

Aircraft Parameter Estimation Considering Process and Measurement Noise

Jared A. Grauer* and Eugene A. Morelli†
NASA Langley Research Center, Hampton, Virginia 23681

<https://doi.org/10.2514/1.C038080>

A practical formulation is proposed for aircraft parameter estimation using the filter-error method, which is a maximum likelihood estimator for dynamic systems with both process noise and measurement noise. The novelty of the proposed formulation is that by accurately estimating the measurement noise covariance matrix using a time series analysis method, the remaining unknowns (which include the unknown parameters in the state-space matrices and the process noise covariance matrix) become decorrelated and can be estimated simultaneously in a straightforward manner. The approach is demonstrated using simulation data and flight test data from a subscale airplane. Results indicate that the proposed algorithm can produce accurate modeling results when both measurement noise and process noise are present in the data.

Nomenclature

$\text{cov}(\cdot)$	=	covariance
$E[\cdot]$	=	expected value
\ln	=	natural logarithm
\Re	=	real number
∂	=	partial derivative
$\delta(t)$	=	Dirac delta function
δ_{ij}	=	Kronecker delta function
$ \cdot $	=	absolute value or determinant

Superscripts

T	=	transpose
$^{-1}$	=	matrix inverse
\cdot	=	time derivative
$\hat{\cdot}$	=	estimated value

I. Introduction

PARAMETER estimation is the process of determining values and uncertainties of unknown constants in a mathematical model so that model outputs best match experimental data. There are numerous uses for parameter estimation in aerospace applications, such as quantifying aircraft stability and control characteristics, determining flight instrumentation errors, representing high-fidelity simulations with reduced-order models, evaluating flying qualities, determining actuator or sensor transfer functions, and many others [1–4].

Maximum likelihood approaches for aircraft parameter estimation, which have been used widely in practice, can be classified by assumptions made in the modeling problem. In the equation-error (EE) method, process noise is considered but measurement noise is neglected. Conversely, in the output-error (OE) method, measurement noise is considered but process noise is neglected. Both of these methods involve a noise assumption that is not strictly correct because flight test applications always include some amount of both process and measurement noise. Despite this fact, EE and OE

have been successfully used for decades, and several improvements have been made to mitigate the shortcomings of these approaches.

The maximum likelihood estimator that considers both process noise and measurement noise is called the filter-error (FE) method. Although FE is the most general approach in this family of estimators and contains the most complete theory, it is also the most difficult to use, in part due to the increased complexity of the estimation problem, issues with convergence and identifiability, difficulty in interpreting the results, and the need for more informative data [4]. For those reasons and others, EE and OE have remained the standard approaches for practical parameter estimation at NASA Langley Research Center (LaRC) and at other institutions [2,4–7].

Nevertheless, there are several routinely encountered problems for which FE would be a useful, if not the preferred, approach for parameter estimation. One example is flight in atmospheric turbulence, which is usually present to some extent, where gust velocities act as process noise and influence the aircraft flight dynamics. Alternatives to using FE for parameter estimation in this case include flight testing in calm conditions, increasing the excitation amplitudes to dominate the effects of turbulence [8], and using measured or reconstructed turbulence in the estimation [9,10]. Another example is data compatibility analysis, where kinematically related measurements are compared to estimate and remove instrumentation errors such as biases, scale factors, and time skews. In this problem, which is examined in more detail later, the process noise is due to noise on the measured inputs used for data compatibility. Nonlinear versions of the Kalman filter or complex iterative filter-smoothers have been applied to this problem [11,12], but OE is a commonly used approach [4,13]. Lastly, model structure error, which is always present in practical modeling problems, can be regarded as a form of process noise, although this can also be addressed with other techniques [4].

In this paper, a practical formulation of the FE parameter estimation method is presented and demonstrated. The key to this formulation is to separately estimate the measurement noise covariance matrix using time series analysis, then treat that as a fixed and known quantity during the subsequent parameter estimation. Measurement noise is a physical quantity that can be independently determined using time series analysis techniques. Fixing this value during the subsequent parameter estimation breaks correlations among the remaining unknowns (which include parameters in the dynamic model equations) and eliminates convergence issues that have plagued other implementations of FE. With this new formulation, parameter estimation can proceed in a straightforward manner using relatively simple optimization methods.

The paper is organized as follows: Section II introduces the parameter estimation problem. Section III summarizes other parameter estimation approaches, including previous implementations of FE, to provide context for the present formulation. Section IV discusses

Presented as Paper 2015-2704 at the AIAA Aviation 2015 Forum, Dallas, TX, June 22–26, 2015; received 31 May 2024; accepted for publication 2 September 2024; published online 20 November 2024. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

*Research Aerospace Engineer, Dynamic Systems and Control Branch, MS 308. Associate Fellow AIAA.

†Research Aerospace Engineer, Dynamic Systems and Control Branch, MS 308. Fellow AIAA.

details of the proposed FE implementation. Section V demonstrates the approach using a simulation example involving the roll dynamics of an airplane. Section VI demonstrates the approach using a data compatibility analysis with flight test data for the X-56A airplane. Section VII concludes the paper.

II. Problem Description

This section briefly summarizes the parameter estimation problem considered. For more details and a complete description of the theory, see Refs. [4,14] and the references therein.

Consider the linear time-invariant (LTI) model

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{G}\mathbf{w}(t) \quad (1a)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{H}\mathbf{w}(t) \quad (1b)$$

$$\mathbf{z}(i) = \mathbf{y}(i) + \mathbf{v}(i), \text{ for } i = 1, 2, \dots, N \quad (1c)$$

Bold-italic lower-case symbols are vectors, bold upper-case symbols are matrices, and nonbold symbols are scalars. The vector $\mathbf{u}(t) \in \mathfrak{R}^{n_u}$ is the deterministic input to the system and is assumed to elicit sufficient data information in the response for accurate identification. The vector $\mathbf{x}(t) \in \mathfrak{R}^{n_s}$ is the state, $\mathbf{y}(t) \in \mathfrak{R}^{n_o}$ is the output, and $\mathbf{z}(i) \in \mathfrak{R}^{n_o}$ is the sampled measurement.

The matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} , \mathbf{G} , and \mathbf{H} are constant matrices that contain model parameters with unknown values. In flight test applications, these parameters are typically stability and control derivatives. The pair (\mathbf{A}, \mathbf{B}) is assumed to be controllable so that the inputs can excite each state. Similarly, the pair (\mathbf{A}, \mathbf{C}) is assumed to be observable so that each state can be reconstructed by the measurements. The system model can be stable or unstable.

The inputs, states, outputs, and measurements represent perturbations about reference values. Equations (1a) and (1b) evolve continuously over time t , whereas Eq. (1c) is valid at discrete instances of time $i\Delta t$, where i is the integer sample index and Δt is the sampling interval. Equations (1a–1c) are therefore a continuous-discrete, state-space representation of the system. Figure 1 depicts this system model as a block diagram.

The vectors $\mathbf{w}(t) \in \mathfrak{R}^{n_w}$ and $\mathbf{v}(i) \in \mathfrak{R}^{n_o}$ in Eq. (1) are the process noise and sampled measurement noise, respectively. These quantities are modeled as random and white noise with

$$E[\mathbf{w}(t)] = \mathbf{0} \quad (2a)$$

$$E[\mathbf{v}(i)] = \mathbf{0} \quad (2b)$$

$$E[\mathbf{w}(t)\mathbf{w}(\tau)^T] = \mathbf{Q}_c(t)\delta(t-\tau) \quad (2c)$$

$$E[\mathbf{v}(i)\mathbf{v}(j)^T] = \mathbf{R}(i)\delta_{ij} \quad (2d)$$

where \mathbf{Q}_c and \mathbf{R} are symmetric and positive definite matrices that are assumed to be uncorrelated with each other. Typically, \mathbf{R} is assumed to be a diagonal matrix to reflect the independence of sensor measurement errors, although both \mathbf{R} and \mathbf{Q}_c may be fully populated.

The initial conditions for the state vector are given by the mean and covariance

$$E[\mathbf{x}(0)] = \mathbf{x}_0 \quad (3a)$$

$$E\{[\mathbf{x}(0) - \mathbf{x}_0][\mathbf{x}(0) - \mathbf{x}_0]^T\} = \mathbf{P}_0 \quad (3b)$$

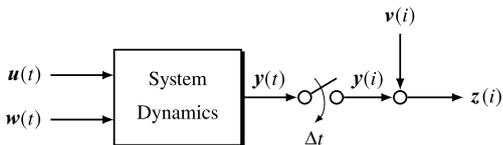


Fig. 1 System model block diagram.

The matrix \mathbf{P}_0 is usually assumed to be diagonal for simplicity. Because the system is stochastic rather than deterministic, a state estimator is needed to solve the equations of motion for the discrete-time output estimate $\hat{\mathbf{y}}(i|i-1)$. This notation means an estimate of \mathbf{y} at sample i using data up to and including sample $i-1$.

All of the unknowns in the system are grouped together in the vector $\boldsymbol{\theta} \in \mathfrak{R}^{n_\theta}$. In general, parameters in $\boldsymbol{\theta}$ can include contributions from the following:

1) Terms in the system matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} , \mathbf{G} , and \mathbf{H} from Eq. (1), listed in the vector $\boldsymbol{\theta}_1$

2) Elements of the process noise spectral density \mathbf{Q}_c from Eq. (2c), or its discrete-time counterpart \mathbf{Q} , listed in the vector

$$\boldsymbol{\theta}_2 = [q_{11} \quad q_{12} \quad \dots \quad q_{n_s n_s}]^T \quad (4)$$

3) Elements of the discrete-time measurement noise covariance \mathbf{R} from Eq. (2d), listed in the vector

$$\boldsymbol{\theta}_3 = [r_{11} \quad r_{12} \quad \dots \quad r_{n_o n_o}]^T \quad (5)$$

4) Initial conditions \mathbf{x}_0 and elements of the associated covariance \mathbf{P}_0 from Eq. (3), listed in the vector

$$\boldsymbol{\theta}_4 = [x_{0_1} \quad x_{0_2} \quad \dots \quad x_{0_{n_s}} \quad p_{0_{11}} \quad p_{0_{12}} \quad \dots \quad p_{0_{n_s n_s}}]^T \quad (6)$$

The total model parameter vector to be estimated is then assembled as

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \boldsymbol{\theta}_3 \\ \boldsymbol{\theta}_4 \end{bmatrix} \quad (7)$$

For a specific application, a subset of the full parameter list in Eq. (7) could be estimated. For example, only the diagonal elements of \mathbf{R} could be estimated. A common practice, which is used in this paper, is that initial conditions can be determined from the data rather than estimated.

The parameter estimation problem is then to determine values and uncertainties for $\boldsymbol{\theta}$ that result in the best match of $\hat{\mathbf{y}}(i|i-1)$ to $\mathbf{z}(i)$ for the entire time series, $i = 1, 2, \dots, N$. Note that the model residuals, which are the differences between the measured responses and the updated estimates $\hat{\mathbf{y}}(i|i)$, are not considered in this problem. Because of this, it is incorrect to use a smoother or state estimator other than the Kalman filter to solve the equations of motion. Different formulations of the estimation problem for which other observers are appropriate can also be developed [14,15] but are not considered in this paper for simplicity and practicality.

In this work, the matching is evaluated using the maximum likelihood approach, which seeks to maximize the likelihood of observing the measured data for the specified model. Maximum likelihood estimators are practical and have been successfully used in aircraft system identification analyses for decades [1,2,4,5,7,13,14,16]. These estimators have a number of favorable properties, including being asymptotically unbiased, consistent, and efficient [4,14].

The likelihood function evaluated at sample i is

$$\mathbb{L}[\mathbf{z}(i); \boldsymbol{\theta}] = (2\pi)^{-\frac{n_o}{2}} |\mathbf{S}(i)|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \boldsymbol{\nu}^T(i) \mathbf{S}^{-1}(i) \boldsymbol{\nu}(i) \right] \quad (8)$$

where

$$\boldsymbol{\nu}(i) = \mathbf{z}(i) - \hat{\mathbf{y}}(i|i-1) \quad (9)$$

is the innovation and $\mathbf{S}(i)$ is the innovation covariance. Instead of maximizing the likelihood function directly, it is more convenient and computationally simpler to minimize the negative logarithm of the likelihood function. These two problems are equivalent because

logarithms are monotonic functions. For N data points, the cost function to be minimized then simplifies to

$$J = \frac{1}{2} \sum_{i=1}^N \nu^T(i) S^{-1}(i) \nu(i) + \frac{1}{2} \ln |S(i)| \quad (10)$$

The first term in the cost function is the innovation sum of squares, weighted by the innovation covariance matrix inverse. The second term reflects the size of the innovation covariance matrix. A third term, which was removed from the cost function and not shown in Eq. (10), is a constant that does not affect the optimization. In this way, the maximum likelihood estimator seeks the parameter estimates that minimize the weighted square and covariance of innovations, based on measurements of the potentially multiple-input multiple-output (MIMO) system.

III. Parameter Estimation Methods

This section briefly summarizes some common approaches for aircraft parameter estimation to provide context for the proposed FE formulation. This summary is not intended to be comprehensive. For more information on the approaches mentioned, see Refs. [4,14,16]. Unless otherwise noted, the noise statistics are assumed to be constant with time, for simplicity, and arguments are omitted in the notation.

A. Equation Error

The EE method is the maximum likelihood estimator when process noise is considered but measurement noise is neglected. Although in practice all measurements contain noise, accurate results can be obtained, even in real time, by smoothing explanatory variables or performing the estimation over a reduced bandwidth in the frequency domain with Fourier transform data [4,17]. The estimation problem typically reduces to ordinary least squares, which has an analytical solution for each state equation or aerodynamic coefficient considered. In addition to θ_1 , an estimate of Q can be obtained from the model residuals. However, all explanatory variables must be accurately measured or reconstructed.

B. Output Error

The OE method is the maximum likelihood estimator when measurement noise is considered but process noise is neglected. This formulation is consistent with the fact that sensor data contain measurement noise, and with the standard practice of conducting flight tests for system identification analysis in relatively calm air. Due to these and several other reasons [4], OE is used widely. The parameter estimation problem considered in OE, which is nonlinear and requires iteration, is to estimate θ_1 and R . Because these unknowns are correlated, a relaxation technique is used where one group of unknowns is held fixed while the other group is estimated, and vice versa, until both sets of unknowns converge [4,14]. Other benefits of using OE include that state variables do not necessarily need to be measured, the estimation can be done in the frequency domain with Fourier transform or frequency response data for LTI models, and a time-domain analysis can include arbitrarily nonlinear and time-varying systems. In the presence of significant process noise or model structure error, though, OE can produce biased estimates with large uncertainties or may fail to converge.

C. Filter Error

The FE method is the maximum likelihood estimator that considers both process noise and measurement noise. Previous formulations of FE are reviewed in Refs. [4,18], which are summarized here. According to Ref. [18], previous implementations can be grouped as follows.

The first group is called the natural formulation. These implementations attempt to estimate all the unknowns (θ_1 , Q , R , x_0 , and P_0) simultaneously. This approach leads to high computational costs due to the large number of unknowns and complicated interdependencies created by the Kalman filter. More importantly, convergence issues

and numerical singularities for Q are typically encountered. As a result, very few practical applications of this formulation have been reported, and these have usually required further compromises, such as in Refs. [19,20].

The second group is called the innovation formulation. To circumvent problems associated with estimating both Q and R , the Kalman filter gain matrix K (defined later in Sec. IV.B) and the innovation covariance matrix S are instead estimated using the relaxation technique, along with θ_1 . This formulation eliminates the convergence and numerical problems associated with the natural formulation but introduces other new problems. For instance, K has potentially many elements and is typically fully populated, which increases the number of unknowns to be estimated and can lead to convergence issues and long computation times. Furthermore, elements of K have little physical connection, which makes it difficult to find adequate starting values and to apply intuition when verifying results. Also, because elements of K are estimated rather than computed from other known quantities in the usual way, a more sophisticated optimization technique is needed to constrain the values and keep the resulting filter realizable.

The third formulation could be called the Maine–Iliff formulation [18]. In this method, elements of θ_1 and Q are estimated, followed by S , in a relaxation technique until all parameters converge. This method reduces the computational burden, removes the numerical singularities, and mitigates convergence issues that were incurred by the previous formulations. This formulation was implemented in software at NASA Dryden Flight Research Center (DFRC) [18,21–23] and the Deutsches Zentrum für Luft- und Raumfahrt (DLR) [24–26]. However, this method still requires a sophisticated optimization to respect constraints associated with the Kalman filter.

D. Dual State and Parameter Estimation

Another approach for FE parameter estimation, which is sometimes called the dual estimation problem, is to use a nonlinear Kalman filter to estimate both $x(i)$ and θ_1 given known statistics and initial conditions. This is mechanized by appending the unknown constants to the state vector and sometimes also by exciting the associated differential equations with a computer-generated white noise sequence during the analysis. Because estimated parameters multiply the estimated states, a nonlinear Kalman filter, such as the extended Kalman filter (EKF) or unscented Kalman filter (UKF), is needed. This approach is suited for real-time estimation and has been used with some success [11,24,27]. This formulation also has the potential to track time-varying dynamics. However, nonlinear Kalman filters were developed for state estimation and can incur convergence issues when applied to parameter estimation. Furthermore, they rely on accurate knowledge of Q and R [28] and are in general inferior to batch-processed maximum likelihood estimates, which take multiple passes through the data [18].

IV. Proposed Formulation

This section presents the details of the proposed FE formulation. The main novelty of this formulation, which drives the subsequent nature of its solution, is that $R = \hat{R}$ is determined separately using a time series analysis method and is held as a fixed value during the subsequent state and parameter estimation. As discussed more in Sec. IV.A, the measurement noise covariance matrix is a physically meaningful quantity that can be estimated accurately using a variety of techniques. With R held fixed, a Kalman filter is used to solve the equations of motion, as discussed in Sec. IV.B, and this process is coupled to a nonlinear optimizer that simultaneously adjusts $\hat{\theta}_1$ and \hat{Q} to minimize the negative log likelihood function, as discussed in Sec. IV.C. Figure 2 depicts this process as a block diagram. Sections IV.D and IV.E discuss accuracy of the parameter estimates and practical aspects of the approach, respectively.

The decision to fix R simplifies the estimation problem in three regards. First, and most important, this choice breaks any correlations between θ_1 and R . Parameter correlations have hindered previous attempts at FE, and relaxation techniques are often required to mitigate these correlations at the cost of decreased rates of overall convergence.

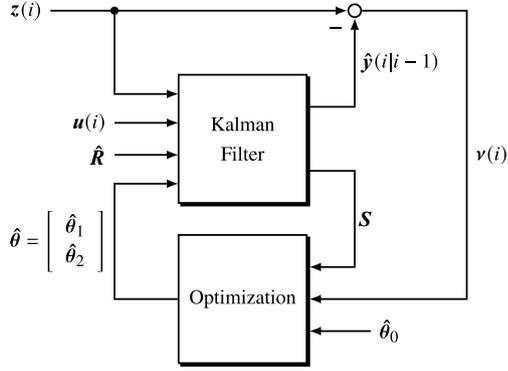


Fig. 2 Proposed filter-error method block diagram.

The impact of this choice for the proposed FE formulation is that the remaining unknowns (θ_1 and Q) are generally well-conditioned and should be estimated simultaneously. Second, this choice simplifies the parameter estimation problem by decreasing the number of unknowns to be estimated. This simplification in turn reduces the computational burden and lowers the uncertainties on estimated parameters. Third, the separate estimation assigns values to \hat{R} that are consistent with the model formulation, so that the implementation matches the theory. In general, if model parameters can be assigned appropriate values, then the overall parameter estimation is simplified and accuracy is improved for the remaining parameter estimates.

It is assumed in the proposed formulation that θ_1 and Q are unknown constants for the data analyzed. Additional information would be needed to identify time-varying process noise levels and model parameters. However, the assumption that the parameters are unknown constants is common, and most maneuvers used for aircraft system identification are relatively short in duration and involve only small perturbations about a reference flight condition. It is also assumed that R is constant. This choice was made in part because R reflects sensor measurement noise, which is not expected to change quickly in time, and in part to facilitate using a steady-state Kalman filter for computational efficiency. At the cost of longer computation times and additional information supplied by the analyst, time-varying measurement noise levels could also be incorporated into the estimation problem.

A. Measurement Noise Covariance

The random error observed in sensor data can often be modeled as additive white noise, which is an uncorrelated random sequence with zero mean and constant power over frequency. The measurement noise covariance matrix can be obtained in a variety of ways. Using manufacturer specifications on sensor noise levels or estimates from ground testing would result in measurement noise estimates that are too low to accurately represent flight test data.

In this work, noise levels are extracted from the measured flight test data and some knowledge of the instrumentation. The basic idea is to determine the noise floor using frequency-domain transforms of the data. For example, Fig. 3 shows a power spectral density (PSD) estimate of the pitch rate measurements for the X-56A airplane during a dynamic maneuver intended for system identification of the short period and first wing bending modes. The primary response of the vehicle occurred at frequencies below about 7 Hz. Low-pass filters attenuated the spectrum above frequencies of about 60 Hz. Assuming white measurement noise, an estimate of the noise floor for this measurement can therefore be obtained over about 7–60 Hz, where the spectrum is relatively flat. From this estimate, the noise variance on this measurement can be computed.

Another approach is to use a Fourier sine series decomposition, as discussed in Refs. [4,29–31]. The data are detrended, reflected about the origin to create odd symmetry without magnitude or slope discontinuity, and then projected onto sine functions using sine series coefficients. By discarding the coefficients outside a frequency band of interest where the noise floor is assumed, for example, outside the range 7–60 Hz in Fig. 3, the noise signal can then be reconstituted as a

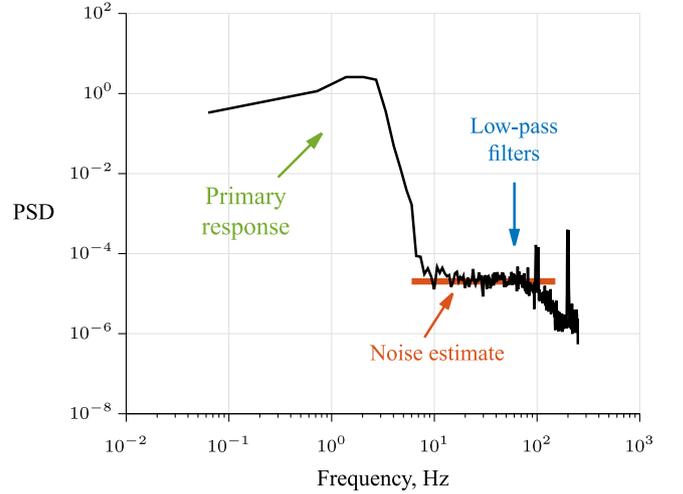


Fig. 3 Example measurement noise estimation from X-56A pitch-rate gyroscope data.

time series and its sample variance computed. Alternatively, the noise variance can be computed directly from the mean square value of the sine series coefficients in the frequency range 7–60 Hz. In doing so, \hat{R} is usually constructed as a diagonal matrix. This implies that measurement noise sequences for different sensors are mutually uncorrelated, which is a good assumption in practice.

Estimating the noise levels from measured data is expected to be the most accurate and convenient approach for determining \hat{R} . This obviates the need for sensor manufacturer specifications, which are often optimistic, or analytically propagating the statistics for derived measurements such as airspeed. It is not detrimental to the estimation process but is worth noting that other sources of error, such as vibrations of the structure or engine, may be incorporated into the noise estimate when determined this way from experimental data. Alternatively, because frequency-domain methods are used to estimate the noise levels, strong components (e.g., the peak near 100 Hz in Fig. 3) could be omitted when determining \hat{R} and regarded as a contribution to the process noise.

B. State Estimation

This section briefly summarizes the Kalman filter used to solve the equations of motion in Eq. (1). For more details on the Kalman filter, see Refs. [4,32,33] and others.

To simplify the estimation problem and the computations, a steady-state Kalman filter is used, which assumes that the plant model is LTI, the pair (A, C) is observable, and that Q and R are constants. These assumptions are sufficient in cases where the flight test maneuver involves only small perturbations about a reference flight condition and for relatively short periods of time, which is typical of most maneuvers used for aircraft system identification. Extensions to this situation could be applied, but they are beyond the scope of this work.

The first step in forming the Kalman filter is discretizing the system model in Eq. (1). The state dynamics and outputs are written using zero-order hold (ZOH) sampling as

$$x(i) = \Phi x(i-1) + \Gamma u(i-1) + \Lambda w(i-1) \quad (11a)$$

$$y(i) = Cx(i) + Du(i) + Hw(i) \quad (11b)$$

$$z(i) = y(i) + v(i), \text{ for } i = 1, 2, \dots, N \quad (11c)$$

where

$$\Phi = e^{A\Delta t} \quad (12a)$$

$$\Gamma = \left(\int_0^{\Delta t} e^{A\tau} d\tau \right) B \quad (12b)$$

$$\Lambda = \left(\int_0^{\Delta t} e^{A\tau} d\tau \right) \mathbf{G} \quad (12c)$$

are the discrete-time versions of the system matrices. The process noise spectral density matrix is discretized as the covariance matrix

$$\mathbf{Q} = \int_0^{\Delta t} e^{A\tau} \mathbf{Q}_c e^{A^T \tau} d\tau \quad (13)$$

The second step in forming the Kalman filter is to solve the algebraic Riccati equation (ARE)

$$\mathbf{Q} = \mathbf{P} - \Phi \mathbf{P} \Phi^T + \mathbf{L} \bar{\mathbf{R}}^{-1} \mathbf{L}^T \quad (14)$$

for the steady-state versions of the state error covariance matrix \mathbf{P} and the associated Kalman gain

$$\mathbf{K} = \mathbf{L} \bar{\mathbf{R}}^{-1} \quad (15)$$

where

$$\mathbf{L} = \Phi \mathbf{P} \mathbf{C}^T + \Lambda \mathbf{Q} \mathbf{H}^T \quad (16)$$

$$\bar{\mathbf{R}} = \mathbf{C} \mathbf{P} \mathbf{C}^T + \hat{\mathbf{R}} + \mathbf{H} \mathbf{Q} \mathbf{H}^T \quad (17)$$

Afterward, the steady-state innovation covariance matrix can be computed as

$$\mathbf{S} = \mathbf{C} \mathbf{P} \mathbf{C}^T + \hat{\mathbf{R}} \quad (18)$$

Having now determined each component part, the Kalman filter can then be used to solve the equations of motion using the two-step predictor/corrector procedure. The prediction step first extrapolates the state and output vectors to the next sample time using the current known information as

$$\hat{\mathbf{x}}(i|i-1) = \Phi \hat{\mathbf{x}}(i-1|i-1) + \Gamma \mathbf{u}(i-1) \quad (19a)$$

$$\hat{\mathbf{y}}(i|i-1) = \mathbf{C} \hat{\mathbf{x}}(i|i-1) + \mathbf{D} \mathbf{u}(i) \quad (19b)$$

The correction step then considers a measurement and updates the state and output estimates as

$$\hat{\mathbf{x}}(i|i) = \hat{\mathbf{x}}(i|i-1) + \mathbf{K} \boldsymbol{\nu}(i) \quad (20a)$$

$$\hat{\mathbf{y}}(i|i) = \mathbf{C} \hat{\mathbf{x}}(i|i) + \mathbf{D} \mathbf{u}(i) \quad (20b)$$

where the innovations were defined in Eq. (9). More information on this process can be found in standard textbooks, for example, Refs. [4,32,33].

This procedure yields a time history of the estimated model output based on measurements of \mathbf{u} and \mathbf{z} and estimates of $\boldsymbol{\theta}_1$, \mathbf{Q} , and \mathbf{R} . This model output estimate is compared with the measured output in the optimization step to update estimates of $\boldsymbol{\theta}_1$ and \mathbf{Q} , as described in the next section.

C. Optimization

The optimization process depicted in Fig. 2 determines estimates $\hat{\boldsymbol{\theta}}$ that minimize the cost function in Eq. (10). While the optimization is running, updated estimates $\hat{\boldsymbol{\theta}}$ are supplied to the Kalman filter, which then returns updates of $\boldsymbol{\nu}(i)$ and \mathbf{S} back to the optimizer.

The estimation problem is nonlinear and therefore requires an iterative solver. As a result, sufficiently good starting values of the estimates $\boldsymbol{\theta}_0$ must be provided for the first iteration. These can come from a previous EE or OE analysis, prior information, wind tunnel tests, analyst judgment, etc. Specifically for $\hat{\mathbf{Q}}_0$, starting values can be obtained from the residuals of an EE analysis or can usually be set to a small, diagonal, positive-definite matrix.

Any relevant optimization routine can be used to minimize the negative log likelihood cost function in Eq. (10). For this paper, the unconstrained simplex method [34], as implemented in the MATLAB[®] function `fminsearch.m`, was used. This approach samples the parameter space around the starting values and then follows a simple set of rules to minimize the cost function. Although convergence can be comparatively slow and local minima may be found, the approach is relatively robust to poor starting values and does not require derivatives to be calculated.

A gradient-based optimization, for example, the Gauss–Newton approach often used in OE [4,31], could also be used. In this case, the optimization block in Fig. 2 also includes inputs of the cost gradient and the (approximated) cost Hessian matrix from the Kalman filter. These terms are best obtained from numerical central finite differences rather than from analytical derivations due to the complex dependence of the cost function on the unknown parameters.

D. Accuracy of Parameter Estimates

The uncertainties in the parameter estimates are given by the Cramér–Rao bound matrix

$$\boldsymbol{\Sigma} = \text{cov}(\hat{\boldsymbol{\theta}}) \geq [\mathbf{M}^{-1}]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad (21)$$

which provides a theoretical lower limit of uncertainty. The standard errors for the parameter estimates are obtained from the square root of the diagonal elements in this matrix. In Eq. (21),

$$\mathbf{M} = -E \left[\frac{\partial^2 \ln \mathbb{L}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right] = \left[\frac{\partial^2 J}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad (22)$$

is the information matrix. For this paper, the information matrix was computed numerically using central finite differences after the optimization procedure converged. If instead a gradient-based optimization is used, the information matrix can be computed from the cost gradient and parameter sensitivity functions computed during the optimization iterations.

E. Practical Aspects

So far, discussion of the proposed FE implementation has primarily focused on the theory of the estimation. There are several practical aspects of the approach that warrant further discussion, which is the subject of this section.

The covariance matrix \mathbf{Q} is, by definition, a positive definite matrix. However, it is possible for its elements to be adjusted during the optimization process such that this is not true, which makes the Kalman filter unrealizable. One solution to this problem is to add constraints to the optimization to enforce the matrix to be positive definite. Another simpler approach is to check for positive definiteness and temporarily set any offending elements to small positive numbers for the remainder of that iteration. Another approach, which was used in this paper, is to estimate the square root of \mathbf{Q} (given by its Cholesky factorization) rather than \mathbf{Q} directly, similar to the square-root formulation of the Kalman filter [33]. This last approach proved to be a simple and straightforward solution to the practical problems.

The optimality of the Kalman filter performance, which depends on the correct selection of the dynamic model and noise statistics for the data, can be cross-checked in various ways [28,33]. For example, $\boldsymbol{\nu}$ should be approximately zero mean, normally distributed, and spectrally white. In addition, \mathbf{S} computed from Eq. (18) should approximately match the sample covariance matrix computed from the innovations. If this is not the case, then the filter is suboptimal, which could be due to error in the system model and noise covariance matrices.

The Kalman filter assumes that the measurement and process noise are uncorrelated white noise processes. If the process noise is not

[®]The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by NASA.

expected to be white, additional states can be appended to the state vector in the dynamic model to adjust the spectral coloring [16,33]. If the measurement and process noise are known to be significantly correlated, the Kalman filter can be generalized to account for the correlation $E[\mathbf{w}(i)\mathbf{v}(j)]$ as discussed in Ref. [33].

The process noise acting on the system dynamics sometimes comes from measurement noise on an input. An example of this occurs in data compatibility analysis, which is discussed later in Sec. VI. For these cases, the same techniques discussed in Sec. IV.A to estimate $\hat{\mathbf{R}}$ can be used to estimate $\hat{\mathbf{Q}}$. This estimate can either serve as a starting value, or \mathbf{Q} could be fixed at this value during the parameter estimation if the model structure is well known, to further reduce computations.

In some cases, particularly when there is only one process noise input and one output measurement, elements of \mathbf{Q} and \mathbf{G} and/or \mathbf{H} are confounded, and each term cannot be uniquely determined. When this happens, one parameter should be estimated and the others fixed at unity.

If there is prior knowledge on some of the unknown parameters, the cost function in Eq. (10) can be augmented as

$$J = \frac{1}{2} \sum_{i=1}^N \mathbf{v}^T(i) \mathbf{S}^{-1} \mathbf{v}(i) + \frac{N}{2} \ln |\mathbf{S}| + \frac{1}{2} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_p)^T \boldsymbol{\Sigma}_p^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_p)^T \quad (23)$$

where $\hat{\boldsymbol{\theta}}_p$ and $\boldsymbol{\Sigma}_p$ are the mean and covariance of the prior estimate, respectively [4]. No other changes to the method are necessary. This augmentation is useful, for example, when the data information content from one maneuver is insufficient or when combining results from multiple maneuvers.

V. Simulation Example

To demonstrate the method and make comparisons against a known solution, a simple simulation model using the roll mode approximation of an aircraft is examined. Section V.A considers a baseline case, Sec. V.B examines diagnostic metrics of the estimation, Sec. V.C looks at the sensitivity functions, and Sec. V.D considers errors in the measurement noise covariance matrix.

A. Baseline Case

The simplified roll mode approximation of an aircraft [3,4,35] is

$$\dot{p}(t) = L_p p(t) + L_{\delta_a} \delta_a(t) + w(t) \quad (24a)$$

where p is the roll rate in rad/s, δ_a is the aileron deflection in rad, and w is process noise in rad/s². The model parameters L_p and L_{δ_a} are stability and control derivatives, which are assumed to be constants for a given flight condition. The system equations are completed by

$$p_m(i) = p(i) + v(i) \quad (24b)$$

which is a discrete-time measurement of the roll rate from a rate gyroscope, including sensor measurement noise. The subscript m is used to distinguish the roll rate measurement from the model output for roll rate of the aircraft. This problem could be a first approximation to the roll dynamics of an aircraft flying through turbulence.

The system was simulated for 30 s at a sampling rate of 100 Hz. True values of the system parameters were $L_p = -2 \text{ s}^{-1}$ and $L_{\delta_a} = -10 \text{ s}^{-2}$. The process and measurement noise sequences were generated as white random sequences with $\mathbf{Q} = 0.2 \text{ rad}^2/\text{s}^4$ and $\mathbf{R} = 30 \times 10^{-6} \text{ rad}^2/\text{s}^2$, which corresponded to standard deviations of about 26 deg/s^2 and 0.31 deg/s , respectively. The initial conditions were specified with the covariance $\mathbf{P}_0 = 3.0 \times 10^{-6} \text{ rad}^2/\text{s}^2$, which corresponded to a standard deviation of 0.1 deg/s . During the maneuver, a multisine input [4] with 2 deg amplitude, 20 s duration, and frequency content between 0.1 and 1.0 Hz was applied to the aileron. The simulation and analysis were repeated 500 times using different realizations of the noise sequences. Figure 4 shows input

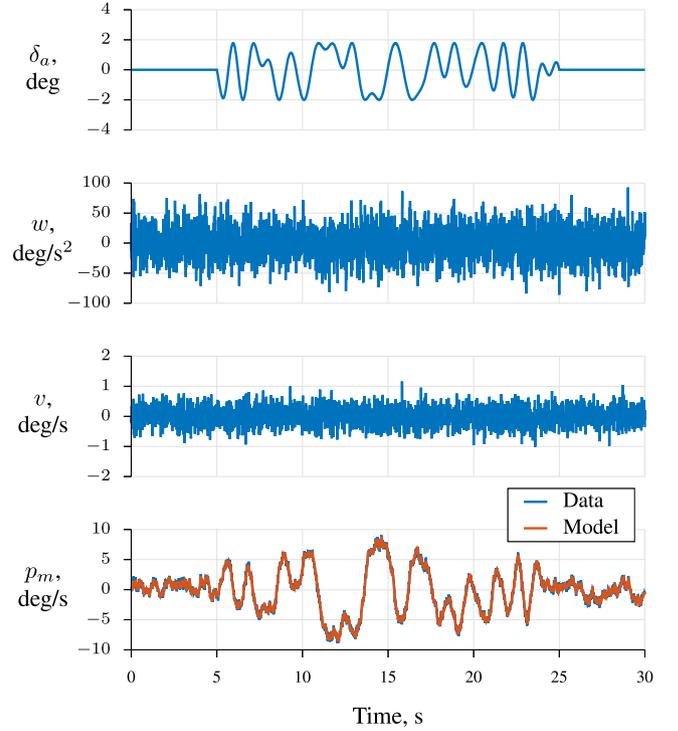


Fig. 4 Data and FE fit for one simulation run.

and output data for one simulation run, drawn in blue and converted from rad to deg. In terms of root-mean-square (RMS) values for the different roll moment terms, $L_p p$ was approximately equal to $L_{\delta_a} \delta_a$, and w was about twice as large.

For this problem,

$$\boldsymbol{\theta}_1 = \begin{bmatrix} L_p \\ L_{\delta_a} \end{bmatrix} \quad (25)$$

$$\boldsymbol{\theta}_2 = \mathbf{Q} \quad (26)$$

and $\boldsymbol{\theta}_3$ and $\boldsymbol{\theta}_4$ were both empty vectors, which made the total parameter vector to be estimated:

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} = \begin{bmatrix} L_p \\ L_{\delta_a} \\ \mathbf{Q} \end{bmatrix} \quad (27)$$

In this initial baseline case, the true measurement noise variance was supplied to the estimator, $\hat{\mathbf{R}} = \mathbf{R}$. Initial conditions were assumed to be zero in the analysis. Starting values for $\hat{\boldsymbol{\theta}}_0$ were obtained from an EE analysis using time-domain data. Over the 500 simulation runs, the optimization converged for every run and used 76 ± 5 iterations and 137 ± 8 function evaluations, which needed $0.55 \pm 0.04 \text{ s}$ of computation time on a standard laptop computer with an Apple® M2 Pro chip. The uncertainties with these numbers indicate one standard deviation of scatter. Figure 4 also includes the model fit to the measured roll rate data using FE for that run, drawn in red. The coefficient of determination for the FE fit to the measured data, R^2 , was 0.996, which is good and consistent with the signal-to-noise ratio for the roll rate data. As expected, the fitted model followed the trend of the measurements closely but avoided fitting the measurement noise.

Figure 5 summarizes the parameter estimates for the 500 simulation runs. True values are shown by the red, horizontal, dashed lines. Mean estimates for this baseline case are given by the blue circles, with the associated left error bounds showing two deviations of scatter in the estimates and right error bounds showing averaged two standard errors (95% confidence intervals) computed in the FE

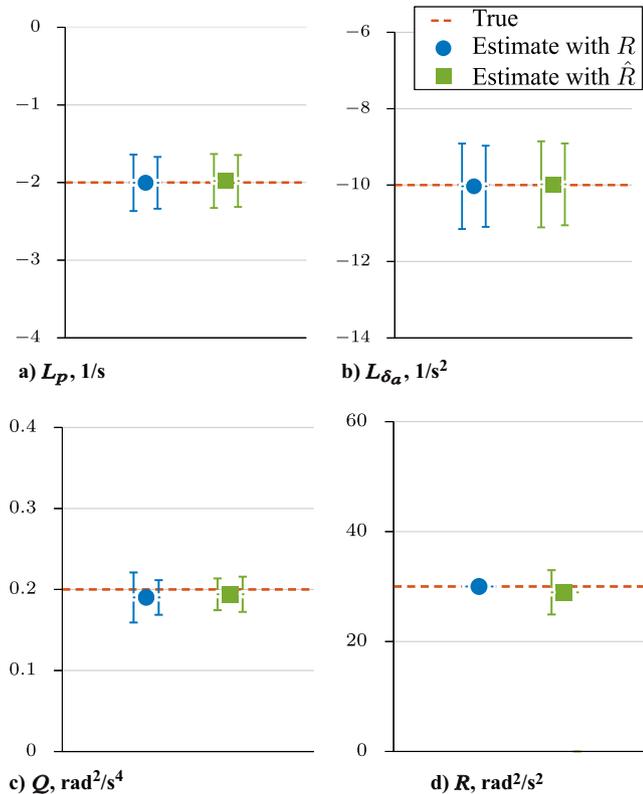


Fig. 5 Summary of parameter estimates for 500 simulation runs: dashed line = true value, marker = mean estimate, left bar = two standard deviations of scatter, and right bar = estimated two standard errors.

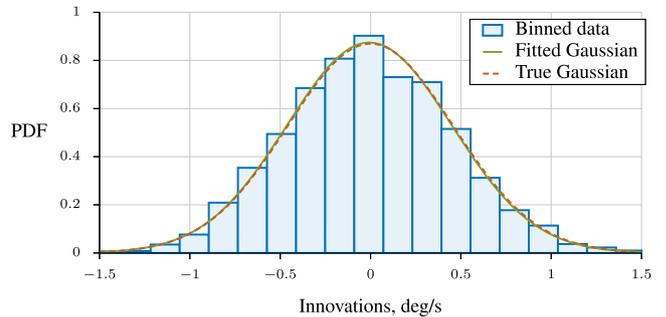
algorithm. The green squares and error bounds, which are discussed later in Sec. V.D, are the results for another set of 500 runs where an estimate of R was used in place of the true value. The baseline estimates were very close to the true values and were in statistical agreement. For the averaged results, biases were small and within 0.2, 0.3, and 5% for \hat{L}_p , $\hat{L}\delta_a$, and \hat{Q} , respectively. The scatter in the estimates (left bounds) approximately matched the estimated uncertainty (right bounds), which indicates that the Cramér–Rao bounds were accurate representations of the uncertainty. The averaged standard errors were 8.3, 5.3, and 5.6% of the averaged parameter values, which is relatively small. Compared to the stability and control derivative estimates, the results for \hat{Q} had a larger bias and a slight mismatch between the two error bars. This is expected to be due to the difficulty in estimating the process noise covariance from noisy data. However, these estimates can still be considered accurate by conventional standards.

B. Innovation Diagnostics

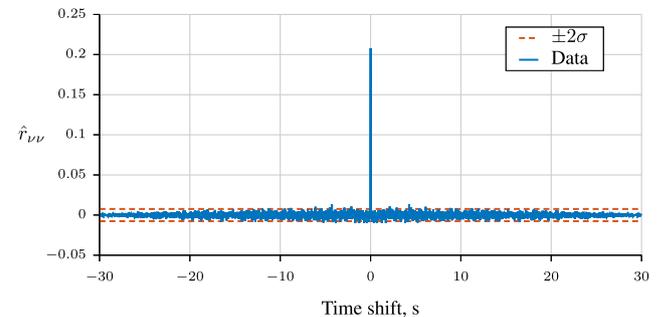
Performance metrics were applied to the innovations for the simulation run shown in Fig. 4. The mean of the innovation sequence was -0.003 deg/s, which was approximately zero and very small compared to the maximum roll rate measurement of about 10 deg/s. The theoretical innovation covariance matrix computed in Eq. (18) was 64.0×10^{-6} rad²/s², whereas the sample covariance of the innovations was 63.4×10^{-6} rad²/s². This difference is less than 1% and suggests that the Kalman filter was working optimally.

Figure 6a shows a probability density function (PDF) of the innovations. The data were partitioned into 20 bins, which are drawn as blue rectangles. A Gaussian function, drawn as a green solid line, was fitted to these binned data. A second Gaussian function, drawn as a red dashed line, shows the theoretical distribution based on the innovation covariance computed using Eq. (18). Only small differences in the mean and variance are evident, which suggests that the innovations were normally distributed with the mean value close to zero.

Figure 6b shows the sample autocorrelation function of the innovations



a) Innovation probability density function estimate



b) Innovation sample autocorrelation function

Fig. 6 Innovation diagnostic plots for one simulation run.

$$\hat{r}_{\nu\nu}(k) = \frac{1}{N} \sum_{i=1}^{N-k} \nu(i)\nu(i+k), \text{ for } k = 0, 1, \dots, N-1 \quad (28)$$

Two standard errors for the innovation autocorrelation estimate, computed from $\hat{\sigma} = \hat{r}_{\nu\nu}(0)/\sqrt{N}$, are shown by the red, horizontal, dashed lines. The autocorrelation plot is characteristic of white noise in that it has a large peak at the origin approximately equal to the variance, and it otherwise generally remains within two standard errors of zero. Only 67 out of the 6001 samples, or about 1%, fell outside the 95% confidence region, which is in statistical agreement with the estimated uncertainty.

In summary, an examination of the innovations supported the quality of the estimation results. The innovations had a small mean and a covariance that matched its theoretical value. This provided evidence that the Kalman filter was running optimally, which occurs when the model parameters and noise statistics are accurate. The estimated PDF and autocorrelation of the innovations indicated that these were normally distributed and spectrally white, as expected.

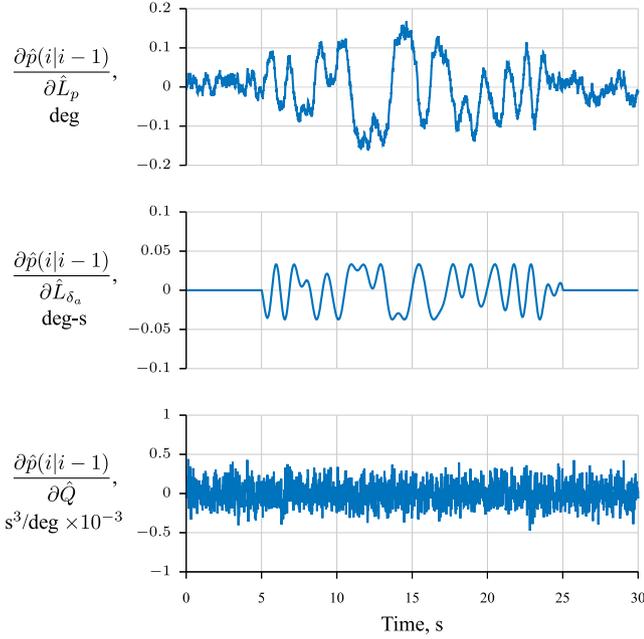
C. Output Sensitivity Functions

The output parameter sensitivity functions quantify the change in the model outputs due to a change in a single model parameter, with all the other model parameters held fixed. Examining these functions can often provide insight into parameter uncertainties and the performance of the optimization. Ideally, sensitivity functions are large in amplitude and distinct from one another to speed convergence.

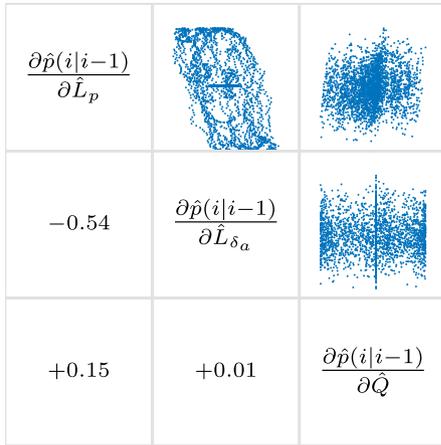
The three sensitivity functions for this example were

$$\frac{\partial \hat{p}(i|i-1)}{\partial \hat{L}_p} \quad \frac{\partial \hat{p}(i|i-1)}{\partial \hat{L}\delta_a} \quad \frac{\partial \hat{p}(i|i-1)}{\partial \hat{Q}}$$

Figure 7a shows time histories of these sensitivity functions for the run in Fig. 4, which were numerically computed using central finite differences. These time histories resemble the roll rate measurements, aileron deflections, and process noise inputs shown in Fig. 4, respectively. Visually, each time history in Fig. 7a appears distinct from the other two.



a) Time histories



b) Pairwise correlations

Fig. 7 Output sensitivity functions for one simulation run.

The degree to which the sensitivity functions are linearly distinct from each other can be quantified by the correlation coefficient. Figure 7b shows this correlation visually in a plot matrix. The diagonal labels denote each sensitivity function. The time histories are cross plotted in the upper triangle, and the correlation coefficient is displayed in the corresponding lower triangle. Overall, the pairwise correlations were relatively low in absolute value, and much less than the 0.9 threshold value recommended in Ref. [4] for obtaining accurate estimates. This is confirmed visually in how the cross plots resemble scatter rather than a straight line. A separate check using a variety of metrics [4] (not shown due to complexity) indicated that multi-collinearity was also not a factor for these data. As discussed in Sec. IV, specifying the measurement noise covariance breaks parameter correlations between the process noise covariance and the stability and control derivatives.

D. Measurement Noise Covariance Accuracy

Another factor investigated using this simulation example was the sensitivity of the results to the accuracy of the specified measurement noise covariance matrix. For the previous baseline results, the true measurement noise covariance was provided to the estimator. In this section, that assumption is removed.

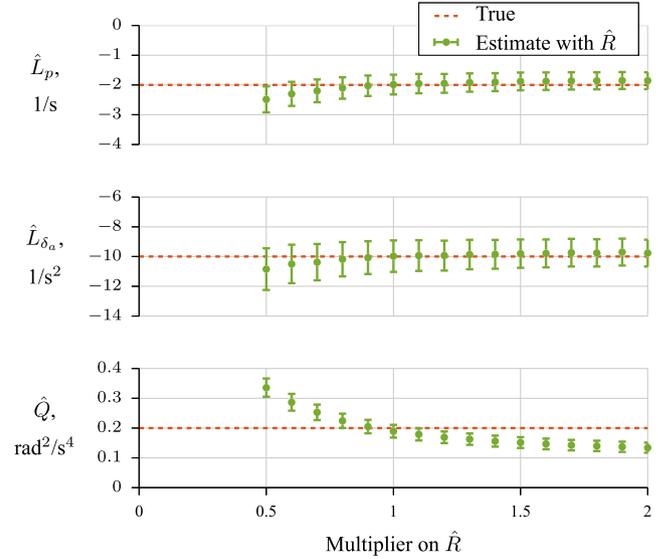


Fig. 8 Sensitivity of the roll mode parameter estimates to the specified measurement noise covariance.

Another set of 500 simulations were run, where this time the \hat{R} supplied to the estimator was obtained from the measurements, as discussed in Sec. IV.A. Upon examining the PSD, which was similar to Fig. 3, \hat{R} was estimated using sine series coefficients [4,29,30] from 10 to 50 Hz. In Fig. 5, the parameter estimation results from this section are summarized in green: again, the markers indicate the mean values, the left bars denote two standard deviations of scatter, and the right bars denote the averaged two standard errors computed by the FE algorithm. There is no right error bound associated with \hat{R} because an uncertainty was not developed for that estimate. On average, the bias in \hat{R} was small at 3.7%. The scatter in the results, which was about 7.2% of the mean estimate, was in statistical agreement with the true value. Estimating R instead of using the true value did not cause noticeable degradation in the estimates. The left bounds were approximately the same as the right bounds, which again suggested that the uncertainty estimates were accurate.

Figure 8 shows another sensitivity test in which \hat{R} was supplied with a multiplicative error. The red, horizontal, dashed lines indicate true values for the parameters. Each marker denotes the mean parameter estimate from a separate set of 500 simulation runs. The error bars denote averaged two standard errors from the FE algorithm. When the multiplier is unity, $\hat{R} = R$ and the results shown in Fig. 5 and drawn in blue are obtained. As the multiplier on \hat{R} increases, more of the random variation in the data is attributed to measurement noise, which gradually lowers the estimated process noise level and increases biases in the parameter estimates. As the multiplier on \hat{R} decreases, the opposite occurs, but happens at a higher rate. Based on this limited sensitivity study, it is best to specify the correct measurement noise covariance, but otherwise it is better to provide an overestimate rather than an underestimate for this simple case.

VI. Flight Test Example

In this section, the proposed FE approach is demonstrated with a data compatibility analysis, also known as a kinematic consistency analysis, using flight test data. In this problem, subsets of the measurements are compared through kinematic relationships to estimate and remove instrumentation errors. For example, in the absence of any errors, integrated velocity measurements should match position measurements. Because the inputs in this problem contain noise, process noise appears in the state equations and FE is the most appropriate choice for estimating the unknown parameters.

Flight data are from the X-56A airplane, which is a subscale aeroelastic demonstrator. Figure 9 is a photograph of the airplane in flight over NASA Armstrong Flight Research Center (AFRC). The maneuver considered includes an automated doublet with 5 s pulse



Fig. 9 X-56A in flight (credit: NASA / Jim Ross).

durations and 3 deg amplitude on the flight path angle command, starting from an initial flight condition of straight and level flight at 4060 ft altitude, 125 ft/s airspeed, and 2.4 deg angle of attack. Longitudinal responses during this maneuver were small, and lateral-directional responses were negligible.

The equations of motion for this data compatibility analysis are represented in the form of Eq. (1), where

$$\mathbf{u} = [a_{z_m} \quad q_m \quad 1]^T \quad (29a)$$

$$\mathbf{x} = [\Delta\alpha \quad \Delta\theta]^T \quad (29b)$$

$$\mathbf{z} = [\alpha_m \quad \theta_m]^T \quad (29c)$$

$$\mathbf{w} = [v_{a_z} \quad v_q]^T \quad (29d)$$

$$\mathbf{v} = [v_\alpha \quad v_\theta]^T \quad (29e)$$

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (29f)$$

$$\mathbf{B} = \begin{bmatrix} \frac{g}{V_0} & 1 & -\frac{g}{V_0}(b_{a_z} + a_{z_0}) - b_q \\ 0 & 1 & -b_q \end{bmatrix} \quad (29g)$$

$$\mathbf{G} = \begin{bmatrix} -\frac{g}{V_0} & -1 \\ 0 & -1 \end{bmatrix} \quad (29h)$$

$$\mathbf{C} = \begin{bmatrix} 1 + \lambda_\alpha & 0 \\ 0 & 1 + \lambda_\theta \end{bmatrix} \quad (29i)$$

$$\mathbf{D} = \begin{bmatrix} 0 & -(1 + \lambda_\alpha)\frac{x_\alpha}{V_0} & b_\alpha + (1 + \lambda_\alpha)\left(\alpha_0 + \frac{x_\alpha}{V_0}b_q\right) \\ 0 & 0 & b_\theta + (1 + \lambda_\theta)\theta_0 \end{bmatrix} \quad (29j)$$

$$\mathbf{H} = \begin{bmatrix} 0 & (1 + \lambda_\alpha)\frac{x_\alpha}{V_0} \\ 0 & 0 \end{bmatrix} \quad (29k)$$

The Appendix includes a derivation of these equations and definitions for the included terms. The inputs to this system are the vertical accelerometer measurements, taken near the center of mass, and the pitch rate gyroscope measurements, taken near the nose of the vehicle. The input vector includes a 1 at the end to account for constant terms in the equations. The states are perturbations in the angle of attack at the center of mass and the pitch angle. The outputs are angle of attack at the noseboom and pitch angle. Noise on the accelerometer and pitch rate gyroscope measurements comprise \mathbf{Q} , whereas noise on the angle-of-attack vane and the pitch-angle measurements comprise \mathbf{R} . The unknowns to be estimated using FE are

$$\boldsymbol{\theta} = [b_{a_z} \quad b_q \quad \lambda_\alpha \quad \lambda_\theta \quad b_\alpha \quad b_\theta \quad Q_{a_z} \quad Q_q]^T \quad (30)$$

which include the bias and scale factor error parameters, as well as the diagonal elements of the process noise covariance matrix. Off-diagonal elements of \mathbf{Q} were not included because the majority of the process noise was expected to be due to measurement noise on the two sensors, which is generally uncorrelated.

Figure 10 shows flight test data in blue for one maneuver. Diagonal elements of $\hat{\mathbf{R}}$ were estimated from PSDs over 7–60 Hz, as in Fig. 3, with standard deviations of 0.070 and 0.020 deg for the angle-of-attack vane and pitch-angle measurements, respectively. Initial conditions for the states were set to zero, as well as all starting values $\hat{\boldsymbol{\theta}}_0$. Values of steady-state terms in Eq. (29), such as α_0 , were estimated from the first 2 s of recorded data. The optimization converged in 719 iterations and 1026 function evaluations, which took about 17 s on the same standard laptop computer mentioned previously. Figure 10 also shows FE fits to the output data as the red lines. These fits followed the data trends closely but avoided the noise. Values of R^2 were 0.963 and 1.000 for the angle of attack and pitch angle, respectively. The residuals on the angle of attack and pitch angle had standard deviations of 0.068 and 0.014 deg, which were within 3 and 30% of the estimated values for $\hat{\mathbf{R}}$, respectively.

Table 1 presents the estimated parameters and uncertainties. All estimates were small, which was expected because high-quality instrumentation was used and calibrations were done carefully. The process noise variance estimates, which are given in Table 1 as standard deviations, were an order of magnitude larger than the values of 0.0095 g and 0.13 deg/s estimated from PSDs over 7–60 Hz. The process noise estimates from the FE algorithm were larger due to

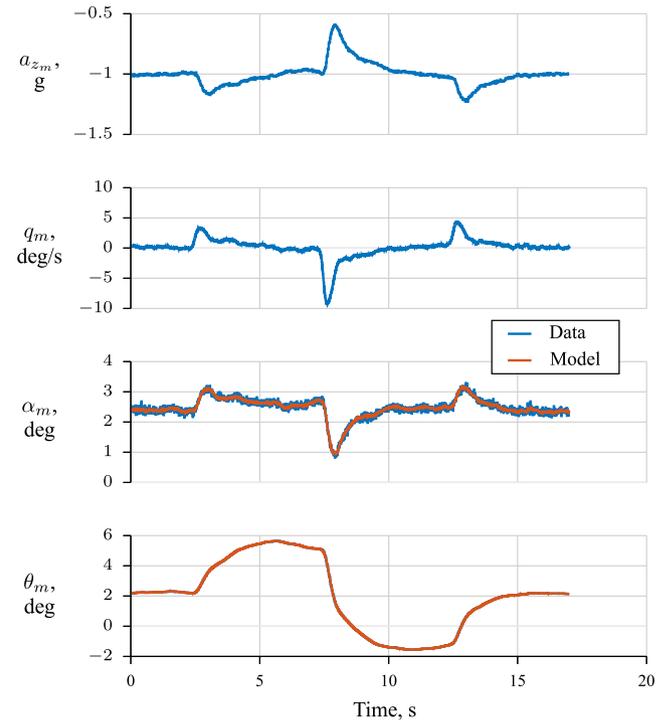


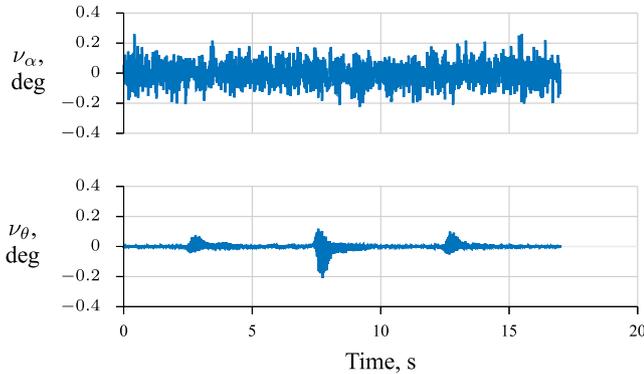
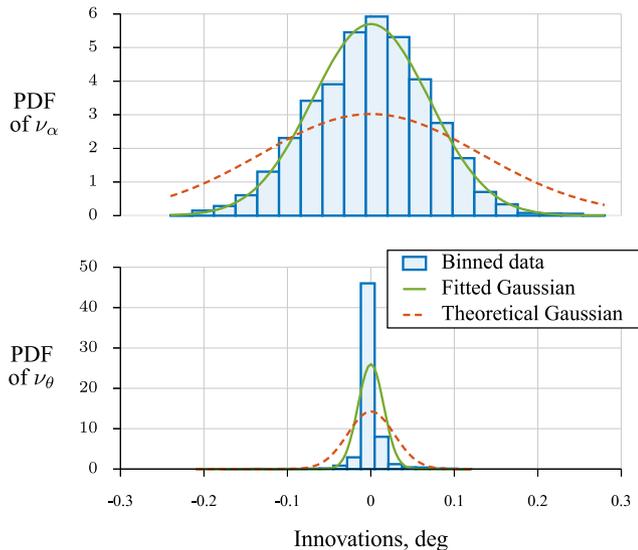
Fig. 10 Flight test data and FE fit.

Table 1 Estimated parameters and uncertainties for the X-56A data compatibility example

Parameter	Estimate	Std. Error	% Std. Error	Unit
b_{a_z}	-0.0031	0.0013	43.4	g
b_q	+0.0964	0.0181	18.8	deg/s
λ_α	-0.0886	0.0212	24.0	—
λ_θ	-0.0210	0.0085	40.5	—
b_α	+0.2121	0.0598	28.2	deg
b_θ	-0.0329	0.0232	70.4	deg
$\sqrt{Q_{a_z}}$	+0.1230	0.0120	9.7	g
$\sqrt{Q_q}$	+1.6690	0.0594	3.6	deg/s

model structure error outside of the 7–60 Hz frequency range that was included in the process noise variance estimates. For this problem, sources of model structure error include linearization, unmodeled structural mode responses, neglected covariances between the two inputs or two outputs, errors in correcting the angle of attack and vertical accelerometer measurements to the center of mass, and a 0.050 s time delay in the pitch-angle measurement.

Figure 11 shows diagnostic plots for these results. Figure 11a shows time histories for the angle-of-attack and pitch-angle innovations, whereas Fig. 11b shows these data as histograms. For the histograms, the blue boxes show 20 bins of data, the green lines are Gaussians fitted to these binned data, and the red dashed lines are Gaussians based on the innovation covariance matrix provided by the Kalman filter. Although the angle-of-attack innovations resembled

**a) Innovation time histories****b) Innovation probability density function estimates****Fig. 11** Innovation diagnostic plots: X-56A example.

white noise, the pitch-angle innovations contained colored noise that occurred about 2, 7, and 12 s into the data record when the gamma command changed value. This colored noise was from the excitation of the first bending mode of the aircraft structure, which was at about 3 Hz, and was recorded by the high-rate gyroscopes near the aircraft nose. The structural responses were below the frequency range used to estimate the measurement noise levels and were therefore regarded as process noise. Because the theory does not fit this situation with colored noise, the innovation covariance estimate is impacted, as can be seen from Eqs. (14–18), which results in a slight degradation in the performance of the Kalman filter.

In summary, this section applied the proposed FE method to flight test data from the X-56A subscale aeroelastic demonstrator for a data compatibility analysis. As in any application to a physical system, there was error in the model structure used for the parameter estimation, which was detected by the innovation diagnostics. However, the approach was able to obtain reasonable estimates of all parameters and produced excellent fits to the measured data.

VII. Conclusions

This paper proposes a practical algorithm for implementing aircraft parameter estimation with the filter-error method, which considers both process noise and measurement noise in the dynamic system model. The key component of the algorithm is that by separately estimating the measurement noise covariance matrix, correlations between the unknown model parameters and the unknown process noise covariance matrix elements are broken, and the estimation can proceed in a straightforward manner using simple nonlinear optimization techniques.

The proposed formulation was demonstrated using a simulation example and a flight test example using data from the X-56A airplane. The findings of this paper can be summarized as follows:

- 1) The proposed method can produce accurate estimates of model parameters and uncertainties, assuming that both measurement noise and process noise are present.
- 2) Diagnostics on the innovations are helpful in understanding results.
- 3) Simple nonlinear optimization techniques, without relaxation steps, can be used because the model parameters and the process noise covariance matrix are decorrelated.
- 4) There is some robustness to error in the measurement noise covariance matrix estimate.
- 5) The proposed method is practical, has relatively low computational costs, maintains physical insight for the model parameters, and avoids convergence problems in the parameter estimation.

The present work considered LTI systems with constant process and measurement noise statistics and time-domain data. The analysis is relegated to batch postflight analysis due to the nonlinear estimation problem requiring iterative solutions. Despite those limitations, the approach shows promise for estimating stability and control derivatives for flight testing in turbulence or with model structure errors, as well as estimating instrumentation error parameters in a kinematic consistency analysis and other applications.

Acknowledgments

This research was supported by the NASA Advanced Air Transport Technology Project. The efforts of the X-56A team at NASA Armstrong Flight Research Center are gratefully acknowledged.

Appendix: Simplified Data Compatibility Equations

This Appendix includes a derivation of the simplified equations used for the data compatibility analysis in Sec. VI. For more information on the data compatibility analysis, see Refs. [4, 11, 13].

The nonlinear differential equations for the translational and rotational kinematics of a rigid-body aircraft are

$$\dot{u} = rv - qw - g \sin \theta + ga_x \quad (\text{A1a})$$

$$\dot{v} = pw - ru + g \sin \phi \cos \theta + ga_y \quad (\text{A1b})$$

$$\dot{w} = qu - pv + g \cos \phi \cos \theta + ga_z \quad (\text{A1c})$$

$$\dot{\phi} = p + \tan \theta (q \sin \phi + r \cos \phi) \quad (\text{A1d})$$

$$\dot{\theta} = q \cos \phi - r \sin \phi \quad (\text{A1e})$$

$$\dot{\psi} = \frac{q \sin \phi + r \cos \phi}{\cos \theta} \quad (\text{A1f})$$

where u , v , and w are the body-axis components of translational velocity; p , q , and r are the angular rates; ϕ , θ , and ψ are the bank, pitch, and heading Euler angles; a_x , a_y , and a_z are linear accelerometer measurements at the center of mass and aligned with the body axes; and g is the acceleration due to gravity.

The system inputs for this problem are the accelerometer and angular rate measurements, whereas the states are the translational velocity components and the Euler angles. The outputs are the airspeed V , angle of attack α , and flank angle μ (which is related to the sideslip angle β [36]),

$$V = \sqrt{(u - y_v r + z_v q)^2 + (v + x_v r - z_v p)^2 + (w - x_v q + y_v p)^2} \quad (\text{A2a})$$

$$\alpha = \arctan\left(\frac{w - x_a q + y_a p}{u - y_a r + z_a q}\right) \quad (\text{A2b})$$

$$\mu = \arctan\left(\frac{v + x_\mu r - z_\mu p}{u - y_\mu r + z_\mu q}\right) \quad (\text{A2c})$$

as well as the bank and pitch Euler angles. The x , y , and z terms represent the body-axis positions of the three sensors relative to the center of mass.

For pure longitudinal flight, the variables v , p , r , ψ , and μ are ignored. Over short durations of time, the forward speed is assumed to be constant, $u \simeq u_0$, where the subscript "0" denotes a steady-state value. In straight and level flight, $\phi \simeq 0$. Linearizing under these conditions and assuming a small pitch angle and a small vertical offset of the air data vane result in the state-space model

$$\begin{bmatrix} \dot{w}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta w(t) \\ \Delta \theta(t) \end{bmatrix} + \begin{bmatrix} g & u_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta a_z(t) \\ q(t) \end{bmatrix} \quad (\text{A3a})$$

$$\begin{bmatrix} \Delta \alpha(t) \\ \Delta \theta(t) \end{bmatrix} = \begin{bmatrix} 1/u_0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta w(t) \\ \Delta \theta(t) \end{bmatrix} + \begin{bmatrix} 0 & -x_a/u_0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta a_z(t) \\ q(t) \end{bmatrix} \quad (\text{A3b})$$

where Δ indicates a perturbation variable. By making the substitution $\Delta \alpha \simeq \Delta w/u_0$ and further assuming flight at low angles of attack such that $u \simeq V$, the model can be rewritten in terms of true airspeed and angle of attack as

$$\begin{bmatrix} \dot{\alpha}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta \alpha(t) \\ \Delta \theta(t) \end{bmatrix} + \begin{bmatrix} g/V_0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta a_z(t) \\ q(t) \end{bmatrix} \quad (\text{A4a})$$

$$\begin{bmatrix} \Delta \alpha(t) \\ \Delta \theta(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta \alpha(t) \\ \Delta \theta(t) \end{bmatrix} + \begin{bmatrix} 0 & -x_a/V_0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta a_z(t) \\ q(t) \end{bmatrix} \quad (\text{A4b})$$

Next, each measurement instrumentation error is generically modeled as

$$y_m = (1 + \lambda_y)y + b_y + v_y \quad (\text{A5a})$$

$$= (1 + \lambda_y)(\Delta y + y_0) + b_y + v_y \quad (\text{A5b})$$

where y_m is the measured value, y is the true value, λ_y is a scale factor error, b_y is a bias error, and v_y is zero-mean white measurement noise. Scale factors for accelerometer and gyroscope measurements are typically small and are neglected [4]. In going from Eq. (A5a) and (A5b), y is decomposed into a perturbation value Δy and a steady-state value y_0 , which could be zero.

Solving Eq. (A5b) for Δy , substituting that expression for each measurement in Eq. (A4), and collecting terms result in the simplified data compatibility equations used for parameter estimation.

References

- [1] Iliff, K., and Maine, R., "Uses of Parameter Estimation in Flight Test," *Journal of Aircraft*, Vol. 20, No. 12, Dec. 1983, pp. 1043–1049. <https://doi.org/10.2514/3.48210>
- [2] Jategaonkar, R., "Flight Vehicle System Identification-Engineering Utility," *Journal of Aircraft*, Vol. 42, No. 1, 2005, p. 11. <https://doi.org/10.2514/1.15766>
- [3] Tischler, M., and Rempfle, R., *Aircraft and Rotorcraft System Identification: Engineering Methods with Flight Test Examples*, 2nd ed., AIAA, Reston, VA, 2012.
- [4] Morelli, E., and Klein, V., *Aircraft System Identification: Theory and Practice*, 2nd ed., Sunflyte Enterprises, Williamsburg, VA, 2016.
- [5] Morelli, E., and Klein, V., "Application of System Identification to Aircraft at NASA Langley Research Center," *Journal of Aircraft*, Vol. 42, No. 1, 2005, pp. 12–25. <https://doi.org/10.2514/1.3648>
- [6] Grauer, J., and Morelli, E., "Introduction to the Advances in Aircraft System Identification from Flight Test Data Virtual Collection," *Journal of Aircraft*, Vol. 60, No. 5, 2023, pp. 1329–1330. <https://doi.org/10.2514/1.C037583>
- [7] Morelli, E. A., and Grauer, J. A., "Advances in Aircraft System Identification at NASA Langley Research Center," *Journal of Aircraft*, Vol. 60, No. 5, 2023, pp. 1354–1370. <https://doi.org/10.2514/1.C037274>
- [8] Morelli, E., and Cunningham, K., "Aircraft Dynamic Modeling in Turbulence," *Atmospheric Flight Mechanics Conference*, AIAA Paper 2012-4650, Aug. 2012. <https://doi.org/10.2514/6.2012-4650>
- [9] Martos, B., and Morelli, E., "Using Indirect Turbulence Measurements for Real-Time Parameter Estimation in Turbulent Air," *Atmospheric Flight Mechanics Conference*, AIAA Paper 2012-4651, Aug. 2012. <https://doi.org/10.2514/6.2012-4651>
- [10] Grauer, J. A., "Aerodynamic Parameter Estimation Using Reconstructed Turbulence Measurements," *Journal of Aircraft*, Vol. 58, No. 5, 2021, pp. 1022–1033. <https://doi.org/10.2514/1.C035933>
- [11] Klein, V., and Schiess, J., "Compatibility Check of Measured Aircraft Responses Using Kinematic Equations and Extended Kalman Filter," NASA TN D-4356, Aug. 1977.
- [12] Bach, R., "State Estimation Applications in Aircraft Flight-Data Analysis: A User's Manual for SMACK," NASA RP-1252, March 1991.
- [13] Klein, V., and Morgan, D., "Estimation of Bias Errors in Measured Airplane Responses Using Maximum Likelihood Method," NASA TM-89059, Jan. 1987.
- [14] Maine, R., and Iliff, K., "Identification of Dynamics Systems: Theory and Formulation," NASA RP-1138, Feb. 1985.
- [15] Cox, A. B., and Bryson, A. E., "Identification by a Combined Smoothing Nonlinear Programming Algorithm," *Automatica*, Vol. 16, No. 6, 1980, pp. 689–694. [https://doi.org/10.1016/0005-1098\(80\)90010-2](https://doi.org/10.1016/0005-1098(80)90010-2)
- [16] Jategaonkar, R., *Flight Vehicle System Identification: A Time-Domain Methodology*, 2nd ed., AIAA, Reston, VA, 2015.
- [17] Morelli, E., and Grauer, J., "Practical Aspects of Frequency-Domain Approaches for Aircraft System Identification," *Journal of Aircraft*, Vol. 57, No. 2, 2020, pp. 268–291. <https://doi.org/10.2514/1.C035599>
- [18] Maine, R., and Iliff, K., "Formulation and Implementation of a Practical Algorithm for Parameter Estimation with Process and Measurement Noise," *SIAM Journal on Applied Mathematics*, Vol. 41, No. 3, 1981, pp. 558–579. <https://doi.org/10.1137/0141045>
- [19] Schultz, G., "Maximum Likelihood Identification Using Kalman Filter—Least Squares Estimation: A Comparison for the Estimation of Stability Derivatives Considering Gust Disturbances," European Space Agency Technical Translation ESA TT-258, 1976.

- [20] Iliff, K., "Identification and Stochastic Control with Application to Flight Control in Turbulence," Univ. of California, Los Angeles, May 1973.
- [21] Maine, R., and Iliff, K., "A FORTRAN Program for Determining Aircraft Stability and Control Derivatives from Flight Data," NASA TN D-7831, April 1975.
- [22] Maine, R., "Programmer's Manual for MMLE3, a General FORTRAN Program for Maximum Likelihood Parameter Estimation," NASA TP-1690, June 1981.
- [23] Murray, J., and Maine, R., "pEst Version 2.1 User's Manual," NASA TM-88280, Sept. 1987.
- [24] Jategaonkar, R. V., and Plaetschke, E., "Algorithms for Aircraft Parameter Estimation Accounting for Process and Measurement Noise," *Journal of Aircraft*, Vol. 26, No. 4, 1989, pp. 360–372. <https://doi.org/10.2514/3.45769>
- [25] Jategaonkar, R., and Plaetschke, E., "Identification of Moderately Nonlinear Flight Mechanics Systems with Additive Process and Measurement Noise," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 2, 1990, pp. 277–285. <https://doi.org/10.2514/3.20547>
- [26] Jategaonkar, R., and Thielecke, F., "ESTIMA—An Integrated Software Tool for Nonlinear Parameter Estimation," *Aerospace Science and Technology*, Vol. 6, No. 8, 2002, pp. 565–578. [https://doi.org/10.1016/S1270-9638\(01\)01143-9](https://doi.org/10.1016/S1270-9638(01)01143-9)
- [27] Chowdhary, G., and Jategaonkar, R., "Aerodynamic Parameter Estimation from Flight Data Applying Extended and Unscented Kalman Filter," *Aerospace Science and Technology*, Vol. 4, No. 2, 2010, pp. 106–117. <https://doi.org/10.1016/j.ast.2009.10.003>
- [28] Mehra, R., "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, Vol. 15, No. 2, 1970, pp. 175–184. <https://doi.org/10.1109/TAC.1970.1099422>
- [29] Laczos, C., *Applied Analysis*, Prentice–Hall, Engelwood Cliffs, NJ, 1956, Chap. 11.
- [30] Morelli, E. A., "Estimating Noise Characteristics from Flight Test Data Using Optimal Fourier Smoothing," *Journal of Aircraft*, Vol. 32, No. 4, 1995, pp. 689–695. <https://doi.org/10.2514/3.46778>
- [31] Morelli, E., "System Identification Programs for AirCraft (SIDPAC)," Ver. 4.1, *NASA Software Catalog*, <http://software.nasa.gov> [retrieved Aug. 2024].
- [32] Gelb, A., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1973, Chap. 4.
- [33] Simon, D., *Optimal State Estimation*, Wiley, Hoboken, NJ, 2006.
- [34] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W., *Numerical Recipes: The Art of Scientific Computing*, Cambridge Univ. Press, Cambridge, England, U.K., 1986, Chap. 10.
- [35] McRuer, D., Ashkenas, I., and Graham, D., *Aircraft Dynamics and Automatic Control*, Preton Univ. Press, Princeton, NJ, 1973, Chap. 5.
- [36] Grauer, J., "Position Corrections for Airspeed and Flow Angle Measurements on Fixed-Wing Aircraft," NASA TM-2017-219795, Nov. 2017.