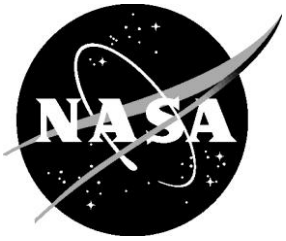
NASA/TM–20210009593



Nonlinear Unsteady Aerodynamic Modeling using Empirical Orthogonal Functions

Eugene A. Morelli
Langley Research Center, Hampton, Virginia

February 2021

NASA STI Program Report Series

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 NTRS Registered and its public interface, the NASA Technical Reports 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 information 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

- Help desk contact information:

https://www.sti.nasa.gov/sti-contact-form/
and select the "General" help request type.

NASA/TM–20210009593

Nonlinear Unsteady Aerodynamic Modeling using Empirical Orthogonal Functions

Eugene A. Morelli
Langley Research Center, Hampton, Virginia

National Aeronautics and Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

February 2021

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Table of Contents

Abstract ……..……..……. ii

Nomenclature ……..……..... iii

Introduction ……..……..... 1

Method ……..……..... 2

Results ……..……..... 10

Conclusions ……..……..... 17

Acknowledgments ……..... 17

References ……..……..... 18

Abstract

Empirical orthogonal function modeling is explained and applied to identify compact discrete-time nonlinear unsteady aerodynamic models from data generated by an unsteady three-dimensional compressible Navier-Stokes flow solver for an airfoil undergoing various pitching motions. Model structures, model parameter estimates, and model parameter uncertainty estimates for nondimensional lift, drag, and pitching moment coefficient models were determined directly from the data. Prediction tests using data that were not used in the modeling process showed that the identified models exhibited excellent prediction capability, which is a strong indicator of an accurate model.

Nomenclature

α	angle of attack, deg or rad
Δt	sampling interval, s
θ	model parameter
$\boldsymbol{\Sigma}$	covariance matrix
C_D	nondimensional aerodynamic drag coefficient
C_L	nondimensional aerodynamic lift coefficient
C_m	nondimensional aerodynamic pitching moment coefficient
i	discrete time index
n	number of model terms
N	number of data points
u	model input
v	model residual $= z - y$
y	model output
z	measured output

SUPERSCRIPTS

$\dot{}$	time derivative
$\tilde{}$	Fourier transform
^{T}	transpose

I. Introduction

Complex nonlinear dynamic systems can be modeled using Nonlinear Auto-Regressive Moving Average with eXogenous input (NARMAX) models [1,2]. For a single-input, single-output system, a NARMAX model is a generally nonlinear discrete-time model relating the current output to current and past values of the input, as well as past values of the model output. For nonlinear unsteady aerodynamics, this form can be simplified to modeling the current output using generally nonlinear terms involving only the current and past values of the input.

An important step in discrete-time nonlinear unsteady aerodynamic modeling is selecting an appropriate form for the model structure, both in terms of the number of past values to include, as well as which functions of the current and past values to select for the model. This is called model structure determination. The model structure including only current and past values of the input is conducive to efficient model structure determination, mainly because the associated model parameter estimation problem is linear [3-5]. Furthermore, this general model structure is the desired form for uses such as aerodynamic analysis, nonlinear simulation, and control system design, among others.

The idea of using empirical orthogonal polynomial functions to model generally nonlinear dependencies was developed independently in the UK [1,2] in the context of industrial modeling and in the US [3-6] in the context of aerodynamic modeling. The statistical metrics used for the model term selection were different in these separate developments, but the general idea of orthogonalizing candidate functions to clarify their importance to modeling the variation in the measured output, and thereby determining an acceptable subset of the candidate functions for the model, was similar.

Past work in nonlinear aerodynamic modeling using empirical multivariate orthogonal functions focused on quasi-steady nonlinear modeling with multivariate dependencies, where nondimensional aerodynamic coefficients were modeled as general nonlinear functions of the current values of explanatory variables, using empirical orthogonal functions generated from the data. Current measured aircraft states and controls were considered as candidate explanatory variables [3-9].

The present work extends this quasi-steady nonlinear modeling approach by applying empirical orthogonal function modeling to identify compact discrete-time nonlinear unsteady models using both current and past values of a single input or physical explanatory variable. The method is applied to discrete-time data generated by an unsteady three-dimensional compressible Navier-Stokes flow solver for an airfoil undergoing various pitching motions [10] to identify models for the longitudinal nondimensional aerodynamic coefficients, namely lift coefficient C_L, drag coefficient C_D, and pitching moment coefficient C_m, as a function of current and past values of angle of attack. The aerodynamics are generally nonlinear and unsteady, and the model forms are identified from the data using empirical orthogonal functions, without any prior information or analyst judgment. This application is intended to investigate the use of the empirical orthogonal function approach to identify discrete-time nonlinear unsteady aerodynamic models from time series data. The case of multivariate nonlinear unsteady aerodynamic dependencies for multiple outputs will be shown as an extension of the simpler case examined in this work. Such extensions do not require any further analytic development, but would require additional computational effort and good experiment design.

The next section contains a brief description of the empirical orthogonal function modeling approach. In the Results section, the empirical orthogonal function modeling technique is applied to computational fluid dynamics data for an airfoil to identify compact discrete-time models for nondimensional longitudinal aerodynamic coefficients, including prediction tests using data that were not used for the model identification. The final section provides a summary and conclusions.

All of the data processing and modeling tasks for this work were done using a software toolbox written in MATLAB® called System IDentification Programs for AirCraft (SIDPAC) [5,6]. SIDPAC was developed at NASA Langley and has been applied successfully to a wide variety of flight test and wind tunnel experiments. SIDPAC has been used at more than 100 organizations worldwide to solve aircraft system identification problems [11].

II. Method

The general form of a discrete-time NARMAX model for a single input u and a single output y with discrete-time index i is

$$y(i) = F\left[y(i-1), y(i-2), ..., y\left(i-n_y\right), u(i), u(i-1), ..., u\left(i-n_u\right) \right] \tag{1a}$$

where $F[\bullet]$ represents a general nonlinear function, and the maximum number of past values of the input and output needed are denoted by n_y and n_u respectively. An example of a NARMAX model would be

$$y(i) = \theta_1 y(i-1) + \theta_2 y(i-1)\, y(i-2) + \theta_3 y(i-3) + \theta_4 u(i)\, y(i-1) + \theta_5 u^2(i-1) \tag{1b}$$

where $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$ are unknown model parameters to be determined from the data.

For the nonlinear unsteady aerodynamic modeling problem, the general NARMAX model form can be simplified to use only current and past values of the input, each of which can also be considered as an explanatory variable,

$$y(i) = F\left[u(i), u(i-1), ..., u\left(i-n_u\right) \right] \tag{2a}$$

In this work, the explanatory variables are current and past values of the angle of attack, and the responses are nondimensional longitudinal aerodynamic force and moment coefficients. An example of a nonlinear unsteady aerodynamic model would be

$$C_L(i) = \theta_1 \alpha(i) + \theta_2 \alpha(i-3) + \theta_3 \alpha(i)\, \alpha(i-3) + \theta_4 \alpha^2(i)\, \alpha(i-1) \tag{2b}$$

Nonlinear unsteady aerodynamic effects depend in general on the current value of the angle of attack (the quasi-steady assumption), as well as the history of the angle of attack. The history of angle of attack at some time τ prior to the current time t can be approximated using current first-order and higher-order derivatives in a truncated Taylor series,

$$\alpha(\tau) \approx \alpha(t) + \dot{\alpha}(t)(\tau - t) + \frac{1}{2}\ddot{\alpha}(t)(\tau - t)^2 + \ldots \qquad \tau \le t \tag{3}$$

This is the basis for using derivative terms to model unsteady effects. Other approaches include using analogies to expressions from simple aerodynamic theory to model the aerodynamic

dependence on explanatory variable history, and applying frequency-dependent stability and control derivatives, among others [5,12].

However, identifying a generally nonlinear model using past values of the explanatory variable(s) is a more direct interpretation of the physics. The problem lies in identifying how far into the past to extend the modeling (i.e., which past values should be used), and what forms the model terms should take to characterize the dependence on the current and past values. This is a discrete-time nonlinear model identification problem using current and past values of the explanatory variable(s). An analog of this problem from a different perspective is identifying an indicial function for nonlinear unsteady aerodynamics [5,12].

The form and number of the generally nonlinear model terms for a discrete-time nonlinear unsteady model, along with estimates of the model parameters and uncertainties can be determined using empirical orthogonal function modeling, as described next. All of the associated calculations are implemented in the SIDPAC program called mof.m.

A. Orthogonal Function Modeling

The form of an empirical orthogonal function model is

$$\boldsymbol{y} = \theta_1 \, \boldsymbol{p}_1 + \theta_2 \, \boldsymbol{p}_2 + ... + \theta_n \, \boldsymbol{p}_n \tag{4}$$

where $\boldsymbol{y}$ is an N-dimensional vector of the model output, $\boldsymbol{y} = \left[y_1, y_2, ..., y_N \right]^T$, composed of a linear combination of n mutually orthogonal modeling functions $\boldsymbol{p}_j, \; j = 1, 2, ..., n$. Each $\boldsymbol{p}_j$ is an N-dimensional vector which in general depends on the explanatory variables. The parameters $\theta_j, \; j = 1, 2, ..., n$ are unknown constant model parameters to be determined from the data.

Noisy output measurements are related to the model output by

$$\boldsymbol{z} = \boldsymbol{y} + \boldsymbol{\varepsilon} = \theta_1 \, \boldsymbol{p}_1 + \theta_2 \, \boldsymbol{p}_2 + ... + \theta_n \, \boldsymbol{p}_n + \boldsymbol{\varepsilon} \tag{5}$$

where $\boldsymbol{z}$ is an N-dimensional vector of output or response measurements (e.g., nondimensional force or moment coefficient values), $\boldsymbol{z} = \left[z_1, z_2, ..., z_N \right]^T$, and $\boldsymbol{\varepsilon}$ denotes the modeling error vector.

Equation (5) is a mathematical model for the functional dependencies in the measured data. The important questions of how to compute the mutually-orthogonal functions $\boldsymbol{p}_j$ from the explanatory variable data, as well how to select which orthogonal functions to include in Eq. (5), which implicitly determines n, will be addressed later. At this point, the properties of an empirical orthogonal function model will be examined.

Define an $N \times n$ matrix $\boldsymbol{P}$,

$$\boldsymbol{P} = \left[\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_n \right] \tag{6}$$

and let $\boldsymbol{\theta} = \left[\theta_1, \theta_2, ..., \theta_n \right]^T$. Equation (5) can then be written as a standard least squares regression problem,

$$\boldsymbol{z} = \boldsymbol{P} \boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{7}$$

The error vector $\boldsymbol{\varepsilon}$ is to be minimized in a least squares sense. The goal is to determine $\boldsymbol{\theta}$ that minimizes the least squares cost function

$$J = \frac{1}{2}(\boldsymbol{z} - \boldsymbol{P\theta})^T (\boldsymbol{z} - \boldsymbol{P\theta}) = \frac{1}{2}\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \tag{8}$$

The parameter vector estimate $\hat{\boldsymbol{\theta}}$ that minimizes this cost function is [5]

$$\hat{\boldsymbol{\theta}} = \left[\boldsymbol{P}^T \boldsymbol{P}\right]^{-1} \boldsymbol{P}^T \boldsymbol{z} \tag{9}$$

The estimated parameter covariance matrix is [5]

$$\boldsymbol{\Sigma}_{\hat{\theta}} = E\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^T\right] = \sigma^2 \left(\boldsymbol{P}^T \boldsymbol{P}\right)^{-1} \tag{10}$$

where E is the expectation operator, and the fit error variance σ^2 can be estimated from the residuals

$$\boldsymbol{v} = \boldsymbol{z} - \boldsymbol{P}\hat{\boldsymbol{\theta}} \tag{11}$$

using

$$\hat{\sigma}^2 = \frac{1}{(N-n)}\left[\left(\boldsymbol{z} - \boldsymbol{P}\hat{\boldsymbol{\theta}}\right)^T \left(\boldsymbol{z} - \boldsymbol{P}\hat{\boldsymbol{\theta}}\right)\right] = \frac{\boldsymbol{v}^T \boldsymbol{v}}{(N-n)} \tag{12}$$

Parameter standard errors are computed as the square root of the diagonal elements of the $\boldsymbol{\Sigma}_{\hat{\theta}}$ matrix in Eq. (10), using $\hat{\sigma}^2$ from Eq. (12). The identified model output $\hat{\boldsymbol{y}}$ is computed as

$$\hat{\boldsymbol{y}} = \boldsymbol{P}\hat{\boldsymbol{\theta}} \tag{13}$$

In conventional least-squares modeling, the modeling functions are often polynomials in the explanatory variables, which in general are not mutually orthogonal and therefore are partially correlated. In the context of discrete-time nonlinear unsteady aerodynamic modeling for an airfoil, this corresponds to polynomials in the current and past values of the angle of attack. However, in this modeling approach, the model terms are not restricted to polynomial functions and in fact can be any arbitrary generally nonlinear functions of the explanatory variables.

If the modeling functions are instead empirical orthogonal functions generated from the explanatory variable data, it is easier to determine an appropriate model structure, because the explanatory capability of each modeling function is completely distinct from all of the others. This decouples the least squares modeling problem, as will be shown next.

For mutually orthogonal modeling functions,

$$\boldsymbol{p}_i^T \boldsymbol{p}_j = 0 \qquad i \neq j \quad , \quad i, j = 1, 2, ..., n \tag{14}$$

and $\boldsymbol{P}^T \boldsymbol{P}$ is a diagonal matrix with the inner product of the orthogonal functions on the main diagonal. Using Eqs. (6) and (14) in Eq. (9), the jth element of the estimated parameter vector $\hat{\boldsymbol{\theta}}$ is computed as

$$\hat{\theta}_j = \left(\boldsymbol{p}_j^T \boldsymbol{z}\right) \Big/ \left(\boldsymbol{p}_j^T \boldsymbol{p}_j\right) \tag{15}$$

Using Eqs. (6), (14), and (15) in Eq. (8),

$$\hat{J} = \frac{1}{2}\left[\boldsymbol{z}^T \boldsymbol{z} - \sum_{j=1}^{n} \left(\boldsymbol{p}_j^T \boldsymbol{z} \right)^2 \Big/ \left(\boldsymbol{p}_j^T \boldsymbol{p}_j \right) \right] \tag{16}$$

Equation (16) shows that when the modeling functions are orthogonal, the reduction in the least-squares cost function resulting from including the term $\theta_j \boldsymbol{p}_j$ in the model depends only on the response variable data $\boldsymbol{z}$ and the added orthogonal modeling function $\boldsymbol{p}_j$. The least-squares modeling problem is therefore decoupled, which means that each orthogonal modeling function can be evaluated independently in terms of its ability to reduce the least-squares model fit to the data, regardless of which other orthogonal modeling functions are already selected for the model. When the modeling functions are instead polynomials in the explanatory variables (or any other non-orthogonal function set), the least-squares problem is coupled, and iterative analysis is required to find the modeling functions for an adequate model structure.

The orthogonal modeling functions to be included in the model are chosen to minimize predicted squared error, PSE, defined by [3-9,13,14]

$$\text{PSE} \equiv \frac{\left(\boldsymbol{z} - \boldsymbol{P}\hat{\boldsymbol{\theta}} \right)^T \left(\boldsymbol{z} - \boldsymbol{P}\hat{\boldsymbol{\theta}} \right)}{N} + \sigma_{max}^2 \frac{n}{N} \tag{17}$$

or

$$\text{PSE} = \frac{2\hat{J}}{N} + \sigma_{max}^2 \frac{n}{N} \tag{18}$$

The constant σ_{max}^2 is the upper-bound estimate of the mean squared error between future data and the model, i.e., the upper-bound mean squared error for prediction cases. The upper bound is used in the model overfit penalty term to account for the fact that PSE is calculated when the model structure is not correct, i.e., during the model structure determination stage. Using the upper bound is conservative in the sense that model complexity will be reduced as a result of using an upper bound for this constant in the penalty term. Because of this, the value of PSE computed from Eq. (17) for a particular model structure tends to overestimate actual prediction errors on new data. Therefore, the PSE metric conservatively estimates the squared error for prediction cases.

A simple estimate of σ_{max}^2 that is independent of the model structure can be obtained by computing σ_{max}^2 as the residual variance estimate for a constant model equal to the mean of the measured response values,

$$\sigma_{max}^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left[z_i - \bar{z} \right]^2 \tag{19}$$

where

$$\bar{z} = \frac{1}{N} \sum_{i=1}^{N} z_i \tag{20}$$

The PSE in Eq. (18) depends on the mean squared fit error, $2\hat{J}/N$, and a term proportional to the number of terms in the model, n. The latter term prevents overfitting the data with too many model terms, which is detrimental to model prediction accuracy [5,13,14]. The mean squared fit error $2\hat{J}/N$ must decrease with the addition of each orthogonal modeling function to the model (by Eq. (16)), whereas the overfit penalty term $\sigma_{max}^2 n/N$ must increase with each added model term (n increases). Introducing the orthogonal modeling functions into the model in order of most effective to least effective in reducing the mean squared fit error (quantified by $\left(\boldsymbol{p}_j^T \boldsymbol{z}\right)^2 \Big/ \left(\boldsymbol{p}_j^T \boldsymbol{p}_j\right)$ for the jth orthogonal modeling function) results in the PSE metric always having a single global minimum.

Based on Eqs. (16) and (18), an individual orthogonal function term $\theta_j \boldsymbol{p}_j$ will only be included in the model when

$$\left(\boldsymbol{p}_j^T \boldsymbol{z}\right)^2 \Big/ \left(\boldsymbol{p}_j^T \boldsymbol{p}_j\right) < \sigma_{max}^2 \tag{21}$$

Model terms that do not satisfy the inequality (21) will increase the PSE, which means that only model terms that reduce the mean squared fit error more than the estimated maximum noise variance for future data will be selected for the model. In other words, only orthogonal functions that can characterize variation in the response variable greater than the expected maximum noise level are selected for inclusion in the model.

Figure 1 depicts this graphically, using actual modeling results from Ref. [14]. The figure shows that after the first 6 modeling functions, the added model complexity associated with an additional orthogonal modeling function is not justified by the associated reduction in mean squared fit error. This point is marked by minimum PSE, which defines an adequate model structure with good predictive capability. Further statistical arguments for the form of PSE given in Eq. (17), including justification for its use in modeling problems, can be found in Ref. [13].

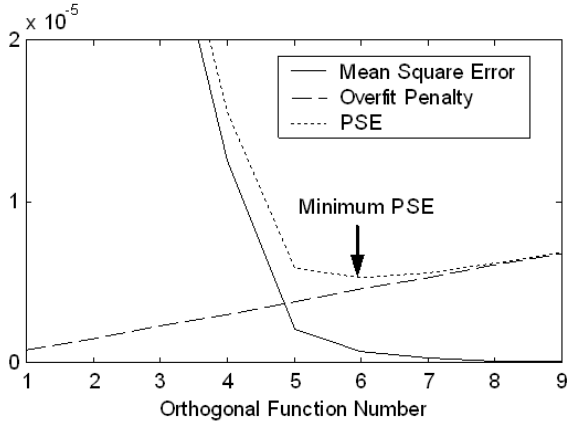


Figure 1. Model structure determination using orthogonal functions and PSE

Modeling accuracy can be enhanced with a more sophisticated estimate of the maximum noise variance for future data, compared to the simple estimate obtained from Eqs. (19)-(20). An accurate noise variance estimate for the measured response data that is independent of any modeling can be found using the global Fourier smoothing technique implemented in SIDPAC program smoo.m [5,6,15]. This noise variance estimate is very accurate, so that σ_{max}^2 can be specified as

$$\sigma_{max}^2 = 25\hat{\sigma}^2 \tag{22}$$

where $\hat{\sigma}^2$ is the noise variance estimate from global Fourier smoothing applied to the measured response data. Assuming the noise is Gaussian, Eq. (22) would correspond to conservatively placing the maximum noise variance at 25 times the estimated noise variance, corresponding to a $5\hat{\sigma}$ maximum deviation. Simulation investigations have shown that this value produces accurate model structure determination results. Equation (22) was used with noise variance estimates from the global Fourier smoothing technique implemented in SIDPAC program smoo.m for all of the modeling in this work.

Using orthogonal functions to model the response variable makes it possible to evaluate the merit of including each modeling function *independently*, using the predicted squared error PSE. The goal is to select a model structure with minimum PSE, and the PSE always has a single global minimum for orthogonal modeling functions. This makes the model structure determination a well-defined and straightforward process that can be (and was) automated.

B. Generating Empirical Orthogonal Modeling Functions

Empirical orthogonal functions can be generated from ordinary functions in the explanatory variables using a Gram-Schmidt orthogonalization procedure. This approach is described in Refs. [3-7], which are the basis for the material presented here. References [5,8,9] explain how empirical function orthogonalization can be done recursively in real time.

The process begins by choosing one of the ordinary functions as the first orthogonal function. Typically, a vector of ones (associated with the bias term in the model) is chosen as the first orthogonal function,

$$\boldsymbol{p}_1 = \mathbf{1} \tag{23}$$

In general, any function of the explanatory variables can be chosen as the first orthogonal function, without any change in the procedure. To generate the next orthogonal function, an ordinary function is made orthogonal to the preceding orthogonal function(s). Define the *j*th orthogonal function $\boldsymbol{p}_j$ as

$$\boldsymbol{p}_j = \boldsymbol{\xi}_j - \sum_{k=1}^{j-1} \gamma_{kj}\, \boldsymbol{p}_k \qquad\qquad j = 2, 3, ..., n_t \tag{24}$$

where $\boldsymbol{\xi}_j$ is the *j*th ordinary function vector. For example, each $\boldsymbol{\xi}_j$ could be some ordinary polynomial function of the explanatory variables. The γ_{kj} for $k = 1, 2, ..., j-1$ are scalars determined by multiplying both sides of Eq. (24) by $\boldsymbol{p}_k^T$, then invoking the mutual orthogonality of the $\boldsymbol{p}_k$, $k = 1, 2, ..., j$, and solving for γ_{kj}

$$\gamma_{kj} = \frac{\boldsymbol{p}_k^T \boldsymbol{\xi}_j}{\boldsymbol{p}_k^T \boldsymbol{p}_k} \qquad\qquad k = 1, 2, ..., j-1 \tag{25}$$

The same process can be implemented in sequence for each ordinary function $\boldsymbol{\xi}_j$, $j = 2, 3, \ldots, n_t$. The total number of ordinary functions used as raw material for generating the empirical orthogonal functions, including the bias term, is n_t. From Eqs. (23)-(25), it follows that each orthogonal function can be expressed exactly in terms of a linear expansion of the original

functions. The orthogonal functions are generated sequentially, and each orthogonal function can be considered an orthogonalized version of an original function.

The empirical orthogonal function generation method described here normally starts by generating all candidate ordinary functions in the explanatory variables. Note that the ordinary functions can be assembled from multiple explanatory variables, and this has been done in previous flight test applications [3-9]. An example of a pool of candidate ordinary functions might be all possible ordinary polynomial functions in the explanatory variables, up to a selected maximum order. If it turns out that a particular candidate modeling function is not needed in the model, the algorithm will not select the orthogonal function associated with that candidate modeling function. This occurs naturally and automatically in the course of the model structure determination process described earlier. Therefore, there is no harm in including explanatory variables and modeling functions that might not be important, except that additional computation time will be required to identify the model structure, because additional orthogonal functions will be generated and sorted. The final identified model will be the same. Consequently, the choices that the analyst needs to make initially to define the pool of candidate modeling functions are easy and not critical to the quality of the final modeling results. The form of the candidate ordinary functions is not limited to polynomials, as described in this example, and can be any arbitrary function of the explanatory variables [5-7].

If the $\boldsymbol{p}_j$ vectors and the $\boldsymbol{\xi}_j$ vectors are arranged as columns of matrices $\boldsymbol{P}$ and $\boldsymbol{X}$, respectively, and γ_{kj} are elements in the *k*th row and *j*th column of an upper triangular matrix $\boldsymbol{G}$ with ones on the diagonal,

$$\boldsymbol{G} = \begin{bmatrix} 1 & \gamma_{12} & \gamma_{13} & \cdots & \gamma_{1n_t} \\ 0 & 1 & \gamma_{23} & \cdots & \gamma_{2n_t} \\ 0 & 0 & 1 & \cdots & \gamma_{3n_t} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{26}$$

then

$$\boldsymbol{X} = \boldsymbol{P}\boldsymbol{G} \tag{27}$$

which leads to

$$\boldsymbol{P} = \boldsymbol{X}\boldsymbol{G}^{-1} \tag{28}$$

The columns of $\boldsymbol{G}^{-1}$ contain the coefficients for an exact linear expansion of each column of $\boldsymbol{P}$ (i.e., each empirical orthogonal function) in terms the original functions in the columns of $\boldsymbol{X}$, which have physical meaning.

C. Orthogonal Function Generation Order

The method described in the preceding section for generating empirical orthogonal functions is sequential, and therefore has an order dependence. Consequently, processing candidate ordinary functions in a different order will produce a different set of empirical orthogonal functions, and in some cases, a different model. The reason is that the candidate ordinary functions often are mutually correlated to some extent, so that parts of them are proportional. In the extreme case where two candidate ordinary functions are linearly dependent, the orthogonalization process will include the first one and discard the second, because the second will have no distinct information to contribute to the model. A more common case occurs when a relatively important term is orthogonalized late in the process, which incurs unnecessary numerical inaccuracy when converting the orthogonal modeling terms selected for the model back to ordinary functions, which are physically meaningful.

To address this problem, the orthogonalization process described in the preceding section was modified to what might be called a forward dynamic programming orthogonalization. At each stage of the orthogonalization, each remaining candidate ordinary function is orthogonalized as the next in line for orthogonalization, rather than orthogonalizing the entire pool of candidate ordinary functions all at once at the start. This modified process evaluates each candidate ordinary function at each stage of the orthogonalization, as if it were the next in line for orthogonalization, and chooses the orthogonal function with the most capability to reduce the mean squared fit error (quantified by the quantity on the left side of inequality (21)) as the next orthogonal function in the sequence. The modified approach requires more computation, because the orthogonalization is repeated at each step for each candidate ordinary function that has not yet been orthogonalized. However, there is a significant advantage to doing the orthogonalization in this way, in that the most important functions are orthogonalized first, which keeps the number of model terms in the final model as small as possible.

To implement the approach, Eqs. (24)-(25) are applied individually to each candidate ordinary function not yet orthogonalized. Then the reduction in squared fit error that would result from adding the orthogonalized function associated with the *j*th candidate ordinary function to the model is computed as (see Eq. (16) or the left side of inequality (21))

$$\Delta_j = \left(\boldsymbol{p}_j^T \boldsymbol{z} \right)^2 \Big/ \left(\boldsymbol{p}_j^T \boldsymbol{p}_j \right) \tag{29}$$

At each step, the largest value of Δ_j is used to select the next empirical orthogonal function, and that function is retained as the most recent in the orthogonalization. Then the process repeats until all of the candidate ordinary functions have been orthogonalized or the data information has been exhausted.

This approach uses the decoupled modeling capability of empirical orthogonal functions to determine the order in which the candidate ordinary functions are orthogonalized. The candidate ordinary functions are orthogonalized in the order of most to least potential for their corresponding orthogonal function to be selected for the model. This makes the model term selection process accurate and reduces the number of terms in the final model, which is good for both prediction accuracy and model parameter accuracy.

D. Conversion to Physically-Meaningful Models

After the model structure is determined using empirical orthogonal modeling functions for minimum PSE, the identified model output is

$$\hat{\boldsymbol{y}} = \boldsymbol{P}\hat{\boldsymbol{\theta}} \tag{30}$$

where the $\boldsymbol{P}$ matrix now includes only the n orthogonal functions selected in the model structure determination, $n \le n_t$. Each retained orthogonal modeling function can be decomposed without error into an expansion of the original functions in the explanatory variables, using the columns of $\boldsymbol{G}^{-1}$ in Eq. (28) corresponding to the retained orthogonal functions. Common terms are combined using double precision arithmetic to arrive finally at a model using only ordinary functions in the explanatory variables. Terms that contribute less than 0.1 percent of the final model root-mean-square magnitude are dropped.

The final form of the model is a sum of ordinary functions in the explanatory variables, with associated model parameter estimates. There is no parameter uncertainty introduced from the conversion from orthogonal functions to ordinary functions, because Eq. (28) is exact. The uncertainty in the model parameters associated with the empirical orthogonal function model can be computed simply from Eq. (10), with only the empirical orthogonal functions selected for the model included in the calculation, and noting that $\left(\boldsymbol{P}^T\boldsymbol{P}\right)^{-1}$ will be a diagonal matrix because the columns of $\boldsymbol{P}$ are mutually orthogonal by design. The fit error variance estimate in Eq. (12) can be upgraded to account for colored residuals [5,6] after the modeling process is complete. Because the uncertainty in each model parameter in the orthogonal function model can be computed in a straightforward way, and each ordinary function is a linear combination of orthogonal functions by Eq. (28), the computation of the uncertainties in the model parameters for the final ordinary function model is a simple bookkeeping exercise.

III. Results

Computational fluid dynamics (CFD) data for an airfoil undergoing various pitching motions at Mach number 0.3 and Reynolds number 6 million were collected at a sampling rate of 10,000 Hz, then downsampled to 200 Hz for analysis and modeling. The data were composed of time series for angle of attack in degrees and nondimensional aerodynamic coefficients for lift force C_L, drag force C_D, and pitching moment C_m. Details of the unsteady, three-dimensional, compressible Navier-Stokes flow solver used to generate the data can be found in Ref. [10]. Because the modeling is based on information embodied in the data, modeling different aerodynamics at other flight conditions would require additional CFD runs to produce data with those characteristics.

The left side of Fig. 2 shows modeling data for angle of attack, which was a chirp signal with increasing frequency. The plots on the right show the data for nondimensional aerodynamic coefficients $C_L, C_D,$ and C_m, plotted against angle of attack in degrees. The plots indicate approximately linear dependency for C_L, but nonlinear unsteady dependencies for C_D and C_m.

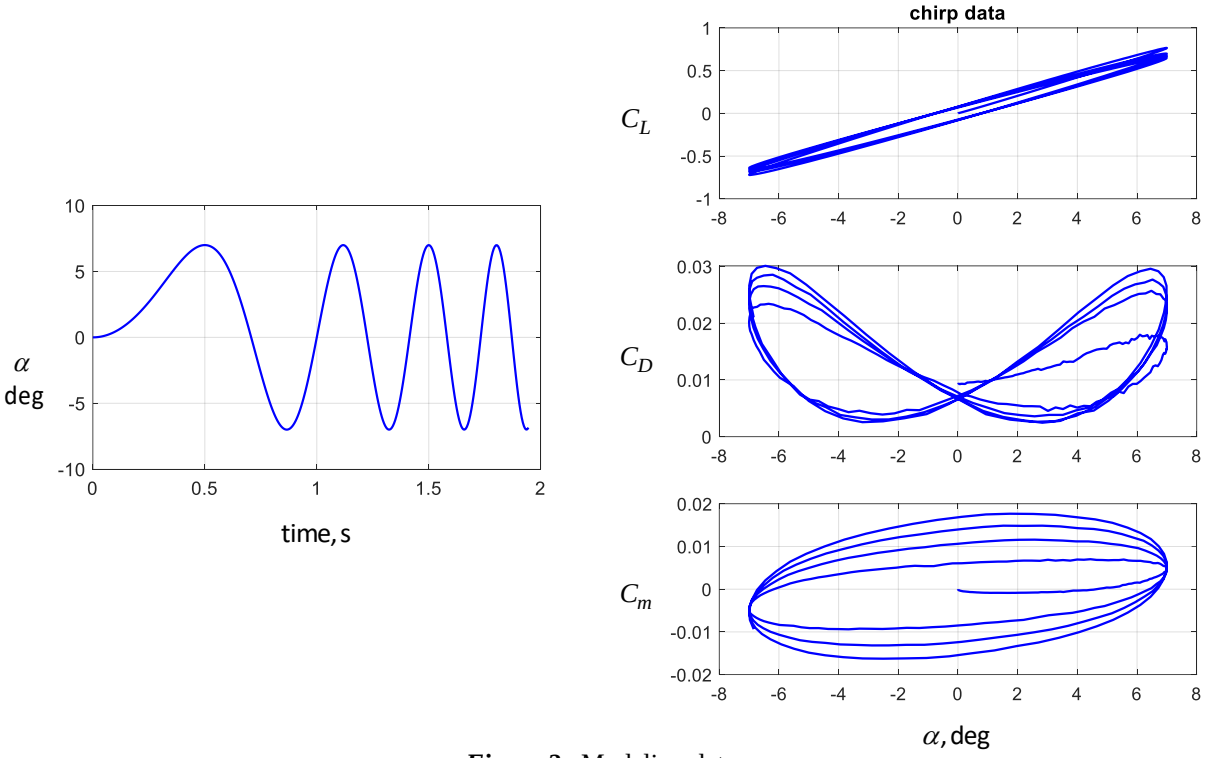


Figure 2. Modeling data

Figure 3 shows the same data transformed into the frequency domain using a high-accuracy finite Fourier transform [5,6]. The C_L data are clearly proportional to the angle of attack data, whereas the data for both C_D and C_m have distortions and higher-frequency content typical of nonlinear unsteady effects. The significant data information for all of the time series appears at frequencies below 10 Hz.

Nonlinear quasi-steady modeling was attempted first, using generally nonlinear functions of the current values of the angle of attack to model current values of $C_L, C_D,$ and C_m . Empirical orthogonal function modeling was applied with candidate polynomial functions up to 5th order. As expected, this approach failed because of the unsteady aerodynamics apparent in the response plots shown in Fig. 2.

Empirical orthogonal function modeling was then applied using candidate polynomial modeling functions up to 3rd order, with explanatory variables composed of the current value of angle of attack and past values of angle of attack in radians with time lags up to 0.3 s. For both C_L and C_D modeling, time lags up to 0.3 s were found to be necessary, but the modeling did not need the finest resolution of time lag that was available using 200 Hz data (5 ms). This was found by repeated modeling using angle of attack with various time lags as the explanatory variables for the candidate model terms. In these cases, the angle of attack time lags for the explanatory variables were $\tau = [0 : 0.025 : 0.3]$ s . For C_m modeling, a finer mesh of time lags was needed, but the maximum lag required was smaller. The angle of attack lags for the explanatory variables used in the C_m modeling were $\tau = [0 : 0.005 : 0.2]$ s .

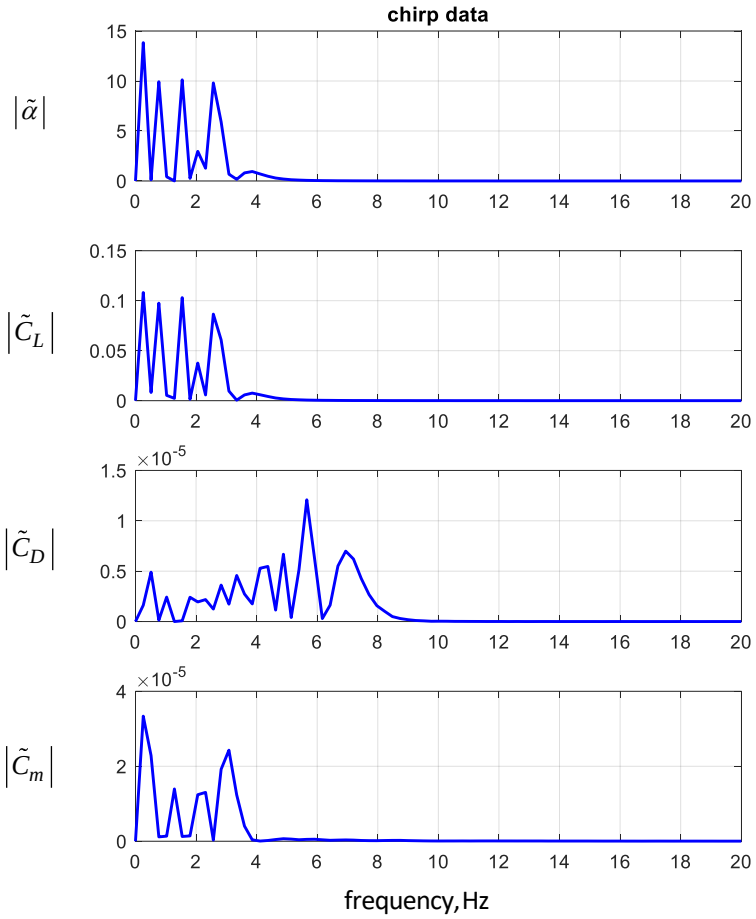


Figure 3. Spectral content of computational fluid dynamics data for an airfoil

In theory, all of the modeling could have been done by including every possible time lag of the angle of attack up to some maximum lag, for example 0.5 s, and some maximum polynomial order, such as 5th order. However, it was more efficient to investigate subsets of those possibilities to identify the important time lags and polynomial model orders. This can be done readily with the empirical orthogonal function modeling program in SIDPAC called mof.m, which selects the best functions for the model from a pool of candidate model terms using the approach described earlier.

The identified models were

$$C_L(i) = \theta_{L_1}\alpha(i) + \theta_{L_2}\alpha(i-15) + \theta_{L_3}\alpha^2(i-40)\alpha(i-45) + \theta_{L_4}\alpha(i-5)\alpha^2(i-60) \tag{31}$$

$$\begin{aligned} C_D(i) = {} & \theta_{D_1} + \theta_{D_2}\alpha^2(i) + \theta_{D_3}\alpha(i)\alpha(i-20) + \theta_{D_4}\alpha(i)\alpha(i-25) + \theta_{D_5}\alpha(i)\alpha(i-30) \\ & + \theta_{D_6}\alpha(i-10)\alpha(i-45) + \theta_{D_7}\alpha(i-15)\alpha(i-40) + \theta_{D_8}\alpha^2(i-35) \end{aligned} \tag{32}$$

$$C_m(i) = \theta_{m_1}\alpha(i-1) + \theta_{m_2}\alpha(i) + \theta_{m_3}\alpha(i-13) + \theta_{m_4}\alpha(i)\alpha^2(i-8) \tag{33}$$

Model parameter values and the associated standard errors for these discrete-time nonlinear unsteady aerodynamic models are given in Table 1. Figures 4, 5, and 6 show the model fits to the

response data for C_L , C_D , and C_m , respectively. The plots on the left show the model fit as a function of time, and the plots on the right show the model fit as a function of angle of attack. The lower plots show the residuals, which are the difference between the measured response data and the model output shown in the upper plots. The plots show excellent model fits to the measured response data, and the residuals show little deterministic content. Percent errors for the C_L , C_D , and C_m model fits, calculated as 100 times the root-mean-square fit error divided by the root-mean-square of the measured response, were 0.62 percent, 1.5 percent, and 1.66 percent, respectively.

For the C_L model, the linear quasi-steady term associated with θ_{L_1} modeled 89 percent of the variation in measured C_L , indicating a mostly linear quasi-steady response, whereas for the C_D model, the squared quasi-steady term associated with θ_{D_2} modeled only 24 percent of the variation in measured C_D . For the C_m model, the linear quasi-steady term associated with θ_{m_2} modeled only 15 percent of the variation in measured C_m . These results are consistent with the earlier conclusions regarding nonlinear and unsteady effects based on analysis of Figs. 2 and 3.

Table 1. **Discrete-Time Nonlinear Model Parameters**

Parameter	Estimate (Standard Error)	Parameter	Estimate (Standard Error)	Parameter	Estimate (Standard Error)
θ_{L_1}	5.553 (0.003)	θ_{D_1}	0.0095 (0.0000)	θ_{m_1}	1.402 (0.003)
θ_{L_2}	0.751 (0.002)	θ_{D_2}	1.938 (0.006)	θ_{m_2}	–1.352 (0.003)
θ_{L_3}	14.943 (0.187)	θ_{D_3}	–8.919 (0.048)	θ_{m_3}	–0.0261 (0.0003)
θ_{L_4}	–13.651 (0.342)	θ_{D_4}	13.487 (0.084)	θ_{m_4}	0.902 (0.021)
		θ_{D_5}	–6.346 (0.042)		
		θ_{D_6}	0.668 (0.011)		
		θ_{D_7}	–0.472 (0.011)		
		θ_{D_8}	–0.111 (0.004)		

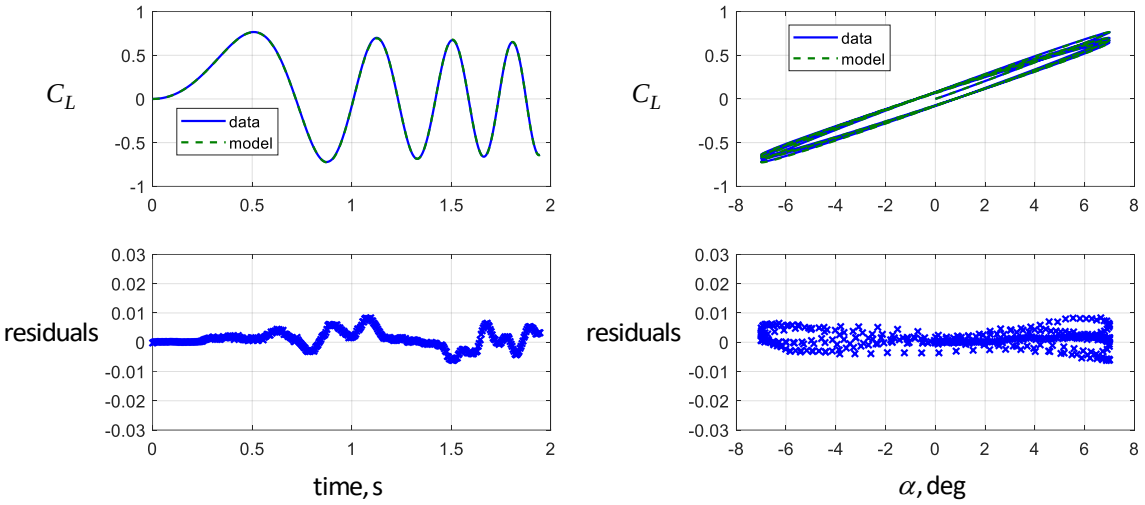


Figure 4. C_L model

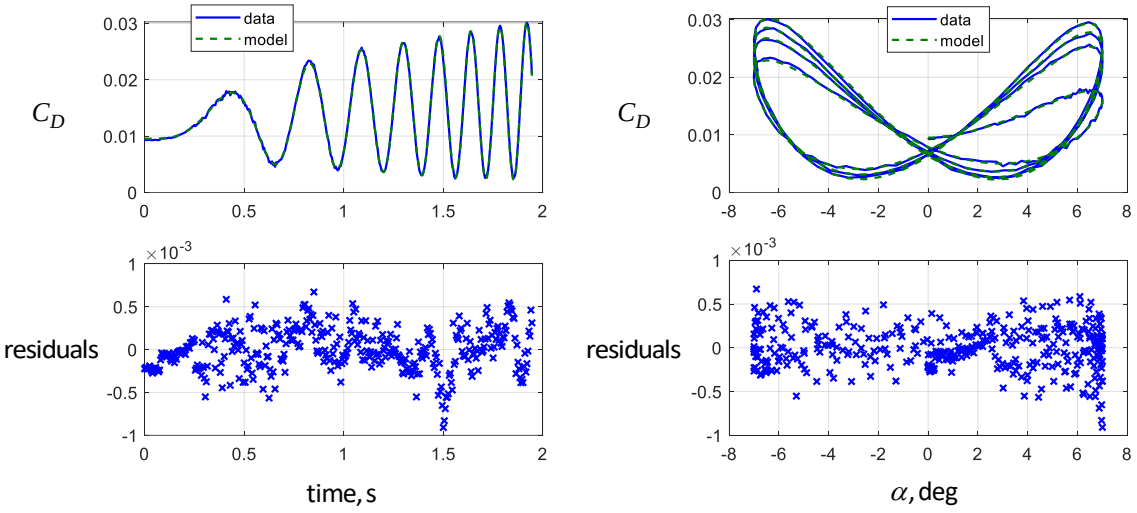


Figure 5. C_D model

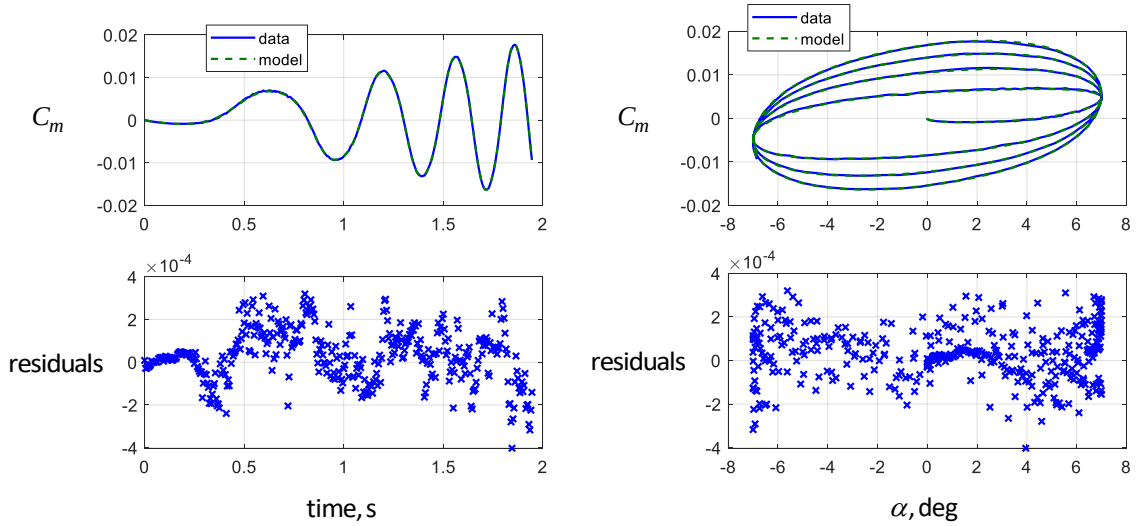


Figure 6. C_m model

Figure 7 shows data used for prediction testing of the identified models. The left side of Fig. 7 shows time series data for angle of attack in degrees, which was a oscillation with increasing amplitude. The plots on the right show the data for nondimensional aerodynamic coefficients $C_L, C_D,$ and C_m , plotted against angle of attack in degrees. The variation in angle of attack for this prediction data was different than the angle of attack variation for the data used to identify the models. The variation in angle of attack shown in Fig. 7 can produce spirals for the nonlinear unsteady aerodynamic response, as shown on the right side of the figure. None of the data shown in Fig. 7 was used in the modeling process.

The identified models from Eqs. (31)-(33) and Table 1 were applied to the data shown in Fig. 7. Figures 8, 9, and 10 show predictions of the response data for C_L , C_D , and C_m , respectively. The predictions are similar in quality to the fits obtained during the model identification process, which is a strong indicator of accurate modeling. Percent errors for the C_L, C_D, and C_m model predictions, computed as 100 times the root-mean-square prediction error divided by the root-mean-square of the measured response, were 1.25 percent, 2.07 percent, and 2.07 percent, respectively.

Although the modeling demonstrated in this work was done for a single response at a time and a single physical explanatory variable with various time lags, the same approach could be used with more than one physical explanatory variable and associated lagged versions of those quantities. This would of course involve more computations in the empirical orthogonal modeling process, and would require rich information content in the multivariate data, but there are otherwise no restrictions to the applicability of the approach to this more complex modeling problem.

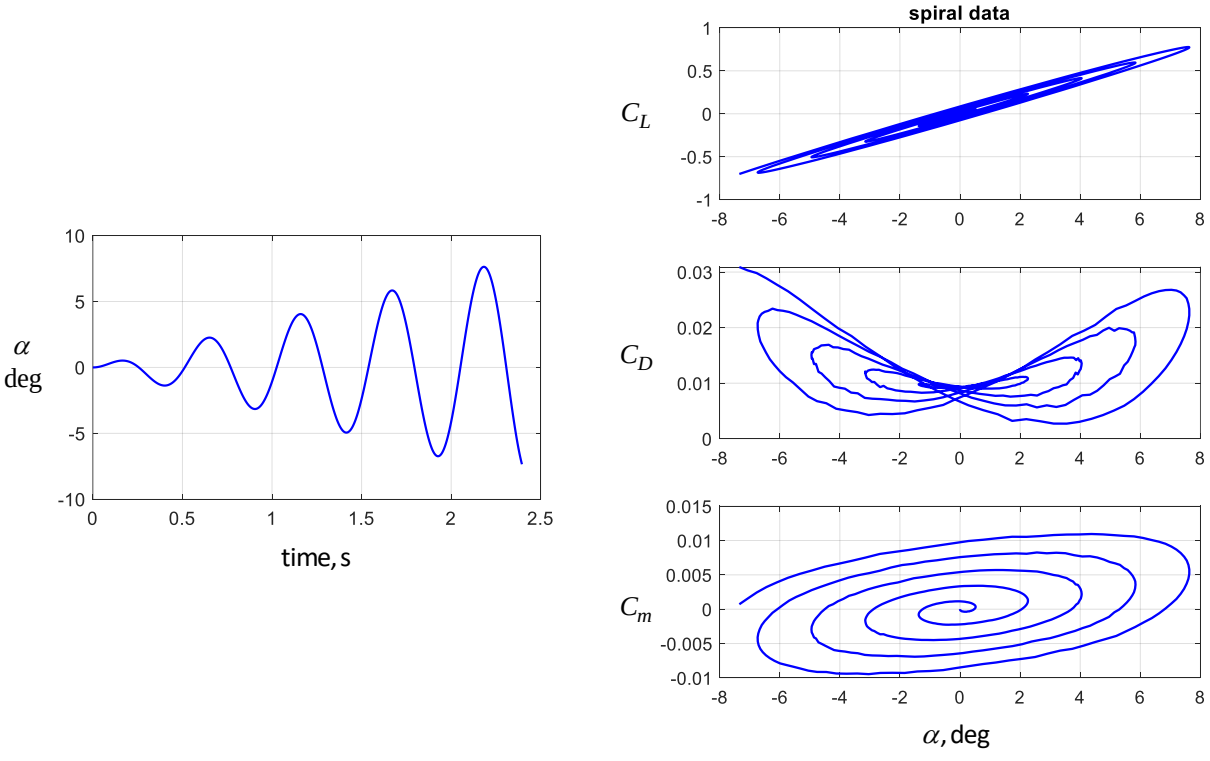


Figure 7. Prediction data

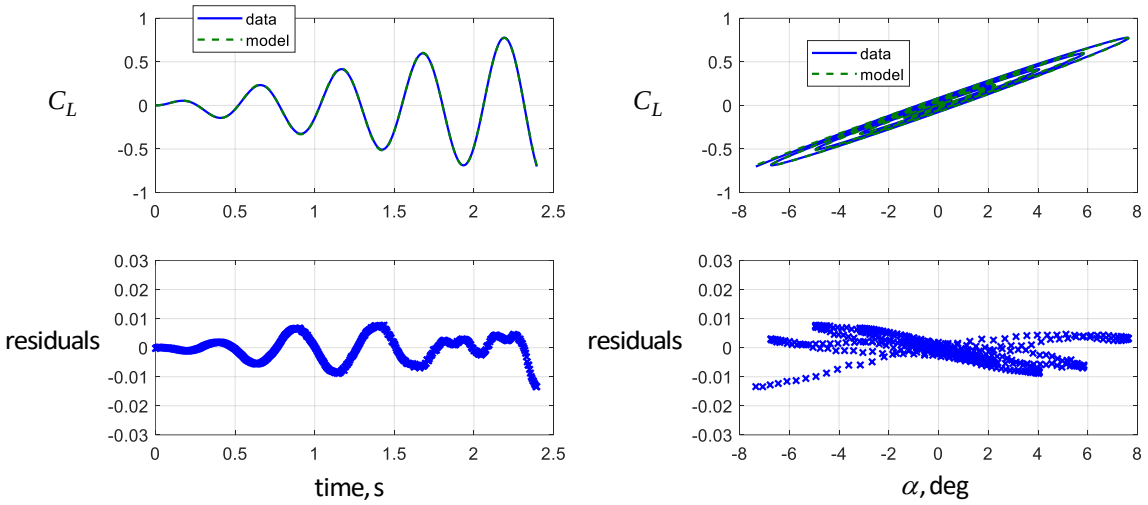


Figure 8. Lift coefficient prediction

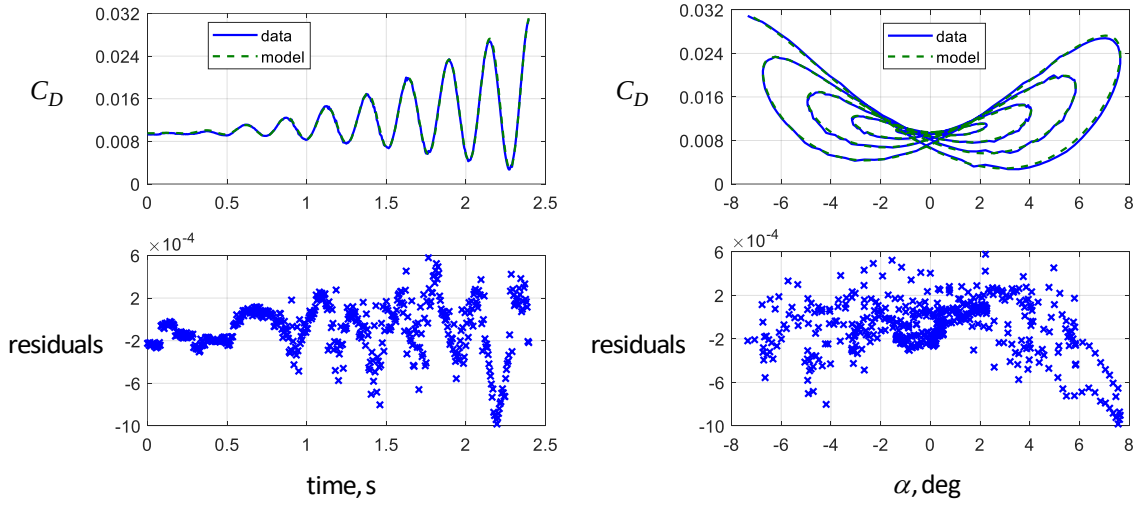


Figure 9. Drag coefficient prediction

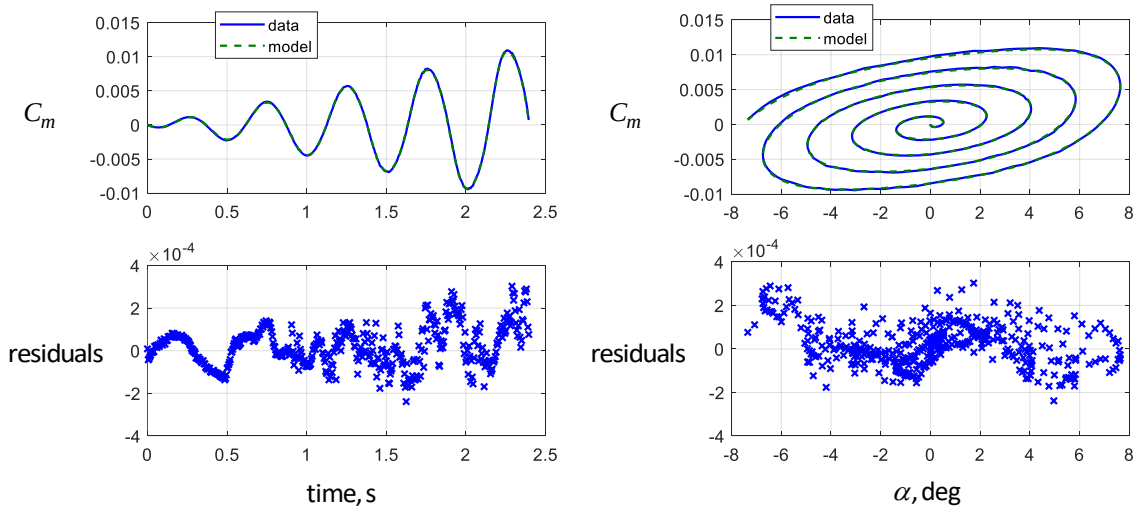


Figure 10. Pitching moment coefficient prediction

IV. Conclusions

Empirical orthogonal function modeling was applied to identify compact and accurate discrete-time nonlinear unsteady models from computational fluid dynamics data for an airfoil undergoing various pitching motions. Nonlinear unsteady aerodynamic models for nondimensional longitudinal aerodynamic coefficients were identified using generally nonlinear polynomial modeling terms involving current and past values of the angle of attack. The identified models characterized the modeling data well and exhibited excellent prediction capability for data that were not used in the modeling process. Model structures, model parameter values, and uncertainty measures were determined from the measured data alone, using data record lengths less than 2 s.

The modeling approach demonstrated that past values of the explanatory variable(s) can be used in lieu of time-derivative information to model the unsteady aerodynamic effects. This has practical significance, because time-derivative information can be difficult to compute accurately, e.g., for low data sampling rates, real-time simulation, or real-time aerodynamic modeling.

The application examples shown in this work were for a single physical explanatory variable (angle of attack) and a single response variable (nondimensional longitudinal aerodynamic coefficients for lift, drag, and pitching moment, considered one at a time), but the approach can be extended to use multiple physical explanatory variables to model each measured response. This would require additional computations, along with adequate data information content in the multivariate data. The modeling process can be applied to individual measured responses one at a time, as was demonstrated in the application examples.

The empirical function modeling approach explained and demonstrated in this work is general and can be used to model complex dependencies in other applications. For aerodynamic modeling, nonlinear model terms using both present and past values of the explanatory variable data provide the capability to model general nonlinear unsteady aerodynamic effects. No analyst judgment is required for the model identification, apart from initially postulating the maximum possible complexity of the model. This can be done generously without any adverse effects on the final identified model, but incurs increased computations. A more efficient approach is to use repeated applications of empirical orthogonal function modeling for subsets of all possible time lags and model term complexity to identify an appropriate set of explanatory variables and candidate model terms.

Acknowledgments

Thanks to Dr. Mehdi Ghoreyshi from the U.S. Air Force Academy for providing the computational fluid dynamics data and to Dr. Jae Lee from NAVAIR for funding for this work.

References

1. Billings, S.A. *Nonlinear System Identification – NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*, John Wiley & Sons, Ltd., West Sussex, UK, 2013.

2. Chen, S., Billings, S.A. and Luo, W. "Orthogonal Least Squares Methods and Their Application to Nonlinear System Identification," ACSE Report 343, Dept. of Automatic Control and System Engineering, University of Sheffield, Sheffield, UK, 1988.

3. Morelli, E.A. "Nonlinear Aerodynamic Modeling using Multivariate Orthogonal Functions," AIAA Paper 93-3636, *AIAA Atmospheric Flight Mechanics Conference*, Monterey, CA, August 1993.

4. Morelli, E.A. "Global Nonlinear Aerodynamic Modeling using Multivariate Orthogonal Functions," *Journal of Aircraft*, Vol. 32, No. 2, March-April 1995, pp. 270-77.

5. Morelli, E.A. and Klein, V. *Aircraft System Identification - Theory and Practice*, 2nd Edition, Sunflyte Enterprises, Williamsburg, VA, December 2016.

6. SIDPAC, NASA Software Catalog, https://software.nasa.gov/software/LAR-16100-1, accessed 06 January 2021.

7. Morelli, E.A. "Efficient Global Aerodynamic Modeling from Flight Data," AIAA-2012-1050, *50th AIAA Aerospace Sciences Meeting*, Nashville, TN, January 2012.

8. Brandon, J.M. and Morelli, E.A. "Real-Time Global Nonlinear Aerodynamic Modeling from Flight Data," *Journal of Aircraft*, Vol. 53, No. 5, September-October 2016, pp. 1261-1297.

9. Morelli, E.A. "Practical Aspects of Real-Time Modeling for the Learn-To-Fly Concept," AIAA-2018-3309, *Atmospheric Flight Mechanics Conference, AIAA Aviation Forum*, Atlanta, GA, June 2018.

10. Ghoreyshi, M., Jirásek, A., and Cummings, R.M. "Reduced Order Unsteady Aerodynamic Modeling for Stability and Control Analysis using Computational Fluid Dynamics," *Progress in Aerospace Sciences*, Vol. 71, November 2014, pp. 167-217.

11. SIDPAC User List, http://sunflyte.com/SIDBook_SIDPAC.htm#SIDPAC_Users, accessed 06 January 2021.

12. Greenwell, D.I. "A Review of Unsteady Aerodynamic Modelling for Flight Dynamics of Manoeuvrable Aircraft," AIAA-2004-5276, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Providence, RI, August 2004.

13. Barron, A.R., "Predicted Squared Error : A Criterion for Automatic Model Selection," *Self-Organizing Methods in Modeling*, Farlow, S.J., Ed., Marcel Dekker, Inc., New York, NY, 1984, pp. 87-104.

14. Morelli, E.A. and DeLoach, R., "Wind Tunnel Database Development using Modern Experiment Design and Multivariate Orthogonal Functions," AIAA Paper 2003-0653, 41st *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January 2003.

15. Morelli, E.A. "Estimating Noise Characteristics from Flight Test Data using Optimal Fourier Smoothing," *Journal of Aircraft*, Vol. 32, No. 4, July-August 1995, pp. 689-695.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/02/2021	TECHNICAL MEMORANDUM	

4. TITLE AND SUBTITLE

Nonlinear Unsteady Aerodynamic Modeling using Empirical Orthogonal Functions

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Morelli, Eugene A.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

031102.02.R7.05.9287.17

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-001

10. SPONSOR/MONITOR'S ACRONYM(S)

NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

NASA/TM-20210009593

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Empirical orthogonal function modeling is explained and applied to identify compact discrete-time nonlinear unsteady aerodynamic models from data generated by an unsteady three-dimensional compressible Navier-Stokes flow solver for an airfoil undergoing various pitching motions. Model structures, model parameter estimates, and model parameter uncertainty estimates for nondimensional lift, drag, and pitching moment coefficient models were determined autonomously and directly from the data. Prediction tests using data that were not used in the modeling process showed that the identified models exhibited excellent prediction capability, which is a strong indicator of an accurate model.

15. SUBJECT TERMS

empirical; orthogonal; modeling; function; nonlinear; unsteady; aerodynamics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			HQ - STI-infodesk@mail.nasa.gov
U	U	U	UU	27	19b. TELEPHONE NUMBER *(Include area code)* 757-864-9658