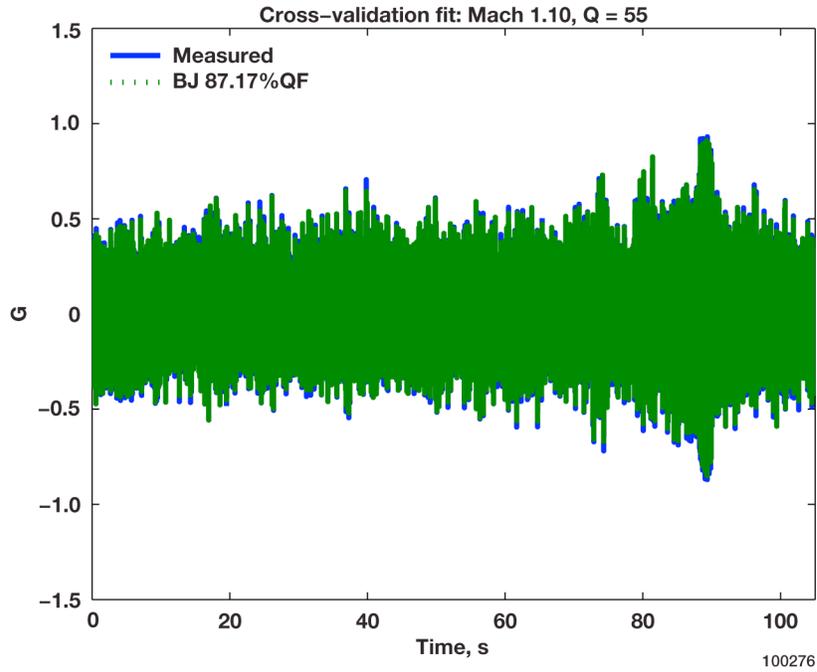


Figure 4(a). Cross-validation fit: Mach 1.10, Q = 55 psf.

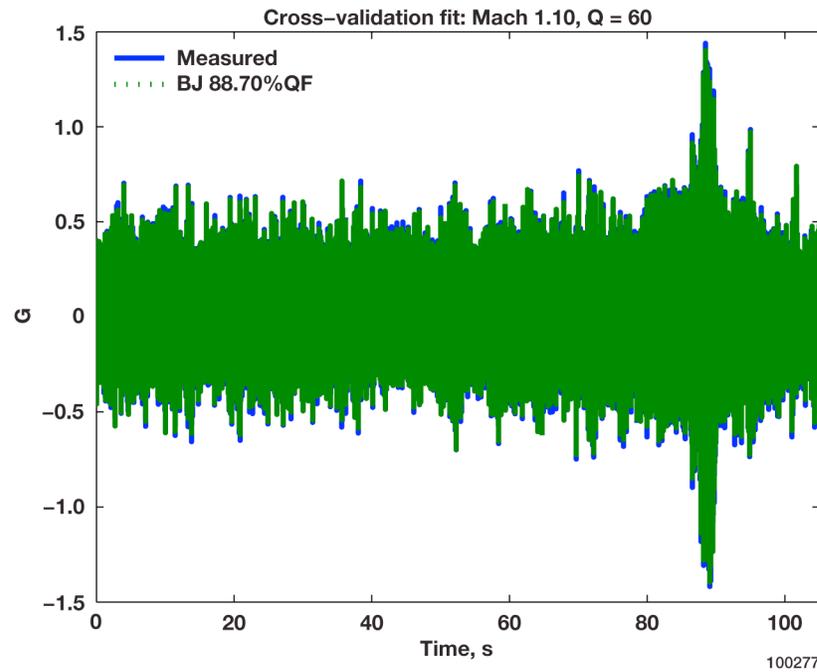


Figure 4(b). Cross-validation fit: Mach 1.10, Q = 60 psf.

Figure 4. Predicted output of the BJ model at Mach 1.10 superimposed on top of measured output; (a): Q = 55 psf, (b): Q = 60 psf, and (c): Q = 65 psf. Solid line (“—”) measured output. Dot-dot line (“...”) predicted BJ model output.

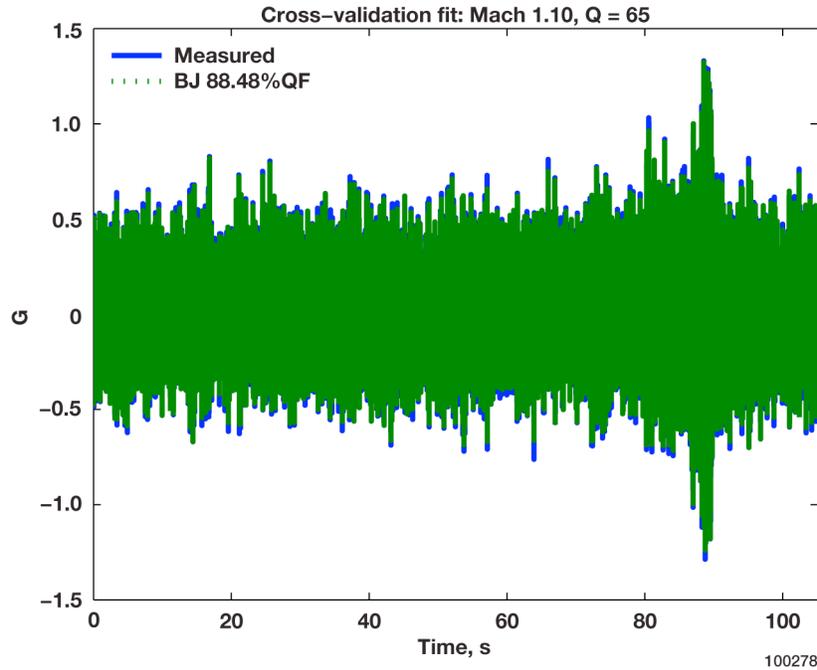


Figure 4(c). Cross-validation fit: Mach 1.10, Q = 65 psf.

Figure 4. Concluded.

Figures 4a–4c illustrate %QF of the predicted output for the BJ model superimposed on top of measured data. The %QF’s obtained for the BJ model at  $Q = 55, 60, 65$  psf are 87.17%, 88.70%, and 88.48%, respectively. Although the BJ model offers greater complexity to model the observed data, it offers a slightly lower %QF. Using validation data, if the fit of a higher order deteriorates, it is an indication that the model complexity is too high (ref. 33). Table 1 summarises the findings of our analysis.

Table 1. Summary of cross-validation results with Mach number and dynamic pressure versus model structure.

Mach	Dynamic pressure, psf	Quality of fit, %			
		ARX5	ARX10	ARMAX	BJ
0.80	55	86.04	87.63	88.26	88.05
	60	86.81	88.25	88.87	88.61
	65	86.16	87.33	87.79	87.66
0.95	55	83.31	85.56	86.80	86.05
	60	83.60	85.97	87.11	86.36
	65	X	X	X	X
1.10	55	82.97	86.61	87.38	87.17
	60	84.94	88.24	88.86	88.70
	65	84.94	88.05	88.67	88.48

The results from table 1 illustrate that model fit increases with added complexity for the ARX model structure, improves incrementally for the ARMAX structure, but decreases for the BJ. From these results we conclude that the higher-order ARX model may be sufficient for our purposes, and as such we

compare the predictive capability of this model to the FE based ASE model used for control design during wind-tunnel test (refs. 38–39).

#### 4.4 Comparison of Tenth–Order ARX and ASE Models

Lastly, we compared the ARX10 (eq. (14)), and ASE model’s ability to predict measured data. The original FE based model had 17,196 degrees of freedom. Using a modal approach an eightieth-order ASE model was developed and used for comparison (refs. 38–39). The ASE model contained 60 structural and 20 unsteady aerodynamic lag states.

Figures 5a–5h illustrate results of how accurately, as %QF, the two models correspond to wind-tunnel data at Mach 0.80, 0.95, 1.1 and,  $Q = 55, 60, 65$  psf.

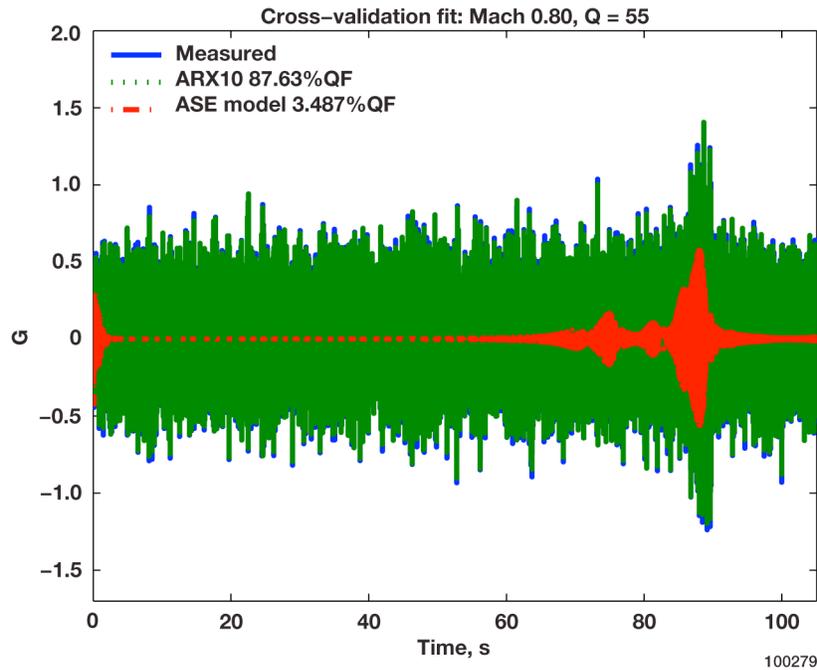


Figure 5(a). Cross-validation fit: Mach 0.80,  $Q = 55$  psf.

Figure 5. Predicted outputs of the ASE and ARX10 model superimposed on top of measured output; (a–c): Mach 0.80,  $Q = 55$ –65 psf; (d–e): Mach 0.95,  $Q = 55$ –60 psf; and (f–h): Mach 1.10,  $Q = 55$ –65 psf. Solid line (“—”) measured output. Dot-dot line (“...”) predicted ARX10 model output. Dash-dot line (“-..”) predicted ASE model output.

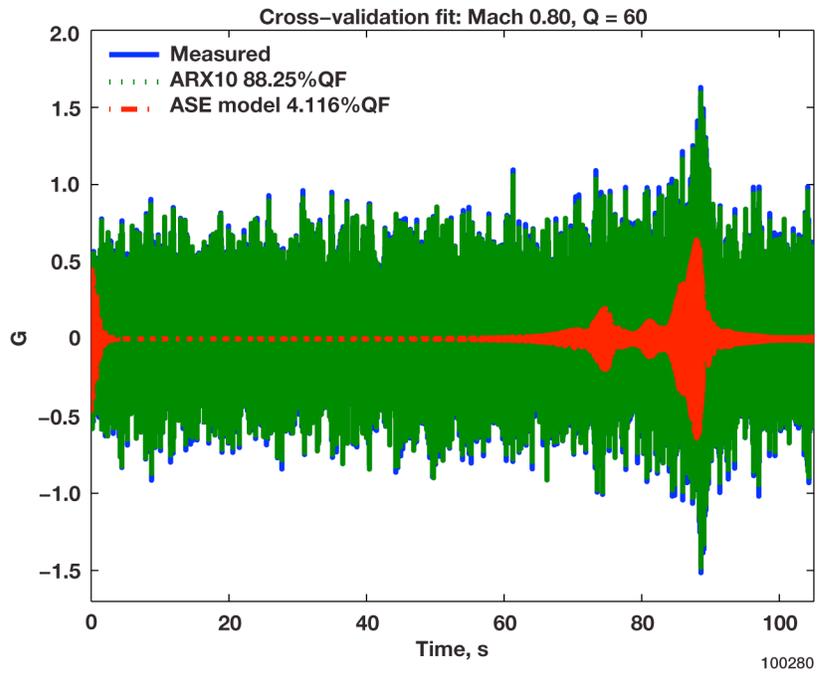


Figure 5(b). Cross-validation fit: Mach 0.80, Q = 60 psf.

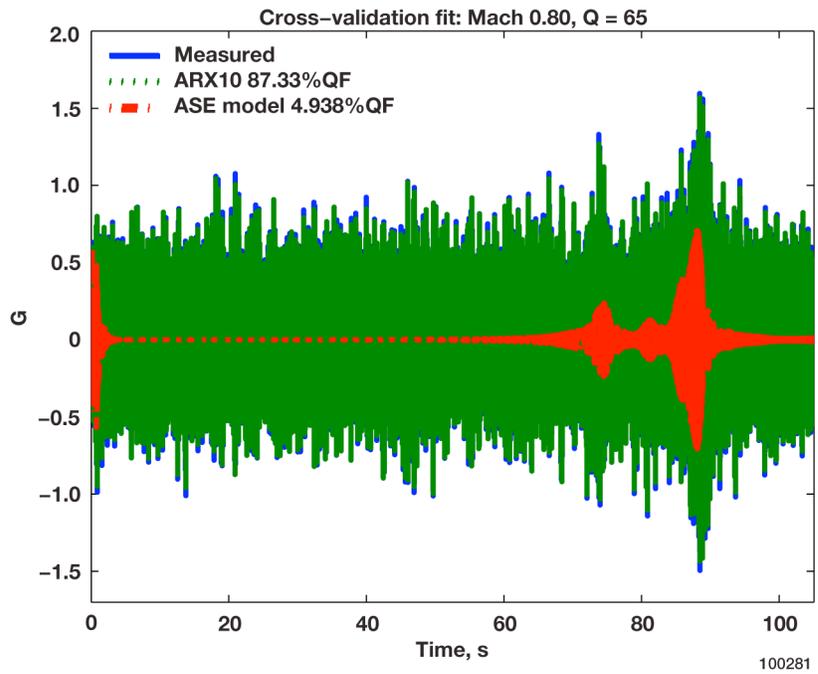


Figure 5(c). Cross-validation fit: Mach 0.80, Q = 65 psf.

Figure 5. Continued.

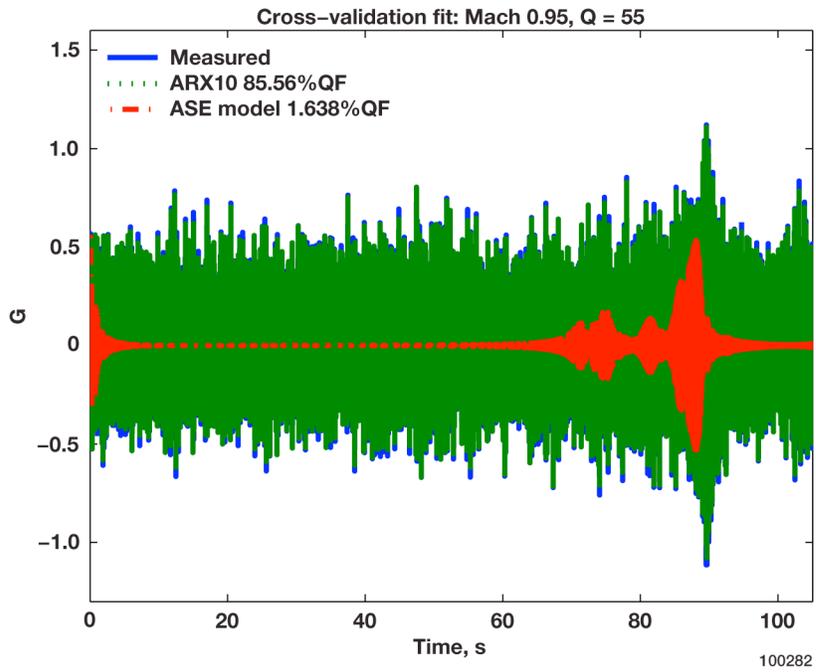


Figure 5(d). Cross-validation fit: Mach 0.95, Q = 55 psf.

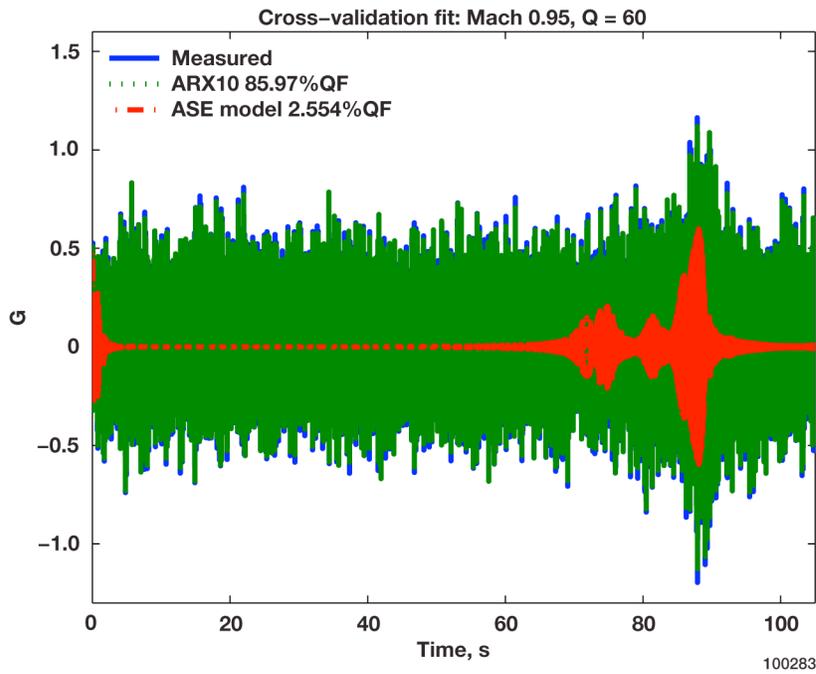


Figure 5(e). Cross-validation fit: Mach 0.95, Q = 60 psf.

Figure 5. Continued.

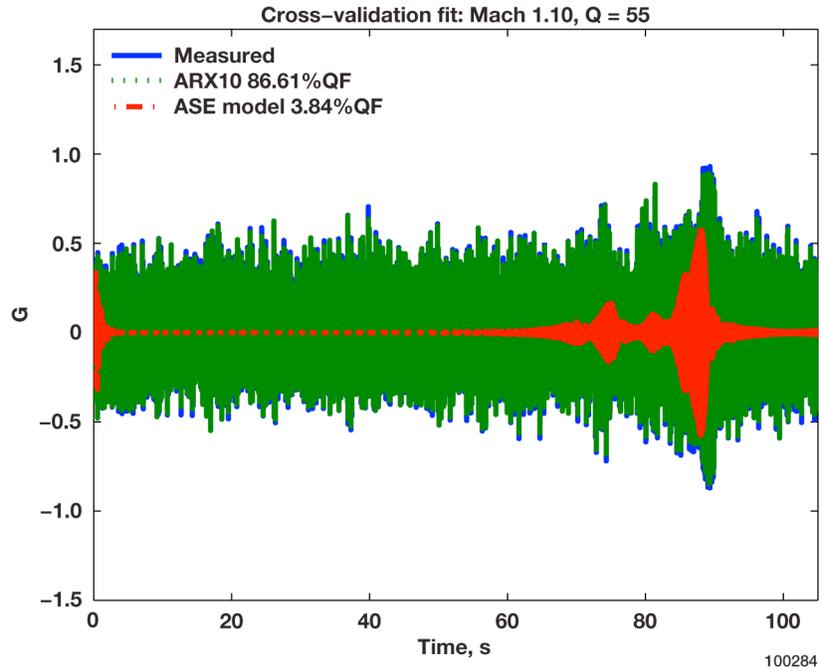


Figure 5(f). Cross-validation fit: Mach 1.10, Q = 55 psf.

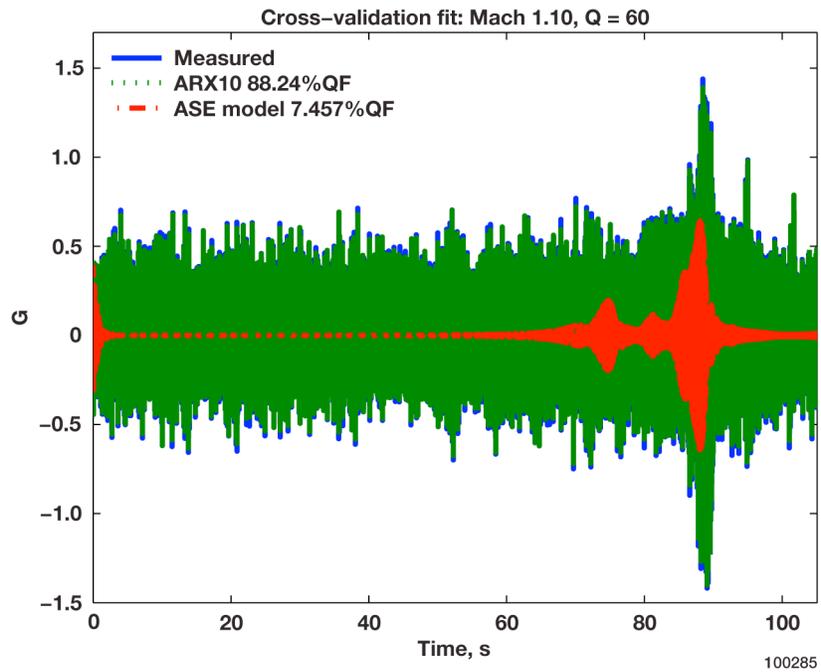


Figure 5(g). Cross-validation fit: Mach 1.10, Q = 60 psf.

Figure 5. Continued.

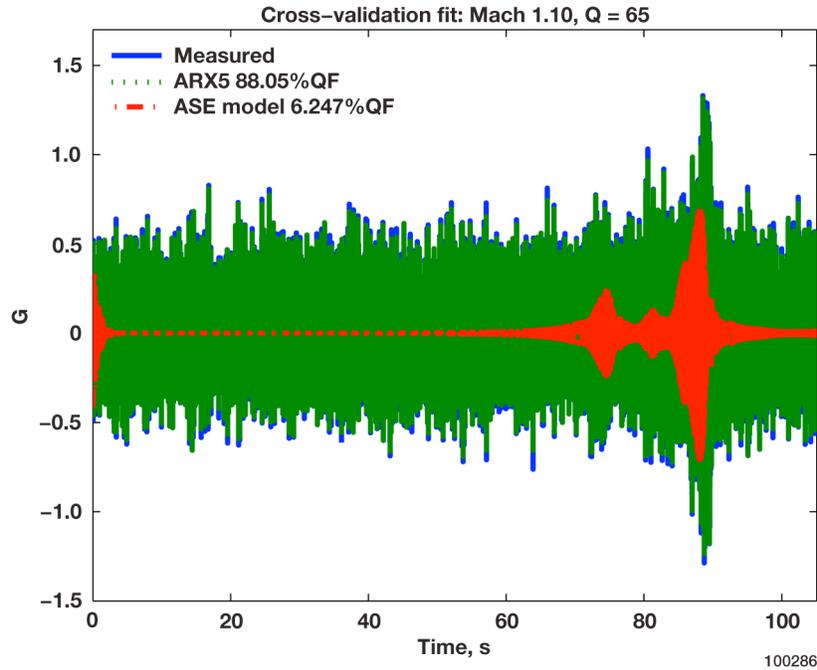


Figure 5(h). Cross-validation fit: Mach 1.10, Q = 65 psf.

Figure 5. Concluded.

Figures 5a–5h shows the predicted output for the ARX10 and ASE models superimposed on top of measured data. The %QF’s obtained for the ARX10 model, at all Mach numbers and  $Q$ ’s, range from 85% to 88%. For the ASE model the %QF’s, for all Mach numbers and  $Q$ ’s, range from 2% to 7%. Although the FE ASE has many more degrees of freedom, it was not able to provide better predictive capability than the data-driven ARX model. Clearly, the tenth-order ARX model obtained using system identification methodology outperforms the FE ASE model.

## 5. Discussion

This study explored the utility of system identification techniques to develop robust models for control synthesis. Initially, at each Mach number, an ARX model with maximum order of five was posed to the AIC, FPE and MDL techniques to estimate optimal lag and dynamic order. Next the ARX models were allowed greater flexibility with a maximum order of ten to assess whether the same structure could explain more of the output variance. Analysis of these results indicated that the tenth-order ARX models provided sufficient improvement in model fit to justify the added complexity. Therefore, at each Mach number the tenth-order ARX models were used as a starting point for ARMAX model development. The best fit ARMAX models were then used to develop BJ models.

The results show whilst the ARMAX models provide an improved fit, it was less than 1% for each case. The predictive capability of the BJ models was less than that of the ARMAX structures, which is a symptom of too much complexity. As such, these results indicate that the tenth-order ARX models were optimal in terms of parsimony and predictive capability to describe the dynamics of recorded wind-tunnel data.

A comparison of the tenth-order ARX and 80th-order ASE model’s ability to explain the output variance revealed that the more complex ASE model was not able to achieve as high of a fit. The ARX

model was able to attain a higher fit because the model was developed using measured data, which often contains dynamics that are not captured in a FE based model. Conversely, the ASE model was developed from idealised assumptions about mass, damping, et cetera, which are idealisation and often do not closely hold with observations under experimental conditions.

Although a different ASE model was developed for each Mach and  $Q$ , they yielded significantly reduced fits to measured data compared with the tenth-order ARX model. Notice that the ARX model at each Mach number was developed with estimation data at  $Q = 30$  psf but was able to accurately predict the measured data at higher  $Q$ 's, namely,  $Q = 55-65$  psf. This is the power of utilising a data driven approach to model development.

Often a model, which yields a minimum system description yet provides good predictive capabilities, is deemed as the best or optimal model. This is a trade-off between the ability to describe the system behaviour and parsimony. If two models can describe the system behaviour almost equally well, why choose the more complex one? For example, in  $H^\infty$ -control synthesis the order of the controller is equal to the order of the model plus the order of the performance weights. Of course, it is not as computationally demanding to compute a controller of lower order. Moreover, selecting a model with greater complexity may lead to numerical issues if the model order is too high. Hence, minimum complexity often renders control synthesis more tractable (33, 47).

A future study will utilise these data-driven models to develop control laws to assess model and controller design robustness. Initially, this study will be performed in a simulation environment but with data collected from the wind-tunnel tests described in this report. The successful demonstration of this work may lead to comparison of the robustness of control law design based on the two approaches, FE ASE and system identification models, on a supersonic flight test vehicle.

## 6. Conclusion

This study demonstrates the application of system identification techniques to develop parsimonious and robust models directly from data with excellent predictive capability. The results show that a data-driven model is capable of predicting observed data for a larger operating point. This robust predictive power was demonstrated with models developed at  $Q = 30$  psf, but able to accurately predict the measured output at higher  $Q$ 's. This superior predictive capability allows for simpler control solutions and reduced modelling effort, whilst traditional ASE based control strategies rely on computationally expensive models with little predictive power rendering control law design more expensive.

## References

1. Danowsky, Brian P., Peter M. Thompson, Charbel Farhat, and Chuck Harris, "Residualization of an Aircraft Linear Aeroelastic Reduced Order Model to Obtain Static Stability Derivatives," *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, AIAA-2008-6370, Honolulu, Hawaii, August 18-21, 2008.
2. Courant, R., "Variational Methods for the Solution of Problems of Equilibrium and Vibrations," *Bulletin of the American Mathematical Society*, Vol. 49, No. 1, pp. 1-23, January 1943.
3. Hrennikoff, A., "Solution of Problems of Elasticity by the Framework Method," *Journal of Applied Mechanics*, Vol. 8, pp. A169-A175, 1941.
4. King, Belinda B., Naira Hovakimyan, Katie A. Evans, and Michael Buhl, "Reduced Order Controllers for Distributed Parameter Systems: LQG Balanced Truncation and an Adaptive Approach," *Mathematical and Computer Modelling*, Vol. 43, No. 9-10, pp. 1136-1143, May 2006.
5. Karhunen, K., "Zur Spektraltheorie Stochastischer Prozesse," *Annales Academiae Scientiarum Fennicae*, Vol. 37, 1946.
6. Ahmed, N., and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*, Springer, New York, 1975.
7. Almroth, B. O., P. Stern, and F. A. Brogan, "Automatic Choice of Global Shape Functions in Structural Analysis," *AIAA Journal*, Vol. 16, No. 5, May 1978.
8. Nagy, Dennis A., "Modal Representation of Geometrically Nonlinear Behavior by the Finite Element Method," *Computers & Structures*, Vol. 10, No. 4, pp. 683-688, August 1979.
9. Lucia, David J., Paul I. King, and Philip S. Beran, "Reduced Order Modeling of a Two-Dimensional Flow With Moving Shocks," *Computers & Fluids*, Vol. 32, No. 7, pp. 917-938, August 2003.
10. Noor, Ahmed K., and Jeanne M. Peters, "Recent Advances in Reduction Methods for Instability Analysis of Structures," *Computers & Structures*, Vol. 16, No. 1-4, pp. 67-80, 1983.
11. Noor, Ahmed K., C. M. Andersen, and Jeanne M. Peters, "Reduced Basis Technique for Collapse Analysis of Shells," *AIAA Journal*, Vol. 19, No. 3, pp. 393-397, March 1981.
12. May, Stephen F., and Ralph C. Smith, "Reduced-Order Model Design for Nonlinear Smart System Models," *Modeling, Signal Processing, and Control for Smart Structures*, Vol. 7286, 2009.
13. Banks, H. T., S. C. Beeler, G. M. Kepler, and H. T. Tran, "Reduced Order Modeling and Control of Thin Film Growth in an HPCVD Reactor," *SIAM Journal on Applied Mathematics*, Vol. 62, No. 4, pp. 1251-1280, 2002.
14. Fukunaga, Keinosuke, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, 1990.

15. Holmes, Phillip, John L. Lumley, and Gal Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge Monographs on Mechanics, Cambridge University Press, 1996.
16. Ito, Kazufumi, and Jeffrey D. Schroeter, “Reduced Order Feedback Synthesis for Viscous Incompressible Flows,” *Mathematical and Computer Modelling*, Vol. 33, No. 1–3, pp. 173–192, January–February 2001.
17. Sirovich, Lawrence, “Turbulence and the Dynamics of Coherent Structures,” *Quarterly of Applied Mathematics*, Vol. 45, pp. 561–571, 573–590, October 1987.
18. Ly, Hung V., and Hien T. Tran, “Modeling and Control of Physical Processes Using Proper Orthogonal Decomposition,” *Mathematical and Computer Modelling*, Vol. 33, No. 1–3, pp. 223–236, January–February 2001.
19. Shvartsman, Stanislav Y., and Ioannis G. Kevrekidis, “Nonlinear Model Reduction for Control of Distributed Systems: A Computer-Assisted Study,” *AIChE Journal*, Vol. 44, No. 7, pp. 1579–1595, July 1998.
20. Graham, W. R., J. Peraire, and K. Y. Tang, “Optimal Control of Vortex Shedding Using Low-Order Models,” *International Journal for Numerical Methods in Engineering*, Vol. 44, No. 7, pp. 973–990, March 1999.
21. Du, Haiping, Nong Zang, and Hung Nguyen, “Mixed  $H_2/H_\infty$  Control of Tall Buildings with Reduced-Order Modelling Technique,” *Structural Control and Health Monitoring*, Vol. 15, No. 1, pp. 64–89, February 1989.
22. Banks, H. T., Michele L. Joyner, Buzz Wincheski, and William P. Winfree, “Nondestructive Evaluation Using a Reduced-Order Computational Methodology,” *Inverse Problems*, Vol. 16, No. 4, pp. 929–945, August 2000.
23. Lucia, David J., Philip S. Beran, and Walter A. Silva, “Reduced-Order Modeling: New Approaches for Computational Physics,” *Progress in Aerospace Sciences*, Vol. 40, No. 1–2, pp. 51–117. February 2004.
24. Romanowski, Michael C., “Reduced Order Unsteady Aerodynamic and Aeroelastic Models Using Karhunen–Loeve Eigenmodes,” AIAA-1996-3981-CP, *Proceedings of the 6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, pp. 7–13, Bellevue, Washington, September 4–6, 1996.
25. Kim, Taehyoun, “Frequency-Domain Karhunen–Loeve Method and Its Application to Linear Dynamic Systems,” *AIAA Journal*, Vol. 36, No. 11, pp. 2117–2123, November 1998.
26. Hall, Kenneth C., Jeffrey P. Thomas, and Earl H. Dowell, “Proper Orthogonal Decomposition Technique for Transonic Unsteady Aerodynamic Flows,” *AIAA Journal*, Vol. 38, No. 10, pp. 1853–1862, October 2000.
27. Pettit, Chris L., and Philip S. Beran, “Reduced-Order Modeling for Flutter Prediction,” AIAA-2000-1446, *Proceedings of the 41<sup>st</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit*, Atlanta, Georgia, April 3–6, 2000.

28. Pettit, C. L., and P. S. Beran, "Application of Proper Orthogonal Decomposition to the Discrete Euler Equations," *International Journal for Numerical Methods in Engineering*, Vol. 55, No. 4, pp. 479–497, October 2002.
29. Dowell, E. H., J. P. Thomas, and K. C. Hall, "Transonic Limit Cycle Oscillation Analysis Using Reduced Order Aerodynamic Models," *Journal of Fluids and Structures*, Vol. 19, No. 1, pp. 17–27, January 2004.
30. Cardoso, M. A., L. J. Durlosfsky, and P. Sarma, "Development and Application of Reduced–Order Modeling Procedures for Subsurface Simulation," *International Journal for Numerical Methods in Engineering*, Vol. 77, No. 9, pp. 1322–1350, February 2009.
31. Cardoso, M. A., and L. J. Durlosfsky, "Linearized Reduced–Order Models for Subsurface Flow Simulation," *Journal of Computational Physics*, Vol. 229, No. 3, pp. 681–700, February, 2010.
32. Rewienski, Michal, and Jacob White, "A Trajectory Piecewise–Linear Approach to Model Order Reduction and Fast Simulation of Non linear Circuits and Micromachined Devices," *IEEE Transactions on Computer–Aided Design of Integrated Circuits and Systems*, Vol. 22, No. 2, pp. 155–170, February 2003.
33. Ljung, Lennart, *System Identification: Theory for the User*, 2<sup>nd</sup> ed., Prentice Hall, Upper Saddle River, New Jersey, 1999.
34. Bendat, Julius S., and Allan G. Piersol, *Random Data: Analysis and Measurement Procedures*, 2<sup>nd</sup> ed., John Wiley and Sons, New York, 1986.
35. Siebert, William McC., *Circuits, Signals, and Systems*, The MIT Press, Cambridge, Massachusetts, 1986.
36. Chen, Rong, and Ruey S. Tsay, "Nonlinear Additive ARX Models," *Journal of the American Statistical Association*, Vol. 88, No. 423, pp. 955–967, September 1993.
37. Perry, Boyd III, Walter A. Silva, James R. Florance, Carol D. Wieseman, Anthony S. Pototzky, Mark D. Sanetrik, and et al., "Plans and Status of Wind–Tunnel Testing Employing an Aeroservoelastic Semispan Model," AIAA–2007–1770, *Proceedings of the 48<sup>th</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Honolulu, Hawaii, April 23–26, 2007.
38. Chen, P. C., Boris Moulin, Erich Ritz, D. H. Lee, and Z. Zhang, "CFD–Based Aeroservoelastic Control for Supersonic Flutter Suppression, Gust Load Alleviation, and Ride Quality Enhancement," AIAA–2009–2537, *Proceedings of the 50<sup>th</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Palm Springs, California, May 4–7, 2009.
39. Moulin, Boris, "CFD–Based Control for Flutter Suppression, Gust Load Alleviation, and Ride Quality Enhancement for the S<sup>4</sup>T Model," AIAA–2010–2623, *Proceedings of the 51<sup>st</sup> AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Orlando, Florida, April 12–15, 2010.

40. Goodwin, Graham C., and Robert L. Payne, *Dynamic System Identifications: Experiment Design and Data Analysis*, *Mathematics in Science and Engineering* book series, Vol. 136, Academic Press, New York, 1977.
41. Hsia, Tien C., *System Identification: Least-Squares Methods*, Lexington Books, Lexington, Massachusetts, 1977.
42. Box, George E. P., and Gwilym M. Jenkins, *Time Series Analysis; Forecasting and Control*, Holden-Day, San Francisco, 1970.
43. Akaike, Hirotugu, "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics*, Vol. 21, No. 1, pp. 243–247, December 1969.
44. Akaike, H., "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, pp. 716–723, December 1974.
45. Rissanen, J., "Modeling by Shortest Data Description," *Automatica*, Vol. 14, pp. 465–471, 1978.
46. Boulet, B., and B. A. Francis, "Consistency of Open-Loop Experimental Frequency-Response Data with Coprime Factor Plant Models," *IEEE Transactions on Automatic Control*, Vol. 43, No. 12, pp. 1680–1691, December 1998.
47. Zhou, Kemin, John C. Doyle, and Keith Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

