Structural Identification for Mobile Sensing with Missing Observations

Thomas J. Matarazzo, S.M.ASCE[1]; and Shamim N. Pakzad, A.M.ASCE[2]

Abstract: There are many occasions in structural health monitoring (SHM) on which collected data sets contain missing observations. Such instances may occur as a result of failed communications or packet losses in a wireless sensor network or as a result of sensing and sampling methods—for example, mobile sensing. By implementing modified expectation and maximization steps, structural identification using expectation maximization (STRIDE) is capable of processing data in these circumstances and is the first modal identification technique to formally accept data with missing observations. This paper presents the STRIDE algorithm, a statistical perspective of missing data, and new STRIDE equations that account for missing observations. Expectation step (E-step) equations are given explicitly for both partially observed time steps and those not fully observed. The maximization step (M-step) provides state-space parameter updates in terms of available observations and missing-data state-variable statistics. This paper also discusses the performance and convergence behavior of STRIDE with missing data. Finally, two applications are presented to exemplify common use in network reliability and mobile sensing, both using data collected at the Golden Gate Bridge. This paper demonstrates that sensor network data containing a significant amount of missing observations can be used to achieve a comprehensive modal identification. A successful real-world identification with simulated mobile sensors quantifies the preservation of spatial information, establishing the benefits of this type of network and emphasizing a line of inquiry for future SHM implementations. **DOI: 10.1061/(ASCE)EM.1943-7889.0001046.**

Author keywords: Mobile sensing; Missing data; Modal identification; Structural health monitoring; Wireless sensor network.

Introduction

Mobile-sensor networks are an emerging topic of interest in structural health monitoring (SHM). Their purpose is to remedy the shortcomings in fixed-sensor array setups. In mobile-sensor networks, a few sensors can be used to collect data containing dense spatial information (Cerda et al. 2012; Fabien et al. 2009; Singhvi et al. 2005; Unnikrishnan and Vetterli 2012). Additionally, in real-world implementations, sampling locations are less restricted; there are an immense number of potential mobile-sensor paths but a limited number of feasible areas for fixed-sensor placement. Consequently, mobile-sensor networks are easier to implement and more cost-effective compared with dense fixed-sensor arrays. The ultimate flaw of fixed sensors is that they provide limited spatial information, which can be addressed by mobile sensors.

The current state of mobile-sensor network applications in SHM is developing; research has been diverse although limited, as discussed in Matarazzo and Pakzad (2013, 2014). Zhu et al. (2010, 2012) developed flexure-based mobile-sensor nodes (FMSNs) representing a moving SHM sensor procedure that collects data at fixed nodes. However, these sensors do not record data while in motion. Sibley et al. (2002) and Dantu et al. (2005) designed a small, inexpensive robot platform, Robomote, permitting mobile non-SHM coverage of large-scale sensor networks. The challenges summarized in this work are not specific to SHM applications. Partial system identification studies include a moving vehicle to investigate the identifiability of frequencies for a single bridge span with frequency domain techniques [Lin and Yang (2005) and Cerda et al. (2012)].

This paper focuses on mobile-sensor data (i.e., data from sensors simultaneously recording in time while moving in space) for comprehensive system identification of real structural systems. Specifically, it addresses an inherent attribute of mobile-sensor data, the anticipated effects of missing observations in time and space.

Another application of processing techniques for data with missing observations is in wireless sensor networks (WSNs), which in SHM have typically been designed to minimize packet losses and budget their energy requirements. Nagayama et al. (2007) and Pakzad et al. (2008) showed that, although packet loss can be successfully prevented in a WSN for a power and communication cost, it may be possible to allow some missing packets by relaxing design constraints and still achieve accurate estimates of the data's desired features with an acceptable error. This might reduce WSN overhead communication and power costs by avoiding data retransmission and the computational efforts necessary to ensure perfect data delivery (Lynch 2007; Pakzad 2010; Xu et al. 2004). Lastly, despite developments in sensor network design, environmental effects may always limit reliable data collection and transmission (Lynch and Loh 2006), leaving missing data as an unavoidable WSN challenge (Zhao and Ramesh 2003). In the current study, packet losses in a sensor network were modeled as missing observations in a time series so that the presented missing-data equations enabled STRIDE to process data from a sensor network with a relatively unreliable communication protocol.

The capabilities of common output-only (or operational or stochastic) system identification (SID) algorithms, such as ERA-OKID-OO (Chang and Pakzad 2013), ERA-NEXT (James III et al. 1993), ERA-NEXT-AVG (Chang and Pakzad 2012), AR

[1]Dept. of Civil and Environmental Engineering, Lehigh Univ., 117 ATLSS Dr., Bethlehem, PA 18015 (corresponding author). E-mail: thomasjmatarazzo@gmail.com

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Lehigh Univ., 117 ATLSS Dr., Bethlehem, PA 18015. E-mail: pakzad@lehigh.edu

Note. This manuscript was submitted on July 3, 2014; approved on October 12, 2015No Epub Date. Discussion period open until 0, 0; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Engineering Mechanics***, © ASCE, ISSN 0733-9399.

(Andersen 1997), SSI (Peeters and De Roeck 1999), N4SID (Van Overschee and De Moor 1992), and others, in regard to data containing missing observations are undocumented. Furthermore, it is assumed in these cases that it is the responsibility of the user to remedy the missing data problem *prior* to SID. For example, "single" or "multiple" imputation is a statistical missing-data strategy that replaces unmeasured observations with computed values (which vary among methods), thus *completing* the data. Rubin (1987) detailed these techniques, which develop data sets compatible with familiar full-data analyses (SID algorithms in the context of this paper); however, Little and Rubin (2002) demonstrated that imputations can introduce biased inferences, systematically underestimate standard errors, and require assumptions about the predictive distribution of the missing values. Such issues become increasingly problematic for large amounts of missing data.

This paper presents a formulation for structural identification using expectation maximization (STRIDE) (Matarazzo and Pakzad 2016). STRIDE permits the use of data sets containing missing observations; therefore, in addition to its documented advantages over other SID algorithms as discussed in Matarazzo and Pakzad (2016), it is the first modal identification method to formally accept data with missing observations. With this feature, STRIDE can process data from mobile-sensor networks or data with missing packets (e.g., acquired from sensor networks with less reliable communication).

The following are the main objectives and contributions of this paper:

- To present new expectation (E-step) and maximization (M-step) formulations for STRIDE with missing data, joining approaches from Shumway and Stoffer (2011), Digalakis et al. (1993), and Sinopoli et al. (2004);
- To establish STRIDE convergence behavior and performance with various missing-data patterns and magnitudes;
- To provide two missing-data examples based on real-world monitoring of a structural system; and
- To quantitatively demonstrate that, given the same number of sensors, a mobile-sensor network produces superior spatial information when compared with a fixed-sensor network.

The remainder of this paper is organized as follows. First, a discussion of mobile sensing as a missing-data problem with specific time series approaches. Second, an overview of the STRIDE procedure presented in Matarazzo and Pakzad (2016). Third, required STRIDE missing-data equations are presented which affect the Kalman filter in the E-step and the state statistics in the M-step; the E-step considers both partially observed time steps and fully unobserved time steps. Fourth, an investigation on the performance and convergence behavior of STRIDE in the case of missing observations. Fifth, this methodology is applied to missing packets and mobile sensing. Finally, concluding remarks with a summary of results and a discussion of corresponding SHM expansions.

Mobile Sensing and Missing Data

Mobile Sensing as a Missing-Data Mechanism

Observations may be missing from a data set for a number of reasons depending on the nature of the collected data. The connection between missing observations and some unmeasured explanatory variable is the missing data mechanism: the underlying phenomenon responsible for the missing data. Rubin (1976) discusses the importance of understanding why data may be missing and why it is desirable to identify this mechanism. In simple terms, the data Y comprises observed and missing components: $Y = (Y_{obs}, Y_{mis})$, as in Little and Rubin (2002). In surveys the missing component can be participant nonresponses or refusals. In the latter case, there are often specific reasons why a participant chooses not to provide information, which if known provide additional insight into the data analysis. Furthermore, Little and Rubin (2002) demonstrated that underlying values may be observed but might be limited by the survey technique or equipment. Although time series data are quite different from surveys, this situation remains valid in SHM.

Mobile sensors simultaneously record measurements in time while moving in space. With respect to a single sampling location, an incomplete time series is created. As mobile sensors repeatedly scan over a set of points in space, structural vibration information is populated. With knowledge of the mobile sensors' sampling paths, the missing data mechanism with respect to the fixed-sensor case is a known function.

In other words, consider a hypothetical dense fixed-sensor network having sampling locations (a spatial grid) coincident with those of the mobile-sensor network. Matarazzo and Pakzad (2013) demonstrated that the anticipated mobile sensing data are a specific subset of the full, hypothetical, dense fixed-sensor data set. The subset corresponds to the movement of the mobile sensors in the network, which defines the missing data mechanism. Lastly, mobile sensing data can be simulated from fixed-sensor data, as shown in Matarazzo and Pakzad (2014), by removing observations that are not consistent with a known path of the mobile sensors. Fig. 1 shows a mobile-sensor network and its resulting data matrix. In this example, a single mobile sensor scans across a simple beam and records vertical acceleration at N unique sampling locations. After a time delay of d samples, the mobile-sensor path is repeated. The resulting data matrix is incomplete because the entries that do not coincide with the mobile-sensor scans are the missing observations. If a perfectly functioning fixed-sensor network were implemented at all sensing locations in Fig. 1, this data matrix would be complete and would not have missing observations.

Missing-Data Analyses in Time Series Models

Although survey data may provide sufficient conceptual examples to illustrate the importance of an underlying missing data mechanism, the current study exclusively analyzes time series data, which

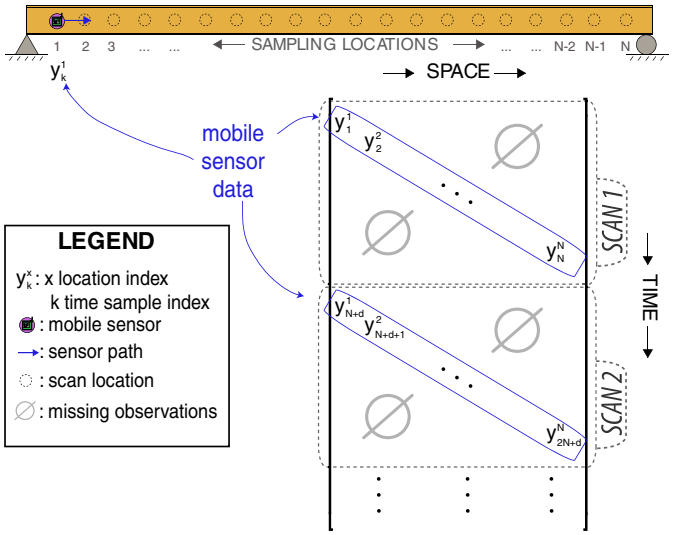


Fig. 1. Mobile-sensor data matrix, assuming that each mobile sensor scan records vertical acceleration at N sampling locations; d = delay between sensor scans

contain fundamentally different characteristics [see Box et al. (2008) or Shumway and Stoffer (2011)]. Consequently, missing-data strategies were developed to preserve such properties. Jones (1962, 1971) derived the asymptotic variance for estimating the spectral density of a discrete stationary stochastic process, and Parzen (1961) analyzed a special case of amplitude-modulated series.

Jones (1980) used maximum-likelihood (ML) ARMA fitting in the case of missing observations, introducing the idea that certain Kalman filter and likelihood calculations may be skipped for missing observations. This work quickly influenced numerous related papers throughout the 1980s. Dunsmuir and Robinson (1981a, b) developed alternative parametric estimation techniques in the presence of missing data based on the missing-data mechanism.

Shumway and Stoffer (1982) provided a foundation for expectation maximization (EM) in the state-space model with missing observations, using the Kalman filter and smoother for the E-step and providing specific M-step parameter update equations. This strategy considered the observation matrix C to be a known, fixed quantity; its current presentation is available in Shumway and Stoffer (2011). Harvey and Pierse (1984) considered ML parameter estimation of ARIMA models with missing data. They used the Kalman filter and fixed-point smoother in the state-space form to estimate and predict missing values. Stoffer (1986) presented a STARMAX model for spatiotemporal data with missing observations in fishing data applications (Mendelssohn and Roy 1986).

Digalakis et al. (1993) revisited EM in the state-space model, like Shumway and Stoffer (1982), but incorporated the observation matrix C as a free parameter; that is, this parameter was included in M-step updates and provided applications in speech recognition. Sinopoli et al. (2004) considered Kalman filtering with intermittent observations, assuming that the observations were binary random variables, and provided upper and lower bounds on a critical arrival probability based on the stability of the filtered covariance. Huang and Dey (2007), Kluge et al. (2010), Plarre et al. (2009), and Mo and Sinopoli (2011, 2012) also studied the behavior of the Kalman filter with missing data.

STRIDE Overview

In this section, the central concepts of the STRIDE algorithm are discussed and the underlying statistics of the mathematical model are reviewed. For more information on STRIDE, the reader is referred to Matarazzo and Pakzad (2016) and Table 1, which provides state-space model details. In the current study, the output-only STRIDE method was formulated in the stochastic state-space model as defined by Eqs. (1) and (2) and also found in Andersen et al. (1999), Chang and Pakzad (2013), and Peeters and De Roeck (1999)

$$\mathbf{x}_k = A\mathbf{x}_{k-1} + \boldsymbol{\eta}_k \quad k = 2, 3, \ldots, K \tag{1}$$

$$\mathbf{y}_k = C\mathbf{x}_k + \boldsymbol{\upsilon}_k \quad k = 1, 2, \ldots, K \tag{2}$$

where $\mathbf{y}_k$ is the observation vector; and $\mathbf{x}_k$ is the state vector for time step k.

The initial state vector $\mathbf{x}_1$ (for the first time step) was assumed to be random Gaussian with mean vector $\bar{\boldsymbol{\mu}}$ and $pN \times pN$ covariance matrix $\bar{V}$. The state-loading and observation noise terms were assumed to be independent, zero-mean, random Gaussian vectors for every time step k [where $N()$ represented the normal distribution] with diagonal covariance matrices Q and R, respectively.

Table 1. Stochastic State-Space Model Terms for STRIDE

Notation	Size	Description
N	Scalar	Observation size; number of sensing locations
k	Scalar	Time step index (subscript)
K	Scalar	Number of total time samples
K^*	Scalar	Number of time steps with missing characteristics (incomplete)
K^c	Scalar	Number of time steps without missing characteristics (complete)
p	Scalar	State-space model order (minimum $p = 2$)
pN	Scalar	Size of state variable (product of p and N)
θ	Scalar	Likelihood slope threshold ($\theta = 5 \times 10^{-4}$ is recommended)
$\mathbf{y}_k$	$N \times 1$	Observation vector at time step k
$\mathbf{x}_k$	$pN \times 1$	State vector at time step k
$V_{k,k}$	$pN \times pN$	State covariance at time step k
$V_{k,k-1}$	$pN \times pN$	Covariance between states at time steps k and $k-1$.
$\boldsymbol{\upsilon}_k$	$N \times 1$	Observation noise terms at time step k
$\boldsymbol{\eta}_k$	$pN \times 1$	State-loading terms at time step k
C	$N \times pN$	Observation matrix (time invariant)
A	$pN \times pN$	State (transition) matrix (time invariant)
R	$N \times N$	Observation noise covariance matrix (time invariant)
Q	$pN \times pN$	State error covariance matrix (time invariant)

Note: Vectors are in bold.

$$\mathbf{x}_1 \sim N(\bar{\boldsymbol{\mu}}, \bar{V}) \qquad \boldsymbol{\eta}_k \sim N(0, Q) \qquad \boldsymbol{\upsilon}_k \sim N(0, R) \tag{3}$$

STRIDE (Matarazzo and Pakzad 2016) is an output-only method for modal identification that embeds the EM algorithm (Dempster et al. 1977). The EM algorithm iterates between the E-step and the M-step. The objective of STRIDE is to determine maximum-likelihood estimates (MLEs) of six model parameters through maximization of the conditional expectation of the log-likelihood function. The parameters are consolidated into a super-parameter Ψ_j specified for iteration j:

$$\Psi_j = (\bar{\boldsymbol{\mu}}_j, \bar{V}_j, A_j, Q_j, C_j, R_j) \tag{4}$$

The algorithm begins with an initial super-parameter estimate Ψ_0 that is updated at each iteration via the M-step [Eqs. (14)–(19) in Matarazzo and Pakzad (2016)], which was designed to guarantee an increase in the conditional log-likelihood function (Dempster et al. 1977; Wu 1983). When the log-likelihood function reaches its maximum value, the algorithm ends, producing the MLE of the super-parameter for the model. More precisely, convergence is determined when the measured slope of the log-likelihood function falls below the slope threshold θ (at this point the likelihood slope is practically zero). As depicted in Figs. 6 and 13 in Matarazzo and Pakzad (2016), for a given data set and model order, smaller slope thresholds require a higher number of iterations and therefore, greater computational effort.

The STRIDE procedure is as follows:

1. Select super-parameter Ψ_0 where $j = 0$ represents initialization;
2. E-step: input the current super-parameter Ψ_0 and the observations $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ into the Kalman filter equations (Kalman 1960) and then the Rauch-Tung-Striebel (RTS) smoother equations (Rauch et al.1965) provided in Matarazzo and Pakzad (2016); these equations produce minimum mean square error (MMSE) estimates for the state variable and its covariances—namely $\hat{\mathbf{x}}_k$, $\hat{V}_{k,k}$, and $\hat{V}_{k,k-1}$ for all time steps (estimates are denoted by hats);
3. Compute the conditional expectation of the log-likelihood function and statistical measures pertaining to the state-variable estimates;

4. M-step: update the super-parameter using the computed statistics and Eqs. (14)–(19) in Matarazzo and Pakzad (2016); this update represents the beginning of a new iteration, $\Psi_j \rightarrow \Psi_{j+1}$; and
5. Repeat Steps 2–4 until the slope threshold θ is met; the final Ψ_{ML} provide the MLE for the model.

STRIDE Missing-Data Equations

Fig. 2 compares data sets with full data and those with missing data. Consider K samples of the $N \times 1$ observation vector $\mathbf{y}_k$. In the full-data case [Fig. 2(a)]. All K time steps are referenced simultaneously as "full data" and are represented mathematically by Y_K. In this case, the data are complete; all entries in the data matrix are available. In the missing-data case [Fig. 2(b)], the data set is represented mathematically by Y_K^* with K total time steps and comprises K^c full-data time steps and K^* missing-data time steps, so $K = K^c + K^*$. As shown in Fig. 2(b), the incomplete data are composed of two parts: an available-data portion and a missing-data portion. For a given time step in the incomplete data, the observation vector contains an available part with size a and a missing part with size m.

Shumway and Stoffer (1982) presented a missing-data EM algorithm strategy for the state-space model that addresses partially observed time steps by partitioning the observation Eq. (2). In addition to the current matrix and vector format (italic and bold, respectfully), submatrices and subvectors are underlined to emphasize reduced dimensions. Eq. (5) revisits the observation Eq. (2) for time steps in the incomplete data depicted in Fig. 2(b).

$$\mathbf{y}_k^* = C^* \mathbf{x}_k + \boldsymbol{\upsilon}_k^* \tag{5}$$

Eq. (5) is split into available and missing partitions, resulting in Eq. (6). This is only necessary for incomplete data time steps.

$$\begin{pmatrix} \underline{\mathbf{y}}_k \\ \underline{\mathbf{y}}_k^* \end{pmatrix} = \begin{bmatrix} \underline{C} \\ \underline{C}^* \end{bmatrix} \mathbf{x}_k + \begin{pmatrix} \underline{\boldsymbol{\upsilon}}_k \\ \underline{\boldsymbol{\upsilon}}_k^* \end{pmatrix} \tag{6}$$

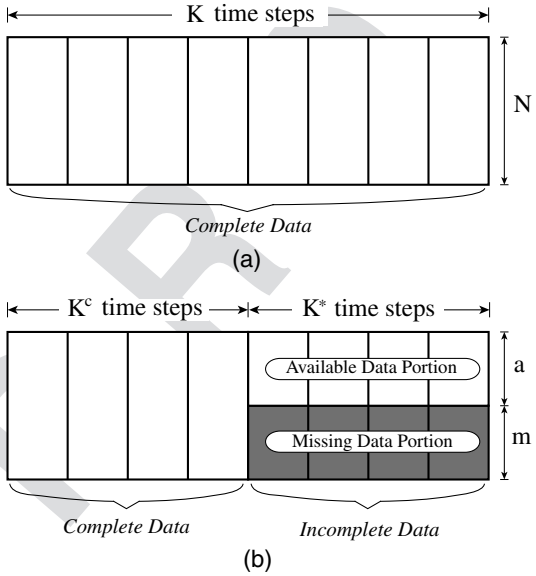


Fig. 2. (a) Schematic of the full-data case: K time steps, each with size N; (b) schematic of the missing data case containing complete data and incomplete data; the incomplete data comprises available and missing portions with sizes a and m, respectively

In Eq. (6), the $N \times 1$ observation vector $\mathbf{y}_k^*$ is partitioned into an $m \times 1$ vector $\underline{\mathbf{y}}_k^*$ of missing observations and an $a \times 1$ vector $\underline{\mathbf{y}}_k$ of available observations, with $N = a + m$. Similarly, the observation matrix C^* is divided (row-wise) into missing and available partitions: an $m \times pN$ matrix $\underline{C}^*$ and an $a \times pN$ matrix $\underline{C}$. Because the noise is assumed to be independent in STRIDE [resulting in the observation noise covariance being diagonal as in Matarazzo and Pakzad (2016)], the off-diagonal submatrices in R are zero matrices, i.e., observed noise (corresponding to available observations) is uncorrelated with unobserved noise (corresponding to missing observations):

$$R = \mathrm{cov}\begin{pmatrix} \underline{\boldsymbol{\upsilon}}_k \\ \underline{\boldsymbol{\upsilon}}_k^* \end{pmatrix} = \begin{bmatrix} \underline{R} & 0 \\ 0 & \underline{R}^* \end{bmatrix} \tag{7}$$

Although the partition positions implied in Fig. 2 and in this section may appear restrictive, apparently restraining the organization of the available and missing observations, this is not the case; it is not necessary to arrange data or matrix positions exactly as presented. Shumway and Stoffer (2011) addressed this concern by introducing a permutation matrix, but for many computational and coding implementations, such a matrix is not required; simple programming and organization of array indices in computational software can circumvent these issues.

In general, a and m are both functions of time step k. The numerical values must be positive integers with a sum equal to N. In other words, the size of the available partitions (and therefore the missing partitions) can vary among time steps within the incomplete data. For mathematical completeness, this formulation can be reduced to the familiar full-data case to treat complete data using $a = N$ and $m = 0$, in which case the missing partitions vanish.

To review, in the case of missing data the time steps are separated into complete and incomplete data. The complete data do not contain missing characteristics and can be modeled normally. The incomplete data are composed of an available-data portion and a missing-data portion—these portions are separated using matrix partitions as discussed in Eqs. (5)–(7). In the case of mobile sensing, all time steps are incomplete thus contain missing observations; thus, $K^* = K$ and $K^c = 0$.

Modified E-Step: Kalman Equations with Missing Observations

The Kalman prediction and filter equations in STRIDE provide minimum mean square error (MMSE) estimates of the states and state covariances from the observations and the current super-parameter (these estimates are denoted with hats) as defined in Eqs. (8)–(12). The Kalman equations are in terms of the state-space model in Shumway and Stoffer (1982), Digalakis et al. (1993), Ghahramani and Hinton (1996), Box et al. (2008), and Matarazzo and Pakzad (2016).

$$\hat{\mathbf{x}}_{k|K} \equiv E[\mathbf{x}_k | Y_K, \Psi_j] \tag{8}$$

$$\hat{V}_{k,k|K} \equiv E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|K})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|K})^T | Y_K, \Psi_j] \tag{9}$$

$$\hat{V}_{k-1,k-1|K} \equiv E[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|K})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|K})^T | Y_K, \Psi_j] \tag{10}$$

$$\hat{V}_{k,k-1|K} \equiv E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|K})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|K})^T | Y_K, \Psi_j] \tag{11}$$

$$\hat{V}_{k-1,k|K} \equiv E[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|K})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|K})^T | Y_K, \Psi_j] \tag{12}$$

In the event that some observations are missing, the Kalman filter remains the optimal filter (Kalman 1960). To process data of this class, the Kalman equations are modified at the K^* incomplete time steps, which contain missing observations. Typical Kalman equations make predictions for the next time step and then update these predictions using the observation at that time step. If the observation information at the next time step is missing, the predictions are unchanged in form; however, there is limited or no information available for the filter calculations to update the state estimates. This result is most evident when all observations are missing at a time step, where the observations provide no information to the filtering step and the predicted value is unchanged; that is, the filtered estimates are identical to the predictions.

$$\hat{\mathbf{x}}^*_{k|K} \equiv E[\mathbf{x}_k|Y^*_K, \Psi_j] \tag{13}$$

$$\hat{V}^*_{k,k|K} \equiv E[(\mathbf{x}_k - \hat{\mathbf{x}}^*_{k|K})(\mathbf{x}_k - \hat{\mathbf{x}}^*_{k|K})^T|Y^*_K, \Psi_j] \tag{14}$$

$$\hat{V}^*_{k-1,k-1|K} \equiv E[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}^*_{k-1|K})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}^*_{k-1|K})^T|Y^*_K, \Psi_j] \tag{15}$$

$$\hat{V}^*_{k,k-1|K} \equiv E[(\mathbf{x}_k - \hat{\mathbf{x}}^*_{k|K})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}^*_{k-1|K})^T|Y^*_K, \Psi_j] \tag{16}$$

$$\hat{V}^*_{k-1,k|K} \equiv E[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}^*_{k-1|K})(\mathbf{x}_k - \hat{\mathbf{x}}^*_{k|K})^T|Y^*_K, \Psi_j] \tag{17}$$

In short, the E-step missing-data equations are derived from modified conditional expectations of the states (because $\hat{\mathbf{x}}_{k|K}$ becomes $\hat{\mathbf{x}}^*_{k|K}$) shown in Eqs. (13)–(17). This is because Y_K has now become Y^*_K.

Despite the change in notation for the state variables (namely, the * superscript), the states and state covariances are not partitioned or changed directly; only the observation dimensions (size N) are partitioned. The superscript indicates that the observed data, which are used as input to the Kalman filter, contain missing information. For zero-mean stationary observations, the following missing data substitutions, recommended in Shumway and Stoffer (1982, 2011), are applicable to the E-step in STRIDE. At time steps with incomplete data, the missing partitions can be assigned to follow the values in Eqs. (18)–(21).

$$\underline{\mathbf{y}}^*_k = 0 \tag{18}$$

$$\underline{C}^* = 0 \tag{19}$$

$$\underline{\mathbf{v}}^*_k = 0 \tag{20}$$

$$\underline{R}^* = I \tag{21}$$

Equation (20) is the expected value of the observation noise, and Eq. (18) is the result of substituting Eqs. (19) and (20) into Eq. (6). Additionally, Eq. (21) is consistent with the initial parameter estimate R_0 suggested in Matarazzo and Pakzad (2016). Finally, although C is assumed known in Shumway and Stoffer (2011), this substitution is still appropriate because its purpose, eliminating these observations from the Kalman filter, is fulfilled. Shumway and Stoffer (1981) proved that by following the missing-data Kalman filter calculations provided in the following sections, the filtered states and covariances can proceed through the RTS smoother equations to produce the MMSE estimates defined in Eqs. (13)–(17). These smoothing calculations are as presented in Matarazzo and Pakzad (2016). Although the observation model may be reduced to exclude the m unavailable observations at the time step, it is computationally more efficient to maintain consistent model dimensions throughout the calculations.

Time Steps Missing Some Observations

This section investigates Kalman filter equations for time steps in which some observations are missing. The following missing-data results are obtained by substituting Eqs. (18)–(21) into partitioned Kalman filter equations

$$\underline{\boldsymbol{\varepsilon}}^*_k = 0 \tag{22}$$

$$\underline{\Sigma}^*_k = I \tag{23}$$

$$\underline{K}^*_k = 0 \tag{24}$$

where $\underline{\boldsymbol{\varepsilon}}^*_k$ is the $m \times 1$ missing partition of the prediction error vector with $m \times m$ covariance submatrix $\underline{\Sigma}^*_k$; and $\underline{K}^*_k$ is the $pN \times m$ Kalman gain submatrix [Eqs. (27)–(29) provide details].

The $pN \times m$ Kalman gain matrix K_k is unique because the observation dimension spans the columns rather than the rows. Equation (24) means that the $pN \times m$ columns of the Kalman gain matrix corresponding to missing observations are set to zero. The Kalman prediction Eqs. (25) and (26) are unchanged in form:

$$\hat{\mathbf{x}}^*_{k|k-1} = A\hat{\mathbf{x}}^*_{k-1|k-1} \tag{25}$$

$$\hat{V}^*_{k,k|k-1} = A\hat{V}^*_{k-1,k-1|k-1}A^T + Q \tag{26}$$

The Kalman filter/update Eqs. (27)–(32) are partitioned in the observation dimension and include the results from Eqs. (22)–(24).

$$\boldsymbol{\varepsilon}^*_k = \begin{pmatrix} \underline{\boldsymbol{\varepsilon}}_k \\ \underline{\boldsymbol{\varepsilon}}^*_k \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{y}}_k - \underline{C}\hat{\mathbf{x}}^*_{k|k-1} \\ 0 \end{pmatrix} \tag{27}$$

$$\Sigma^*_k = \begin{bmatrix} \underline{\Sigma}_k \\ \underline{\Sigma}^*_k \end{bmatrix} = \begin{bmatrix} \underline{C}\hat{V}^*_{k,k|k-1}\underline{C}^T + \underline{R} \\ I \end{bmatrix} \tag{28}$$

$$K^*_k = [\,\underline{K}_k \quad \underline{K}^*_k\,] = [\,\hat{V}^*_{k,k|k-1}\underline{C}^T\underline{\Sigma}^{-1}_k \quad 0\,] \tag{29}$$

$$\hat{\mathbf{x}}^*_{k|k} = \hat{\mathbf{x}}^*_{k|k-1} + K^*_k\boldsymbol{\varepsilon}^*_k \tag{30}$$

$$\hat{V}^*_{k,k|k} = (I - K^*_kC^*)\hat{V}^*_{k,k|k-1} \tag{31}$$

$$\hat{V}^*_{k,k-1|k-1} = (I - K^*_kC^*)A\hat{V}^*_{k-1,k-1|k-1} \tag{32}$$

Time Steps Missing All Observations

This section provides the Kalman filter equations for time steps in which the observations are completely unavailable or all observations are missing (the case of $m = N$, thus $a = 0$), hereafter referred to as time step blanks ["blackouts" in Digalakis et al. (1993)]. As first introduced in Jones (1980) and later adapted by Harvey and Pierse (1984), Shumway and Stoffer (1981), and Digalakis et al. (1993), the Kalman prediction is not updated in the Kalman filter equations. As a result, the predicted estimates are equal to the filtered estimates for these time steps; that is, the prediction error is a $N \times 1$ vector of zeros and the Kalman gain is a $pN \times N$ zero matrix at this time step. The following missing data replacements of Eqs. (33)–(38) correspond to equations in Sinopoli et al. (2004) for $\gamma_t = 0$:

$$\boldsymbol{\varepsilon}_k^* = 0 \tag{33}$$

$$\Sigma_k^* = I \tag{34}$$

$$K_k^* = 0 \tag{35}$$

$$\hat{\mathbf{x}}_{k|k}^* = \hat{\mathbf{x}}_{k|k-1}^* \tag{36}$$

$$\hat{V}_{k,k|k}^* = \hat{V}_{k,k|k-1}^* \tag{37}$$

$$\hat{V}_{k,k-1|k-1}^* = A\hat{V}_{k-1,k-1|k-1}^* \tag{38}$$

In other words, as Jones (1980) first stated, large blocks of consecutive missing observations slowly reduce information about the past in the state vector: the estimated state approaches the initial state vector, and the estimated state covariance approaches the initial state covariance matrix.

Modified M-Step: Likelihood Function with Missing Observations

With estimated (filtered and smoothed) states and state covariances at all time steps, the next iteration of the state-space parameters can be computed from the following conditional expected values. In short, for the M-step missing-data equations, the expected values for the observations are affected (Y_K becomes Y_K^*). Most importantly, the partition substitutions provided for the Kalman filter in the E-step in Eqs. (18)–(21) are not to be used in the M-step. Partitioned values remain as specified rows/submatrices of the current parameter at iteration j unless otherwise noted: $\underline{C}^* = \underline{C}_j^* \neq 0$ and $\underline{R}^* = \underline{R}_j^* \neq I$.

The following updating equations are similar to those in Digalakis et al (1993), but are explicitly extended to the parameter updates in terms of the observed data, estimated states, state covariances, and other model parameters (which may be partitioned). Eq. (39) defines a shorthand notation for a recurring conditional expectation.

$$E^{(*,j)}[(\cdot)] \equiv E[(\cdot)|Y_K^*, \Psi_j] \tag{39}$$

$$E^{(*,j)}[\mathbf{y}_k] = \begin{cases} \underline{\mathbf{y}}_k^* & \text{if available} \\ \underline{C}_j^* \hat{\mathbf{x}}_{k|K}^* & \text{if missing} \end{cases} \tag{40}$$

$$E^{(*,j)}[\mathbf{y}_k \mathbf{y}_k^T] = \begin{cases} \underline{\mathbf{y}}_k^* (\underline{\mathbf{y}}_k^*)^T & \text{if available} \\ \underline{C}_j^* \hat{\mathbf{x}}_{k|K}^{(*)} (\hat{\mathbf{x}}_{k|K}^*)^T (\underline{C}_j^*)^T + \underline{R}_j^* & \text{if missing} \end{cases} \tag{41}$$

$$E^{(*,j)}[\mathbf{y}_k \mathbf{x}_k^T] = \begin{cases} \underline{\mathbf{y}}_k^* (\hat{\mathbf{x}}_{k|K}^*)^T & \text{if available} \\ \underline{C}_j^* \hat{\mathbf{x}}_{k|K}^* (\hat{\mathbf{x}}_{k|K}^*)^T & \text{if missing} \end{cases} \tag{42}$$

$$\begin{aligned} A_{j+1} &= \sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{x}_{k-1}^T] \left(\sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_{k-1} \mathbf{x}_{k-1}^T] \right)^{-1} \\ &= \sum_{k=2}^{K} [\hat{\mathbf{x}}_{k|K}^* (\hat{\mathbf{x}}_{k-1|K}^*)^T + \hat{V}_{k,k-1|K}^*] \left(\sum_{k=2}^{K} [\hat{\mathbf{x}}_{k-1|K}^* (\hat{\mathbf{x}}_{k-1|K}^*)^T + \hat{V}_{k-1,k-1|K}^*] \right)^{-1} \end{aligned} \tag{43}$$

$$\begin{aligned} Q_{j+1} &= \frac{1}{K-1} \left(\sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{x}_k^T] - \sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{x}_{k-1}^T] \left(\sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_{k-1} \mathbf{x}_{k-1}^T] \right)^{-1} \sum_{k=2}^{K} E^{(*,j)}[\mathbf{x}_{k-1} \mathbf{x}_k^T] \right) \\ &= \frac{1}{K-1} \left(\sum_{k=2}^{K} [\hat{\mathbf{x}}_{k|K}^* (\hat{\mathbf{x}}_{k|K}^*)^T + \hat{V}_{k,k|K}^*] - A_{j+1} \sum_{k=2}^{K} [\hat{\mathbf{x}}_{k-1|K}^* (\hat{\mathbf{x}}_{k|K}^*)^T + \hat{V}_{k-1,k|K}^*] \right) \end{aligned} \tag{44}$$

The updates for the observation matrix C and the observation noise covariance R are unique because they directly refer to the observation equation, where missing data effects are most apparent. In Eqs. (45) and (46), the summation over all K observations is split into a sum over K^c and a sum over K^* (with temporary time step indices k^c and k^*, respectively). The sums over K^c provide the statistics for complete data. The sums over K^* provide statistics for incomplete data, which contain two portions and are partitioned in the same manner as the conditional expectations in Eqs. (40)–(42): within these summations, the top submatrix corresponds to the a available rows whereas the bottom submatrix corresponds to the m missing rows.

$$\begin{aligned} C_{j+1} &= \sum_{k=1}^{K} E^{(*,j)}[\mathbf{y}_k \mathbf{x}_k^T] \left(\sum_{k=1}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{x}_k^T] \right)^{-1} \\ &= \left(\sum_{k^c=1}^{K^c} \mathbf{y}_{k^c}^* (\hat{\mathbf{x}}_{k^c|K}^*)^T + \sum_{k^*=1}^{K^*} \begin{Bmatrix} \underline{\mathbf{y}}_{k^*}^* (\hat{\mathbf{x}}_{k^*|K}^*)^T \\ \underline{C}_j^* \hat{\mathbf{x}}_{k^*|K}^* (\hat{\mathbf{x}}_{k^*|K}^*)^T \end{Bmatrix} \right) \left(\sum_{k=1}^{K} [\hat{\mathbf{x}}_{k|K}^* (\hat{\mathbf{x}}_{k|K}^*)^T + \hat{V}_{k,k|K}^*] \right)^{-1} \end{aligned} \tag{45}$$

$$R_{j+1} = \frac{1}{K}\left(\sum_{k=1}^{K} E^{(*,j)}[\mathbf{y}_k \mathbf{y}_k^T] - \sum_{k=1}^{K} E^{(*,j)}[\mathbf{y}_k \mathbf{x}_k^T]\left(\sum_{k=1}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{x}_k^T]\right)^{-1} \sum_{k=1}^{K} E^{(*,j)}[\mathbf{x}_k \mathbf{y}_k^T]\right)$$

$$= \frac{1}{K}\left(\sum_{k^c=1}^{K^c}[\mathbf{y}_{k^c}^*(\mathbf{y}_{k^c}^*)^T - C_{j+1}\hat{\mathbf{x}}_{k^c|K}^*(\mathbf{y}_{k^c}^*)^T] + \sum_{k^*=1}^{K^*}\left\{\begin{array}{l}\underline{\mathbf{y}}_{k^*}\underline{\mathbf{y}}_{k^*}^T - \underline{C}_{j+1}\hat{\mathbf{x}}_{k^*|K}^*\underline{\mathbf{y}}_{k^*}^T \\ [\underline{C}_j^* - \underline{C}_{j+1}^*]\hat{\mathbf{x}}_{k^*|K}^*(\hat{\mathbf{x}}_{k^*|K}^*)^T(\underline{C}_j^*)^T + \underline{R}_j^*\end{array}\right\}\right) \quad (46)$$

$$\bar{\boldsymbol{\mu}}_{j+1} = \hat{\mathbf{x}}_{1|K}^* \quad (47)$$

$$\bar{V}_{j+1} = \hat{V}_{1,1|K}^* \quad (48)$$

Convergence and Performance with Missing Data

Ten Degree-of-Freedom Shear Structure Case Study

In this section, the simulated acceleration response of the 10-story shear frame structure shown in Fig. 3 is considered with 2% damping in all structural modes [similar to the 5-story structure in Chang and Pakzad (2012)]. The numerical responses were generated using a uniform (interval −0.5 to 0.5) random excitation at each degree of freedom (DOF) and Newmark's linear acceleration method. It is important to note that the stochastic state-space model in Eq. (2) assumes Gaussian, not uniform, loading; therefore, a successful identification in this case indicates that this model assumption is not restrictive. To simulate measurement noise, white noise with a standard error equal to 5% RMS of the signal was added to the response at each degree-of-freedom (DOF).

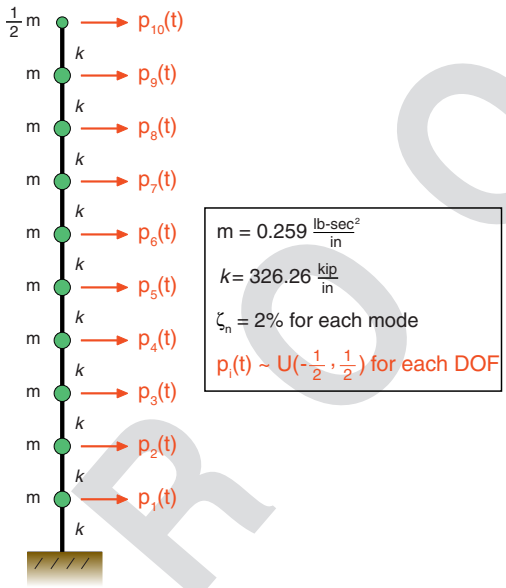


Fig. 3. 10-DOF shear story structure

To simulate the missing data, three missing data mechanism cases were considered:

- Randomly remove observations throughout the response (partial observations);
- Randomly remove full time steps (time step blanks); and
- Remove unobserved entries from deterministic mobile-sensor paths.

For the removals, the entries missing from the noisy response matrix were replaced with zeros in accordance with Eq. (18). The degree of missingness (percentage of missing observations) were varied from 0 to 40% in these cases to represent a range of candidate magnitudes. To better capture performance variability in the missingness magnitudes, the two stochastic missing-data scenarios, (Cases 1 and 2, where the missing data mechanism was a random process) were simulated 30 times.

In the mobile-sensor scenario (Case 3), N_{MS} mobile sensors sampled data at N_{MS} out of N sensing locations at every time step. More specifically, for each group of N_{MS} adjacent sensors, the group shifted back and forth (up and down) in a harmonic fashion, scanning all N locations—the group (all sensors) changed its location at every time step. For the 10-DOF structure, 6, 7, 8, and 9 mobile-sensor groups corresponded to 40, 30, 20, and 10% missing data, respectively.

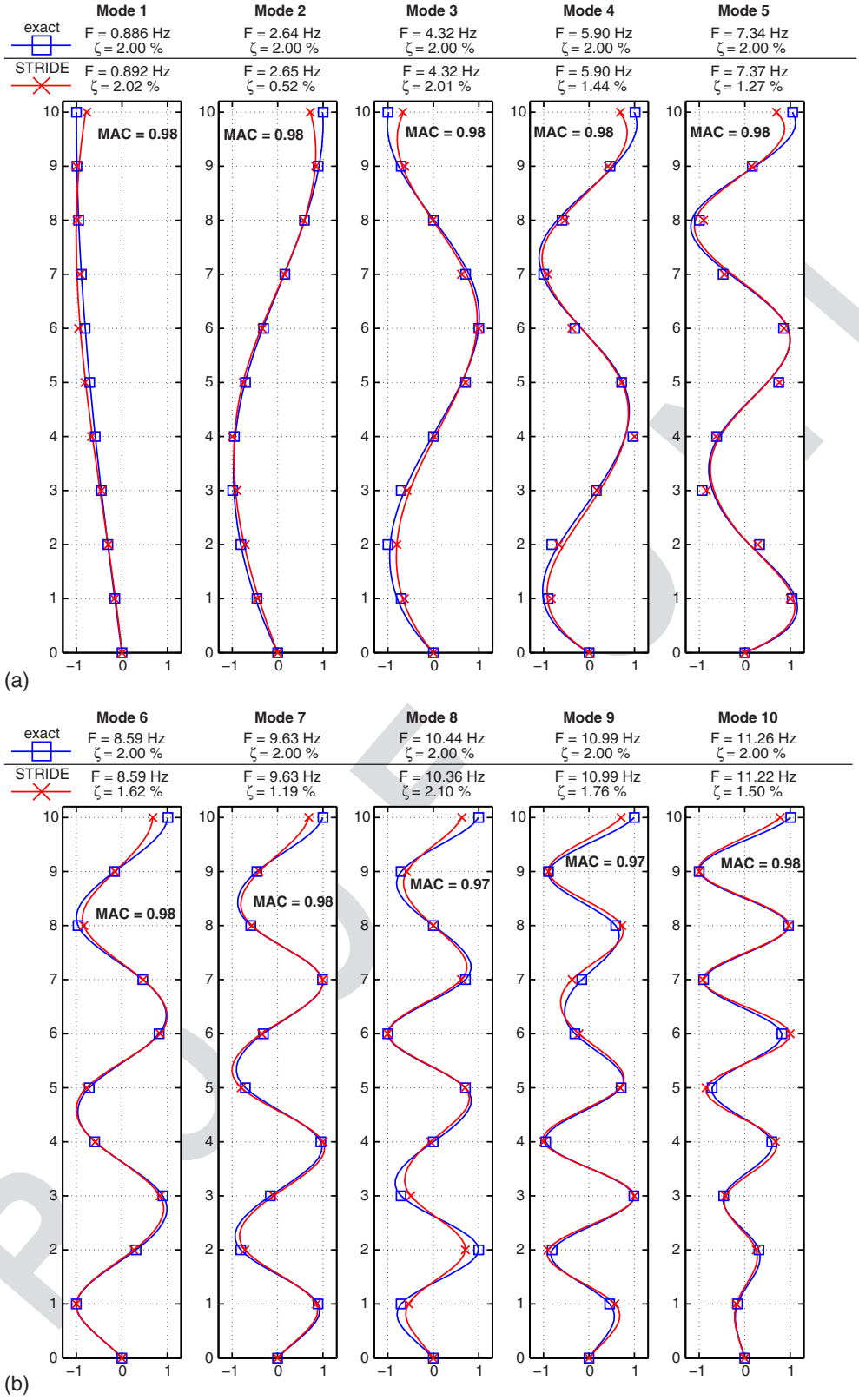
In addition to the three cases already introduced, two STRIDE analyses were included as baseline performances: a case of full data and a case of truncated data. The purpose of the full-data case was to establish the accuracy of modal estimates when there were no missing data—the best-case scenario—and to control additional model variables (e.g., model order and slope threshold remained fixed). The truncated analyses considered shortened data sets containing 5 to 40% missing data from the original sample size K and serve as a fair comparison to data sets with fewer observations, i.e., Cases 1–3. In the truncated case, a higher degree of missingness was equivalent to a reduced sample size (a decrease in K).

In all cases, the STRIDE analyses used a state-space model order equal to 4, denoted ($p = 4$), and the initial state matrix A_0

Table 2. Modal Properties of the 10-DOF Structure

Modal property	1	2	3	4	5	6	7	8	9	10
Exact frequencies	0.886	2.64	4.32	5.90	7.34	8.59	9.63	10.44	10.99	11.26
STRIDE frequencies	0.892	2.65	4.32	5.90	7.37	8.59	9.63	10.36	10.99	11.22
Exact damping	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
STRIDE damping	2.02	0.52	2.01	1.44	1.27	1.62	1.19	1.10	1.76	1.50

Note: Exact frequencies (Hz) and damping ratios (%) are compared with STRIDE ($p = 4$, $\theta = 5 \times 10^{-4}$).



F4:1 **Fig. 4.** STRIDE results compared with exact MDOF values for (a) modes 1–5; (b) modes 6–10

and observation matrix C_0 were estimated by ERA-NExT-AVG ($p = 4$). The remaining state-space parameters were set in accordance with Matarazzo and Pakzad (2016).

In the case of full data (no missing data), hereafter Case 0, STRIDE ($p = 4$) converged to MLE after 122 iterations; the corresponding modal estimates were compared with the exact MDOF values in Table 2, and Fig. 4. Table 2 compares STRIDE modal estimates with the exact natural frequencies and damping ratios for the 10-DOF structure. Overall, STRIDE successfully identified all 10 structural modes with an average absolute

frequency error of 0.3% over all modes and an average damping estimate of 1.54% (2.00% was the exact value for every mode). The average of all 10 modal assurance criteria (MAC) values (Allemang and Brown 1982) calculated between the estimated and exact mode shape ordinates was 0.98. Individual mode shapes with spline fits and MAC values are shown in Fig. 4. The following sections discuss the behavior and performance of STRIDE for missing-data Cases 1, 2, and 3.

Convergence Behavior

This section discusses the influence of missing data on STRIDE iterations and the magnitude of the final likelihood function value. The convergence criterion for STRIDE is based on the slope of the conditional likelihood function [see Matarazzo and Pakzad (2016) for details]. The actual slope is estimated as the difference divided by the average of the two most recent likelihood values. The user defines a specific slope threshold θ to decide when the conditional likelihood function practically attains its maximum value. Based on SHM applications in Matarazzo and Pakzad (2016), $\theta = 5 \times 10^{-4}$ is recommended as a guideline slope threshold for model orders $p = 2$ and $p = 4$; therefore, it was selected for this study. Smaller slope thresholds are conducive to more STRIDE iterations and greater computational effort.

When compared with the full-data case, missing observations prompted an alternative EM sequence and path along the likelihood function. From inspection of the likelihood function [see Eq. (6) in Matarazzo and Pakzad (2015)], it was expected that in cases of missing-data the final likelihood value would be lower than that in the full-data case because there were fewer observations to be explained by the same set of model parameters; that is, the number of time samples K was effectively reduced. In general, a lower final likelihood value implies less accurate MLE, but there was no evidence for its impact on the modal estimates.

However, details regarding the quantity of the reduced likelihood value, or the iterations required to get to this critical point are data-dependent and cannot be derived explicitly. This distinction is especially important for likelihood values because likelihood ratio tests are frequently implemented to assess model parameter estimates in similar time series models (King 1989; Nigro et al. 2014; Shahidi et al. 2015). Thus it is prudent to simultaneously evaluate the behavior of the converged likelihood values and investigate the accuracy of corresponding modal properties for various cases of missing data. Finally, in comparison with the full-data case, a missing-data STRIDE analysis may take greater or fewer iterations to achieve a given slope threshold. The cases in this paper aimed to establish an empirical trend because there was no mathematical guarantee for this relation.

Fig. 5(a) shows the number of required iterations versus the degree of missingness with $\pm 2\sigma$ confidence intervals (σ is the standard deviation) for different degrees of missingness; it compiles the results of 30 STRIDE runs for Cases 1 and 2 (because these two cases represented samples of random missing-data mechanism populations). In general, the plot indicates the impact of the missing-data mechanism on the computational costs of STRIDE—more iterations require more floating-point operations (FLOPs). The leftmost point indicates Case 0. In addition to the three cases examined, a truncated data analysis set a baseline for reduced observations. For the truncated data in Fig. 5(a), the number of iterations remained fairly constant. Individual iterations became computationally less expensive as the degree of missingness became higher (as K decreased). Cases 1 and 2 showed similar decreasing required iteration trends with respect to increasing missing data. In general, Case 3 required the highest number of iterations, with a peak of 149 for 20 and 30% missing data (8 and 7 mobile sensors, respectively). Together the plots in Fig. 5(a) may indicate missing-data mechanisms that are more uniform over time (e.g., Case 3) maintain stable identifications as the degree of missingness becomes higher. In such mechanisms, the number of consecutive missing observations is limited, restricting the growth of the prediction error variance in the Kalman filter.

Fig. 5(b) plots the normalized maximum likelihood value against the degree of missingness for three cases and truncated data. In this case, a higher degree of missingness was linked to a decrease in the number of time samples K. The linearly decreasing behavior of the likelihood function was thus to be expected after inspection of the likelihood function in Eq. (6) of Matarazzo and Pakzad (2016). Maximum likelihood values were normalized with respect to the value attained in Case 0 (the data point at the top left in Fig. 5(b) is equal to 100%).

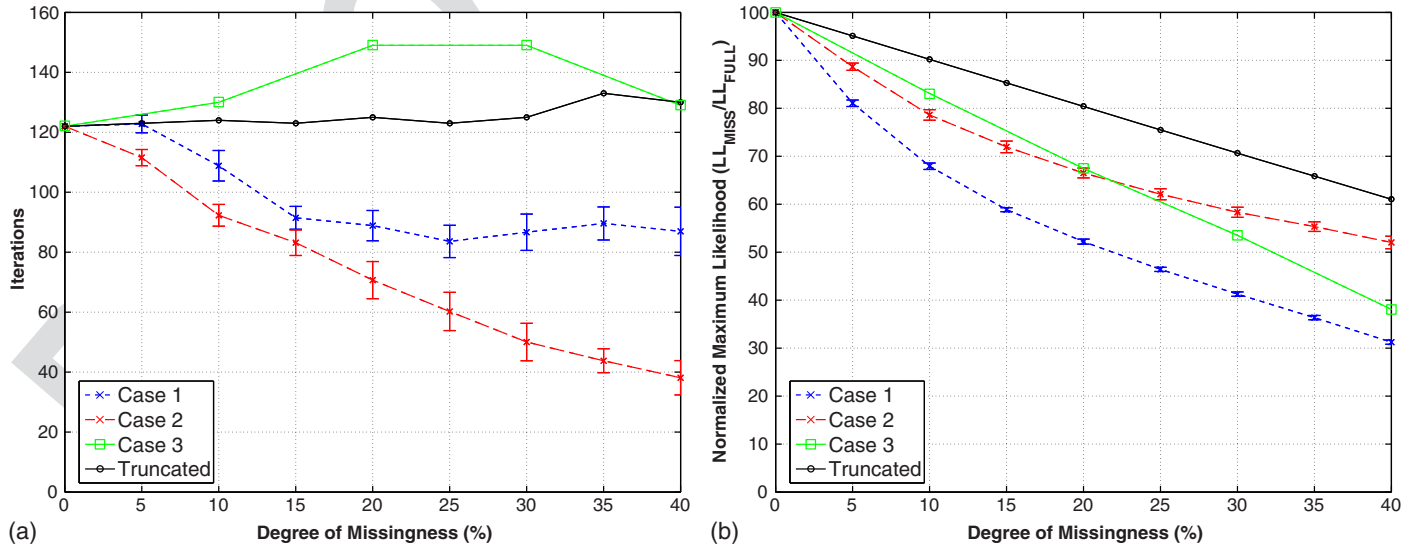
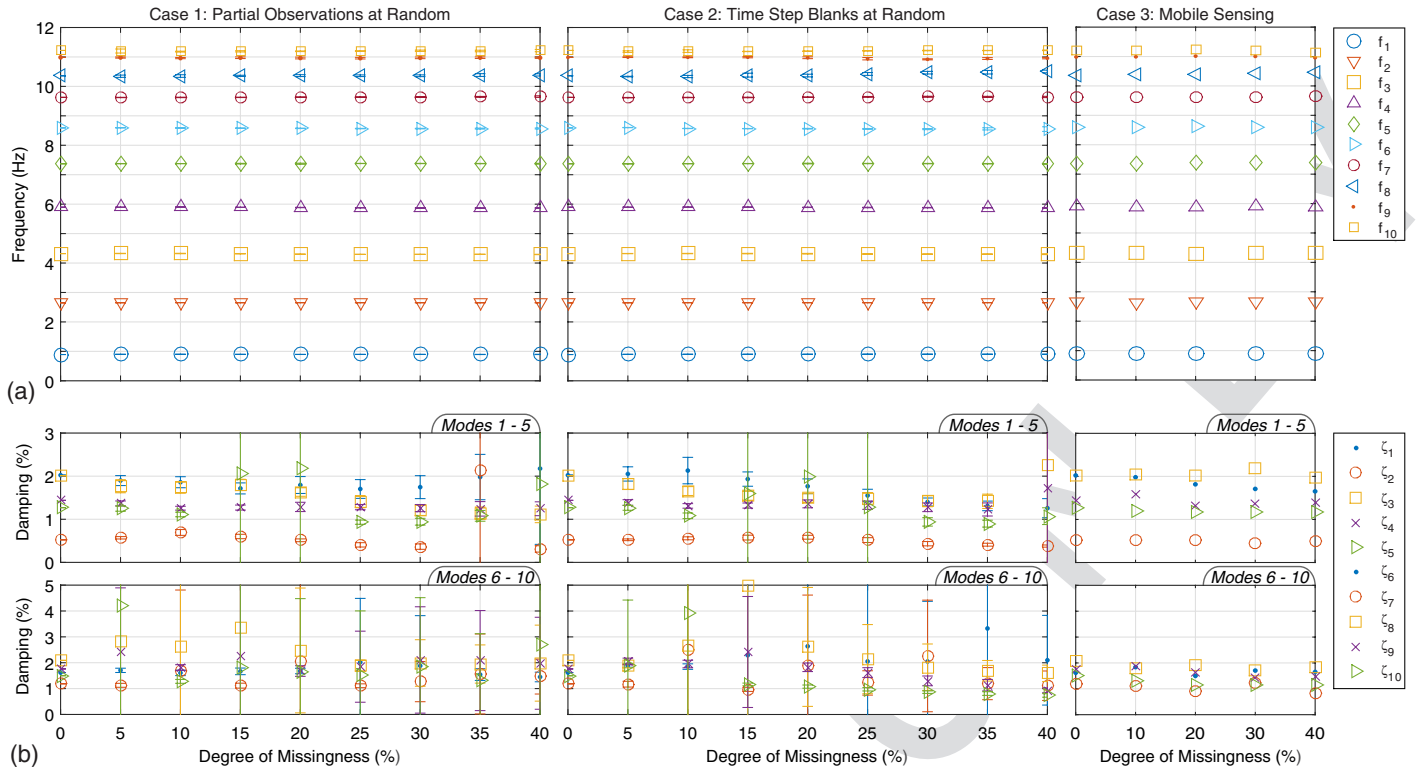


Fig. 5. (a) Required STRIDE iterations for convergence with $\pm 2\sigma$ for Cases 1 and 2; all three cases range from 0 to 40% missingness shown along truncated results; (b) STRIDE maximum likelihood values with $\pm 2\sigma$ for Cases 1 and 2; all three cases range from 0 to 40% missingness shown along truncated results



F6:1 F6:2 **Fig. 6.** (a) STRIDE mean frequency estimates for all three cases ($\pm 2\sigma$ included for Cases 1 and 2); missingness varies 0–40%; (b) STRIDE mean damping estimates for all three cases ($\pm 2\sigma$ included for Cases 1 and 2); missingness varies 0–40%

ML Modal Estimate Performance

For Case 0, STRIDE modal estimates were compared with the exact MDOF values in Table 2 and Fig. 4. The decreasing likelihood values in Fig. 5(b) may seem to imply that the STRIDE results for Cases 1, 2, and 3 yielded less accurate parameter estimates than those for Case 0; however, it is important to investigate individual mode performances to assess potential accuracy decay.

Fig. 6 shows the behavior of the frequency and damping estimates for all three cases (for detailed statistics on the modal estimates for these cases, the reader is referred to Appendix II). As shown in Fig. 6(a), in all three cases, all 10 frequency estimates remained accurate for degrees of missingness 0–40%; the results were nearly identical to those of Case 0. Additionally, the tight confidence bounds in Fig. 6(a) (for Cases 1 and 2) indicate a low variability in these estimates.

In Fig. 6(b) the damping estimates are plotted against missingness for all three cases. In Fig. 6(b), Case 1 (left), the mean damping estimates are mostly stable and maintain accuracy as the magnitude of missing data increases. One exception is the second mode, which is underestimated at approximately 0.50% for most missing-data cases; however, this is consistent with the estimate for Case 0.

In Fig. 6(b), Case 2 (middle), the mean damping estimates for the lower modes remained consistent with those in Case 0 whereas the higher modes fluctuated. Lastly, in Fig. 6(b), Case 3 (right) shows a stable, consistent damping performance as missingness increased with the highest precision out of all three cases.

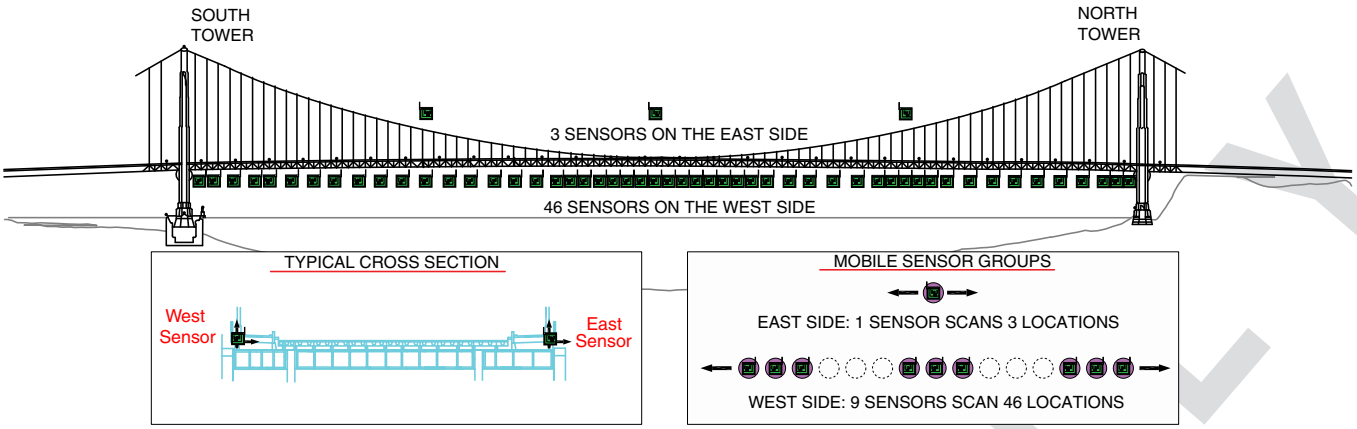
This study evaluated the accuracy of STRIDE modal estimates when the data were subjected to various missing-data patterns. Details on the estimated modal properties and their variation are provided in Appendix II. Overall, the results indicated that an accurate, full modal identification using STRIDE is feasible even when a data set is missing a significant number of observations. More specifically, given a fixed degree of missingness below 40%, the missing-data mechanism considered in Case 3 (mobile sensors) yielded the most accurate results out of the three, in principle supporting mobile sensors as an alternative to fixed sensors. In conclusion, the behavior of likelihood values with respect to missingness [Fig. 5(b)] and the case-wise performances (Fig. 6) show that a reduced model likelihood value does not necessarily indicate less accurate modal estimates.

STRIDE Missing-Data Applications

Golden Gate Bridge Setup

Structural modal identification is advantageous in cases where the collected data contain missing observations because such instances are difficult to avoid entirely. Moreover, the assumption that missing observations leave the remaining data fruitless is wasteful and not necessarily true. As shown in the previous section, comprehensive modal information can be extracted from vibration data even when a significant number of observations are unavailable.

The following sections discuss two applications of structural identification with missing observations using data collected at the Golden Gate Bridge: network communication reliability and mobile sensing. These applications were selected to provide an innovative solution to current SHM hurdles. Fig. 7 shows the wireless sensor network instrumentation from Pakzad et al. (2008), which obtained the data used in these applications for simulation of different missing-data scenarios. Forty-nine sensing locations were considered in the following analyses: 46 on the west side of the main span and 3 on the east side. Details of the sensor network design and implementation can be found in Pakzad et al. (2008) and Pakzad and Fenves (2009).



F7:1 **Fig. 7.** Golden Gate Bridge sensor configuration

Table 3. Comparison of Golden Gate Bridge Modal Estimates: AR (Data from Pakzad and Fenves 2009), STRIDE (Full Data), MP STRIDE (20% Missing), and MS STRIDE (82% Missing)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
T3:2	Modal property	A	S	S	A	T, A	S	T, S	A	T, A	S	A	T, S	S	T, A	A	T, S	S	T, A	A
T3:3	AR frequencies	0.106	0.132	0.170	0.216	0.230	0.301	0.34	0.371	0.445	0.461	0.550	0.566	0.660	0.682	0.769	0.812	0.887	0.945	1.000
T3:4	STRIDE frequencies	0.108	0.136	0.191	0.217	0.229	0.295	0.330	0.369	0.444	0.461	0.549	0.565	0.661	0.682	0.768	0.812	0.884	0.943	1.000
T3:5	$MP_{(20\%)}$ frequencies	0.106	0.132	0.170	0.217	0.228	0.300	0.340	0.368	0.444	0.461	0.550	0.565	0.661	0.681	0.768	0.812	0.884	0.947	1.000
T3:6	$MS_{(82\%)}$ frequencies	0.106	0.132	0.169	0.217	0.220	0.299	0.340	0.369	0.443	0.462	0.542	0.562	0.660	0.681	0.767	0.813	0.885	0.945	1.000
T3:7	AR damping	2.10	2.40	2.30	1.60	2.20	1.60	1.80	0.80	0.80	1.00	1.40	0.90	0.70	0.20	0.60	0.50	0.70	0.90	0.60
T3:8	STRIDE damping	5.74	3.15	8.74	1.80	1.36	1.29	7.60	0.91	0.59	0.87	1.41	0.47	0.42	0.34	0.52	0.44	0.70	0.77	0.74
T3:9	$MP_{(20\%)}$ damping	1.78	1.58	2.47	1.33	0.68	0.90	0.49	0.72	0.40	0.70	1.62	0.37	0.40	0.28	0.46	0.43	0.67	0.62	0.66
T3:10	$MS_{(82\%)}$ damping	0.72	0.67	0.90	0.48	0.91	0.53	0.28	0.53	0.19	0.47	0.66	0.38	0.43	0.14	0.45	0.29	0.37	0.24	0.87
T3:11	$MP_{(20\%)}$ MAC values	1.00	0.99	0.89	1.00	0.90	0.98	0.92	1.00	0.57	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	0.97	1.00
T3:12	$MS_{(82\%)}$ MAC values	1.00	0.98	0.84	0.95	0.34	0.98	0.89	0.99	0.99	0.99	0.75	0.96	0.92	0.97	0.98	0.94	0.97	0.92	0.98

Note: Symmetric (S), antisymmetric (A), and torsional (T) mode labels are shown. MAC values were calculated between full data and MP and between full data and MS.

Before undertaking the missing-data STRIDE analyses, it was essential to establish baseline STRIDE results for the case of no missing data (i.e., full data). The analyses focused on structural modes of the Golden Gate Bridge under 1 Hz; observed acceleration data were filtered and downsampled accordingly. In a comprehensive study, Pakzad and Fenves (2009) determined 12 vertical and 7 torsional structural modes below 1 Hz using AR models with vertical accelerometer channels; these results are provided in the first and fifth rows of Table 3.

STRIDE ($p = 2, \theta = 5 \times 10^{-4}$) refers to a STRIDE analysis at model order 2 (absolute minimum model order for the state-space model) with a slope threshold of 5×10^{-4}. The initial state matrix A_0 and observation matrix C_0 were estimated by ERA-OKID-OO ($p = 2$). As before, the remaining state-space parameters were set in accordance with Matarazzo and Pakzad (2016). STRIDE converged to MLE at a maximum log-likelihood after 60 iterations. This full-data STRIDE analysis successfully identified 19 modes. The estimated frequencies and damping ratios are listed in the second and sixth rows alongside the AR results in Table 3 and are in accordance. There was some variation among damping estimates below the 8th mode; however, damping is known to vary between identification methods (Chang and Pakzad 2013); furthermore, it is difficult to verify the true value of damping ratios for real structures, although the results offer a notable performance for a minimum model-order identification algorithm.

Missing Packets (MP)

The communication reliability of a wireless sensor network was simulated in the current study using the 49-sensor data set discussed in the previous section under the assumption that during data collection the network would malfunction and packets would be dropped at random. To expedite the process of obtaining structural information from the available data, it is advantageous to process the data as is, with missing observations in time and space.

For the simulated data set, 20% of the observations were randomly removed throughout the response (partial observations as in Case 1 in the previous section), indicating significant packet loss. Here "removal" means replacing (imputing) the missing observations with zeros, consistent with Kalman filter substitution Eq. (18). STRIDE ($p = 2, \theta = 5 \times 10^{-4}$) was implemented with initial state matrix A_0 and observation matrix C_0 estimated using ERA-OKID-OO ($p = 2$). In the following section, Table 3, and Fig. 8(a), the results of this analysis are referred to as MP.

Results for Missing Packets

STRIDE converged to the ML estimates after 74 iterations. The 3rd and 7th rows of Table 3 show the frequency and damping estimates for all 19 modes alongside full data STRIDE and AR results. The frequency estimates compared very well with those of full-data STRIDE and were nearly identical to the AR results for all 19 modes.

The damping estimates were satisfactory and, in general, were lower than the full-data STRIDE estimates; however, because the true damping of the bridge was unavailable for comparison, it was not possible to verify either set of estimates.[If the reader is interested in the sensitivity of STRIDE modal estimates, the closed-form confidence intervals can be computed using formulations provided in Matarazzo and Pakzad (2015)].

Fig. 8(a) shows 4 selected MP mode shapes (1, 4, 14, and 19) superimposed on those from full-data STRIDE (full), along with their respective spline fits. Table 3 includes MAC values, which were computed between MP and full mode shape ordinates at all 49 nodes to quantify mode shape consistency in the case of data loss. Overall, the MP spatial results were quite consistent and accurate; the mean MAC value was 0.95.

Mobile Sensing (MS)

A mobile-sensor strategy was simulated using the 49-sensor data set described in the previous section. In this application, a group of 9 mobile sensors scanned 46 nodes on the west side of the main span and a single mobile sensor scanned 3 nodes on the east side. As illustrated in Fig. 7, the group of 9 mobile sensors covered 15 sensing nodes using equally spaced sensor clusters oriented in the following manner: 3 adjacent sensors (Nodes 1–3), 3 empty nodes (Nodes 4–6), 3 adjacent sensors (Nodes 7–9), 3 empty nodes (Nodes 10–12), and 3 adjacent sensors (Nodes 13–15). This sensor group scanned all 46 locations back and forth (like Case 3 in Section 5), simultaneously collecting information across time and space.

For the simulated data, the data set was repeated, effectively doubling the number of samples. The motivation behind the increased sample size was based on the assumption that a more accessible, convenient data collection method such as mobile sensing would promote frequent long-term monitoring (i.e., more data). Given the deterministic missing-data mechanism defined by the mobile-sensor setup, nearly 82% of the observations were removed from the response (dropped or not used in the simulated data set); the missing observations were represented by zeros. With this data, STRIDE ($p = 2$, $\theta = 5 \times 10^{-4}$) was implemented using initial state matrix A_0 and observation matrix C_0 estimated by ERA-OKID-OO ($p = 2$). In the following section, Table 3, and Fig. 8(b), the results of this analysis are referred to as MS.

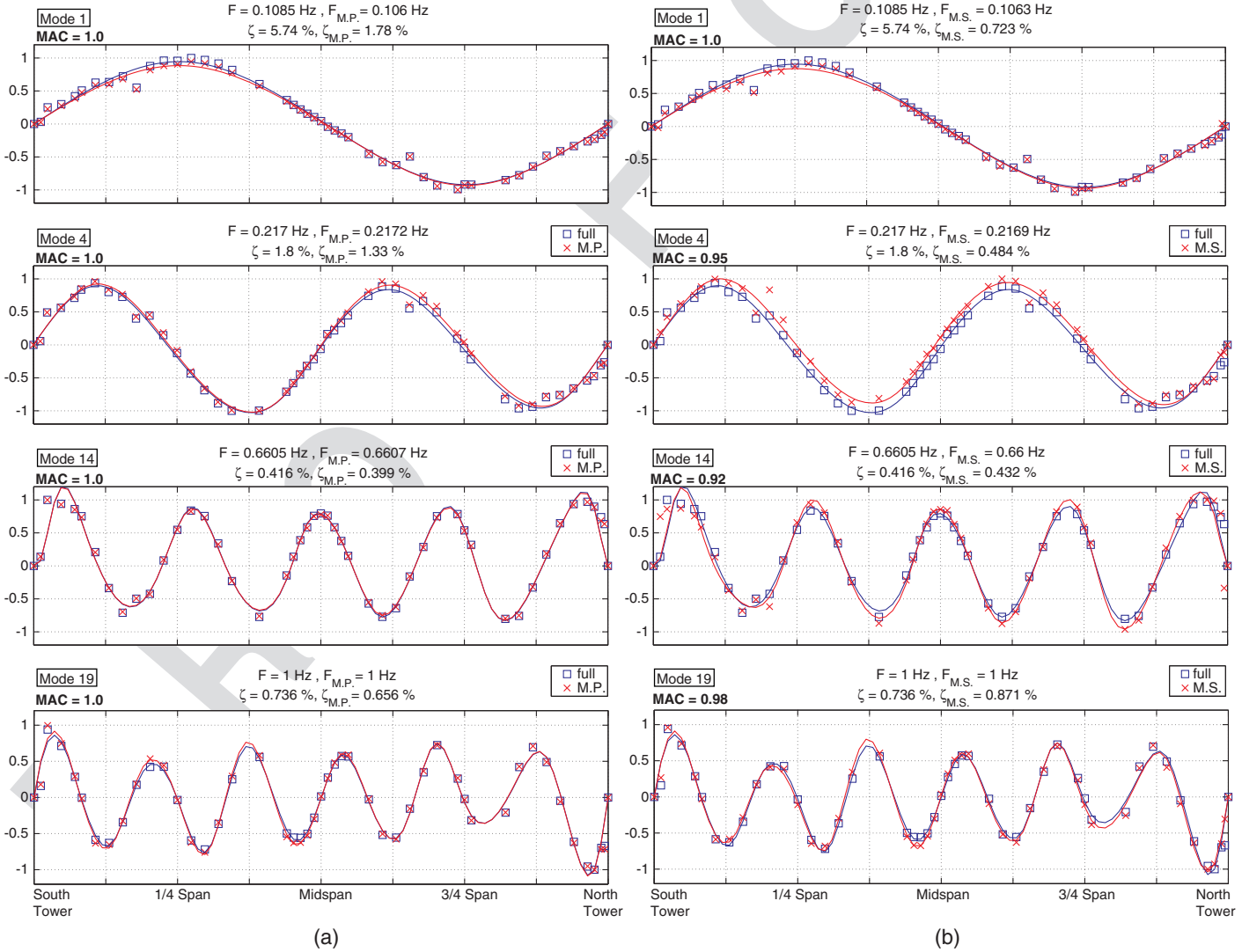


Fig. 8. (a) Superposition of west-side mode shapes 1, 4, 14, and 19; full-data (full) STRIDE and missing-packet (MP) data STRIDE with corresponding spline fits; MAC values computed using all 49 sensing locations; (b) superposition of west-side mode shapes 1, 4, 14, and 19; full-data (full) STRIDE and mobile-sensor (MS) data STRIDE with corresponding spline fits; MAC values computed using all 49 sensing locations

Results for Mobile Sensing

STRIDE converged to ML estimates after 152 iterations. The 4th and 8th rows in Table 3 show frequency and damping estimates for all 19 modes and full-data STRIDE and AR estimates. Like the MP results, the frequency estimates agreed with those of full-data STRIDE and matched the AR results for all 19 modes. The damping estimates were in accordance with the others in Table 3. As with the MP results, these values were generally lower than the full-data STRIDE estimates and MP but were nonetheless acceptable. Again, it is difficult to confirm true damping ratios in real structures.

Fig. 8(b) shows 4 MS mode shapes (1, 4, 14, and 19) superimposed over full-data STRIDE results and their respective spline fits. Overall, the spatial results were accurate and consistent with aforementioned analyses. MAC values were computed between MS and full-mode shape ordinates using all 49 nodes; the mean MAC value, over 19 modes was 0.91.

Perhaps the most propitious results of this analysis were found in mode shapes 15, 17, and 19 [mode shape 19 is shown at the bottom of Fig. 8(b)]. Because of the intricate geometry present in the higher modes, typically a high spatial resolution (achieved using many fixed sensors) is required to uniquely reconstruct such mode shapes based on the number of inflection points (N inflection points require at least $N+1$ fixed sensors). Given the modal coordinates at the supports, the number and location of fixed sensors implemented inherently limit the observed spatial content. For example, to uniquely reconstruct the first mode shape (Fig. 8, top) which contains 1 point of inflection, at least 2 fixed sensors would be required—ideally instrumented at one- and three-quarter span locations. Similarly, because mode shapes 15, 17, and 19 had 9 or more points of inflection, it was not possible to uniquely reconstruct these shapes from the 9 fixed sensors. However, the MS results show that 9 moving sensors could produce an excellent spatial resolution for high-order mode shapes comparable with 46 fixed sensors. MAC values of 0.97 and higher provided quantitative evidence that mobile sensors can preserve a high spatial resolution while using significantly fewer sensors.

Conclusions

In this paper, the STRIDE algorithm was reviewed and the necessary equations to process data sets with missing observations were presented. With these additions, STRIDE becomes the first modal identification method to formally accept data with missing observations, permitting analyses with mobile-sensor data and other incomplete data sets.

The convergence behavior and performance of STRIDE with missing data was investigated using the response of a 10-DOF shear structure. Three missing data mechanisms (Cases 1–3) were examined in which missingness varied from 0 to 40%. Cases 1 and 2 represented random missing-data mechanisms whereas Case 3 exemplified a mobile-sensor platform with systematic missing data corresponding to the sensors' movement. The results demonstrated that an accurate comprehensive modal analysis is possible when a large percentage (40%) of observations is missing. The accuracy of the results for Case 3 demonstrates the efficiency of mobile-sensor networks for SID in SHM. Furthermore, accurate modal estimates were obtained despite relatively low likelihood ratios. For example, normalized likelihood values below 50% in Cases 1 and 3 provided reliable modal estimates [Fig. 5(b)].

Two applications exemplified common STRIDE implementations using ambient vibration data collected at the Golden Gate Bridge. A missing packets (MP) application consisting of 20% missing data and a simulated mobile sensor (MS) network with 10 moving sensors resulted in 82% missing data. Both applications successfully identified 19 vibration modes (vertical and torsional) below 1 Hz with high accuracy in frequency and acceptable damping consistency using STRIDE at a minimum model order ($p=2$). As discussed in Chang and Pakzad (2013) and Matarazzo and Pakzad (2016), obtaining such results using alternate SID methods often requires multiple SID implementations at increasing model orders.

The mobile sensor results for mode shapes 15, 17, and 19 [mode shape 19 is shown in Fig. 8(b), bottom] were especially promising because they verified and quantified the preservation of spatial information in modal analyses for these types of sensor networks. In this application, 9 moving sensors on the west side of the main span accurately estimated high-order mode shapes that would otherwise be unattainable using only 9 fixed sensors. Furthermore, the spatial resolution in the estimated shapes was comparable with those of 49 fixed sensors, as quantified by MAC values of 0.97 and higher calculated for these shapes.

This paper demonstrated that STRIDE is capable of full, accurate modal analyses at low model orders ($p=2$ or 4) when observations are missing. The case studies and applications presented quantified the information reduction as a result of missing observations. The pattern of missing data remains an important factor in the accuracy of modal estimates; however, the successful applications discussed here indicate that valuable structural vibration information is available in SHM data even in circumstances of copious packet losses (limited sensor network reliability). Intentional missing-data mechanisms such as mobile sensing efficiently identified structural modal properties using significantly fewer sensors, encouraging real-world mobile sensor network implementations in future SHM projects.

Appendix I. Details of Eqs. (41)–(43)

$$E^{(*,j)}[\mathbf{y}_k] = \begin{cases} \underline{\mathbf{y}}_k^* & \text{if available} \\ E^{(*,j)}[\underline{C}^*\mathbf{x}_k + \boldsymbol{\upsilon}_k] = \underline{C}_j^*\hat{\mathbf{x}}_{k|K}^* & \text{if missing} \end{cases}$$

$$E^{(*,j)}[\mathbf{y}_k\mathbf{y}_k^T] = \begin{cases} \underline{\mathbf{y}}_k^*(\underline{\mathbf{y}}_k^*)^T & \text{if available} \\ E^{(*,j)}[(\underline{C}^*\mathbf{x}_k + \boldsymbol{\upsilon}_k)(\underline{C}^*\mathbf{x}_k + \boldsymbol{\upsilon}_k)^T] = \underline{C}_j^*\hat{\mathbf{x}}_{k|K}^*(\hat{\mathbf{x}}_{k|K}^*)^T(\underline{C}_j^*)^T + \underline{R}_j^* & \text{if missing} \end{cases}$$

$$E^{(*,j)}[\mathbf{y}_k\mathbf{x}_k^T] = \begin{cases} \underline{\mathbf{y}}_k^*(\hat{\mathbf{x}}_{k|K}^*)^T & \text{if available} \\ E^{(*,j)}[(\underline{C}^*\mathbf{x}_k + \boldsymbol{\upsilon}_k)\mathbf{x}_k^T] = \underline{C}_j^*\hat{\mathbf{x}}_{k|K}^*(\hat{\mathbf{x}}_{k|K}^*)^T & \text{if missing} \end{cases}$$

Appendix II. Supplementary Case Study Results

Table 4. Mean Frequency Estimates (Hz) for Cases 1 and 2, with Estimated Values for Case 3 (Representing Data from 9, 8, 7, and 6 Mobile Sensors, Respectively)

Mode	Exact	0%	5%	10%	15%	20%	25%	30%	35%	40%
Case 1: randomly removed observations										
1	0.886	0.892	0.899	0.900	0.898	0.896	0.897	0.897	0.895	0.896
2	2.638	2.647	2.651	2.654	2.652	2.653	2.653	2.652	2.651	2.651
3	4.324	4.320	4.324	4.324	4.315	4.301	4.295	4.296	4.297	4.291
4	5.903	5.897	5.897	5.902	5.893	5.885	5.877	5.869	5.867	5.868
5	7.338	7.371	7.376	7.377	7.375	7.373	7.382	7.381	7.378	7.372
6	8.591	8.590	8.594	8.587	8.581	8.575	8.567	8.566	8.566	8.570
7	9.634	9.628	9.623	9.625	9.627	9.634	9.636	9.639	9.643	9.653
8	10.438	10.360	10.345	10.352	10.358	10.366	10.362	10.361	10.385	10.391
9	10.986	10.985	10.980	10.962	10.974	10.964	10.959	10.970	10.976	10.972
10	11.264	11.221	11.201	11.181	11.191	11.193	11.199	11.210	11.203	11.216
Case 2: random time step blanks										
1	0.886	0.892	0.899	0.900	0.901	0.901	0.903	0.902	0.902	0.901
2	2.638	2.647	2.651	2.653	2.653	2.653	2.653	2.654	2.653	2.653
3	4.324	4.320	4.321	4.322	4.321	4.316	4.306	4.304	4.303	4.302
4	5.903	5.897	5.896	5.899	5.899	5.892	5.886	5.880	5.877	5.871
5	7.338	7.371	7.370	7.374	7.377	7.379	7.380	7.378	7.373	7.366
6	8.591	8.590	8.601	8.570	8.565	8.562	8.555	8.549	8.558	8.549
7	9.634	9.628	9.617	9.622	9.624	9.633	9.643	9.647	9.643	9.641
8	10.438	10.360	10.334	10.335	10.382	10.368	10.422	10.466	10.481	10.506
9	10.986	10.985	10.994	10.989	10.992	10.974	10.940	10.919	10.946	10.958
10	11.264	11.221	11.181	11.198	11.186	11.195	11.208	11.217	11.218	11.225
Case 3: mobile sensing										
1	0.886	0.892	—	0.896	—	0.891	—	0.892	—	0.891
2	2.638	2.647	—	2.646	—	2.651	—	2.648	—	2.648
3	4.324	4.320	—	4.326	—	4.302	—	4.314	—	4.326
4	5.903	5.897	—	5.891	—	5.893	—	5.898	—	5.885
5	7.338	7.371	—	7.372	—	7.391	—	7.401	—	7.380
6	8.591	8.590	—	8.600	—	8.623	—	8.595	—	8.588
7	9.634	9.628	—	9.633	—	9.635	—	9.625	—	9.652
8	10.438	10.360	—	10.391	—	10.400	—	10.412	—	10.452
9	10.986	10.985	—	10.991	—	11.013	—	11.004	—	10.959
10	11.264	11.221	—	11.198	—	11.224	—	11.216	—	11.130

Table 5. Variation in Frequency Estimates over 30 Simulations, with Coefficient of Variation (CoV) $\times 10^{-2}$ for Each Frequency for Increasing Magnitudes of Missing Data

Mode	5%	10%	15%	20%	25%	30%	35%	40%
Case 1: randomly removed observations								
1	0.157	0.153	0.181	0.239	0.294	0.259	0.397	0.560
2	0.048	0.047	0.056	0.037	0.022	0.029	0.021	0.034
3	0.132	0.172	0.191	0.090	0.058	0.101	0.115	0.159
4	0.106	0.114	0.078	0.106	0.094	0.078	0.108	0.180
5	0.068	0.049	0.151	0.261	0.052	0.070	0.078	0.207
6	0.074	0.084	0.083	0.093	0.091	0.115	0.194	0.217
7	0.059	0.077	0.066	0.067	0.095	0.112	0.168	0.273
8	0.373	0.429	0.197	0.305	0.273	0.287	0.447	0.367
9	0.194	0.188	0.250	0.272	0.248	0.254	0.243	0.198
10	0.605	0.132	0.194	0.094	0.100	0.204	0.228	0.336
Case 2: random time step blanks								
1	0.223	0.301	0.253	0.179	0.136	0.132	0.130	0.114
2	0.034	0.031	0.051	0.061	0.049	0.030	0.022	0.026
3	0.105	0.140	0.192	0.267	0.138	0.115	0.193	0.358
4	0.065	0.098	0.116	0.122	0.145	0.138	0.145	0.195
5	0.066	0.048	0.056	0.151	0.080	0.090	0.113	0.156
6	0.054	0.062	0.071	0.131	0.174	0.219	0.428	0.929
7	0.079	0.070	0.090	0.077	0.149	0.179	0.194	0.222
8	0.086	0.362	0.627	0.438	0.461	0.421	0.564	0.432
9	0.120	0.145	0.196	0.286	0.366	0.175	0.288	0.199
10	0.248	0.394	0.046	0.054	0.051	0.071	0.043	0.087

Note: Variation is not applicable to the deterministic missing data in Case 3.

Table 6. Mean Damping Values (%) for Cases 1 and 2, with Estimated Values from Case 3 (Representing Data from 9, 8, 7, and 6 Mobile Sensors, Respectively)

Mode	Exact	0%	5%	10%	15%	20%	25%	30%	35%	40%
				Case 1: randomly removed observations						
1	2.000	2.023	1.897	1.859	1.716	1.795	1.702	1.745	1.979	2.176
2	2.000	0.523	0.571	0.694	0.601	0.512	0.406	0.366	2.139	0.312
3	2.000	2.011	1.767	1.736	1.798	1.631	1.409	1.208	1.143	1.114
4	2.000	1.439	1.368	1.257	1.272	1.287	1.282	1.259	1.242	1.244
5	2.000	1.269	1.247	1.111	2.067	2.184	0.932	0.943	1.080	1.809
6	2.000	1.620	1.705	1.608	1.666	1.680	2.018	1.881	1.541	1.449
7	2.000	1.186	1.126	1.689	1.114	2.072	1.110	1.270	1.566	1.464
8	2.000	2.091	2.840	2.633	3.368	2.474	1.884	1.989	1.950	1.984
9	2.000	1.761	2.405	1.814	2.240	1.678	1.846	2.107	2.084	1.979
10	2.000	1.498	4.218	1.279	1.805	1.643	1.531	1.865	1.340	2.691
				Case 2: random time step blanks						
1	2.000	2.023	2.051	2.129	1.932	1.766	1.551	1.401	1.310	1.257
2	2.000	0.523	0.524	0.558	0.575	0.583	0.517	0.433	0.410	0.379
3	2.000	2.011	1.819	1.655	1.544	1.494	1.481	1.418	1.417	2.245
4	2.000	1.439	1.389	1.310	1.318	1.357	1.327	1.279	1.227	1.716
5	2.000	1.269	1.266	1.081	1.590	1.986	1.280	0.939	0.899	1.072
6	2.000	1.620	1.898	1.852	2.289	2.637	2.053	2.049	3.327	2.098
7	2.000	1.186	1.172	2.495	0.971	1.907	1.220	2.264	1.193	1.112
8	2.000	2.091	1.874	2.652	4.993	2.612	2.140	1.827	1.690	1.591
9	2.000	1.761	2.051	1.952	2.417	1.825	1.619	1.299	1.133	0.933
10	2.000	1.498	1.874	3.927	1.163	1.078	0.974	0.863	0.810	0.738
				Case 3: mobile sensing						
1	2.000	2.023	—	1.980	—	1.816	—	1.709	—	1.653
2	2.000	0.523	—	0.527	—	0.531	—	0.438	—	0.500
3	2.000	2.011	—	2.033	—	2.012	—	2.182	—	1.969
4	2.000	1.439	—	1.579	—	1.319	—	1.367	—	1.396
5	2.000	1.269	—	1.197	—	1.177	—	1.172	—	1.177
6	2.000	1.620	—	1.838	—	1.498	—	1.705	—	1.652
7	2.000	1.186	—	1.112	—	0.927	—	1.219	—	0.809
8	2.000	2.091	—	1.808	—	1.932	—	1.701	—	1.829
9	2.000	1.761	—	1.878	—	1.647	—	1.425	—	1.481
10	2.000	1.498	—	1.307	—	1.132	—	1.138	—	1.154

Table 7. Variation in Damping Estimates over 30 Simulations, with Coefficient of Variation (CoV) $\times 10^{-2}$ for Each Damping Ratio for Increasing Magnitudes of Missing Data

Mode	5%	10%	15%	20%	25%	30%	35%	40%
			Case 1: randomly removed observations					
1	6.096	6.876	7.449	10.937	12.801	15.247	26.575	80.793
2	4.852	9.743	8.432	6.322	12.417	17.157	460.391	23.751
3	8.600	7.410	7.262	6.062	10.047	9.894	12.417	17.267
4	4.157	3.544	4.580	8.358	4.750	6.588	13.542	12.921
5	4.801	4.759	187.526	170.194	5.242	8.377	11.592	127.538
6	4.853	6.551	7.669	6.308	122.305	103.116	14.705	12.817
7	6.776	184.747	8.546	172.394	6.658	61.179	98.716	46.055
8	131.524	134.428	130.646	97.513	11.328	45.283	37.895	74.067
9	103.529	6.076	123.828	12.227	74.415	97.605	92.840	89.859
10	129.557	5.848	188.917	172.644	161.425	142.114	132.924	146.068
			Case 2: random time step blanks					
1	8.065	14.413	8.603	9.683	9.323	6.571	8.492	17.604
2	3.790	6.218	9.663	7.718	10.876	11.565	8.222	6.662
3	6.858	6.244	6.684	6.692	9.424	8.936	11.746	120.647
4	4.711	4.045	4.503	6.764	7.332	7.910	12.359	167.922
5	4.528	6.051	192.194	193.567	154.496	10.418	8.893	18.135
6	3.756	5.731	142.743	152.610	146.475	113.508	101.167	82.608
7	9.499	190.801	8.141	142.020	38.597	95.300	51.970	51.297
8	3.724	141.278	105.033	88.014	62.463	49.173	23.570	19.939
9	6.252	7.493	88.783	9.046	11.778	14.517	19.540	9.952
10	136.074	136.143	4.436	5.761	6.593	5.623	8.758	9.508

Note: Variation is not applicable to the deterministic missing data in Case 3.

Table 8. Mean MAC Values for Cases 1 and 2, with Estimated MAC Values from Case 3 (Representing Data from 9, 8, 7, and 6 Mobile Sensors, Respectively)

Mode	0%	5%	10%	15%	20%	25%	30%	35%	40%
Case 1: randomly removed observations									
1	0.984	0.984	0.983	0.978	0.957	0.983	0.983	0.979	0.955
2	0.984	0.959	0.984	0.984	0.980	0.985	0.984	0.971	0.951
3	0.983	0.953	0.972	0.956	0.924	0.959	0.928	0.983	0.983
4	0.981	0.968	0.981	0.981	0.980	0.969	0.981	0.948	0.949
5	0.982	0.962	0.953	0.918	0.911	0.980	0.979	0.947	0.912
6	0.980	0.932	0.951	0.889	0.959	0.928	0.955	0.980	0.951
7	0.978	0.955	0.931	0.946	0.895	0.975	0.927	0.898	0.924
8	0.970	0.919	0.940	0.897	0.937	0.975	0.970	0.946	0.945
9	0.968	0.913	0.900	0.901	0.954	0.914	0.929	0.920	0.874
10	0.980	0.776	0.960	0.839	0.917	0.961	0.913	0.957	0.850
Case 2: random time step blanks									
1	0.984	0.984	0.984	0.984	0.984	0.952	0.978	0.984	0.984
2	0.984	0.984	0.984	0.982	0.979	0.982	0.984	0.984	0.967
3	0.983	0.983	0.982	0.983	0.983	0.983	0.984	0.984	0.984
4	0.981	0.953	0.981	0.981	0.960	0.965	0.969	0.959	0.968
5	0.982	0.982	0.964	0.950	0.931	0.934	0.927	0.961	0.959
6	0.980	0.981	0.960	0.952	0.911	0.917	0.908	0.786	0.941
7	0.978	0.929	0.849	0.935	0.888	0.883	0.722	0.914	0.914
8	0.970	0.973	0.933	0.777	0.906	0.959	0.966	0.957	0.957
9	0.968	0.975	0.977	0.956	0.972	0.966	0.963	0.961	0.965
10	0.980	0.893	0.741	0.943	0.986	0.988	0.988	0.986	0.986
Case 3: mobile sensing									
1	0.984	—	0.985	—	0.986	—	0.947	—	0.961
2	0.984	—	0.986	—	0.980	—	0.981	—	0.969
3	0.983	—	0.631	—	0.972	—	0.971	—	0.974
4	0.981	—	0.983	—	0.977	—	0.004	—	0.472
5	0.982	—	0.981	—	0.974	—	0.976	—	0.959
6	0.980	—	0.984	—	0.971	—	0.969	—	0.942
7	0.978	—	0.977	—	0.970	—	0.962	—	0.957
8	0.970	—	0.975	—	0.976	—	0.961	—	0.891
9	0.968	—	0.979	—	0.971	—	0.973	—	0.872
10	0.980	—	0.983	—	0.989	—	0.983	—	0.954

Table 9. Variation in MAC Values over 30 Simulations, with Coefficient of Variation (CoV) $\times 10^{-2}$ for Each MAC Value for Increasing Magnitudes of Missing Data

Mode	5%	10%	15%	20%	25%	30%	35%	40%
Case 1: randomly removed observations								
1	0.043	0.360	2.782	15.437	0.429	0.414	1.018	13.381
2	13.692	0.037	0.072	2.698	0.120	0.091	5.311	17.957
3	11.866	6.104	14.884	22.560	13.263	21.529	0.200	0.242
4	6.936	0.067	0.077	0.367	6.832	0.094	17.804	17.628
5	11.267	16.046	24.256	23.214	0.437	1.414	17.187	19.853
6	18.734	16.425	25.295	11.243	20.008	12.594	0.358	11.927
7	9.813	18.912	17.860	24.354	0.412	18.747	22.770	14.768
8	16.885	17.188	22.126	17.580	0.348	2.117	14.526	13.924
9	15.456	22.925	24.214	9.164	23.689	18.364	21.173	27.067
10	35.452	4.982	30.288	20.891	13.406	19.905	15.127	31.557
Case 2: random time step blanks								
1	0.021	0.033	0.324	0.042	15.961	3.651	0.042	0.028
2	0.019	0.019	1.481	3.049	1.487	0.034	0.034	9.552
3	0.044	0.054	0.070	0.065	0.070	0.055	0.057	0.077
4	15.259	0.035	0.048	11.601	6.072	6.888	10.169	4.727
5	0.049	10.510	17.581	19.301	17.967	19.948	10.415	11.811
6	0.081	11.656	15.285	23.306	19.252	23.068	35.588	17.952
7	13.468	30.191	17.137	24.976	25.710	35.934	12.383	15.000
8	0.213	18.044	34.928	23.199	6.908	0.714	1.471	0.860
9	0.516	0.487	13.691	1.442	0.965	1.131	1.254	1.240
10	23.580	38.374	16.418	0.553	0.437	0.372	0.387	0.365

Note: Variation is not applicable to deterministic Case 3.

Acknowledgments

Research funding for this work is partially provided by the National Science Foundation through Grant No. CMMI-1351537 through the Hazard Mitigation and Structural Engineering Program, and by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastructure Technology Alliance (PITA).

References

Allemang, R. J., and Brown, D. L. (1982). "A correlation coefficient for modal vector analysis." *Proc., 1st Int. Modal Analysis Conf.*, Vol. 1, Springer, Berlin, 110–116.

Andersen, P. (1997). "Identification of civil engineering structures using vector ARMA models." Doctoral dissertation.

Anderson, E., et al. (1999). *LAPACK users' guide*, 3rd edn., Society for Industrial and Applied Mathematics, Philadelphia, PA.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2008). *Time series analysis: Forecasting and control*, Wiley, Hoboken, NJ.

Cerda, F., et al. (2012). "Indirect structural health monitoring in bridges: Scale experiments." *Proc., 7th Int. Conf. on Bridge Maintenance, Safety and Management*, Lago Di Como, 346–353.

Chang, M., and Pakzad, S. N. (2012). "Modified natural excitation technique for stochastic modal identification." *J. Struct. Eng.*, 10.1061/(ASCE)ST.1943-541X.0000559, 1753–1762.

Chang, M., and Pakzad, S. N. (2013). "Observer Kalman filter identification for output-only systems using interactive structural modal identification tool suite (SMIT)." *J. Bridge Eng.*, 19(5), 1–11.

Dantu, K., Rahimi, M., Shah, H., Babel, S., Dhariwal, A., and Sukhatme, G. (2005). "Robomote: Enabling mobility in sensor networks." *Proc., 4th Int. Symp. on Information Processing in Sensor Networks*, IEEE Press, 55.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm." *J. Roy. Stat. Soc.*, 39(1), 1–38.

Digalakis, V., Rohlicek, J. R., and Ostendorf, M. (1993). "ML estimation of a stochastic linear system with the EM algorithm and its application to speech recognition." *IEEE Trans. Speech Audio Process.*, 1(4), 431–442.

Dunsmuir, W., and Robinson, P. M. (1981a). "Estimation of time series models in the estimation presence of missing data." *Am. Stat. Assoc.*, 76(375), 560–568.

Dunsmuir, W., and Robinson, P. M. (1981b). "Parametric estimators for stationary time series with missing observations." *Adv. Appl. Probab.*, 13(1), 129–146.

Fabien, T., Fischer, W., Caprari, G., Siegwart, R., and Moser, R. (2009). "Magnebike: A magnetic wheeled robot with high mobility for inspecting complex-shaped structures." *J. Field Robot.*, 26(5), 453–476.

Ghahramani, Z., and Hinton, G. E. (1996). "Parameter estimation for linear dynamical systems." *Technical Rep. CRG-TR-96-2*, Dept. of Computer Science, Univ. of Totronto, Toronto, Canada, 1–6.

Harvey, A. C., and Pierse, R. G. (1984). "Estimating missing observations in economic time series." *J. Am. Stat. Assoc.*, 79(385), 125–131.

Huang, M., and Dey, S. (2007). "Stability of Kalman filtering with Markovian packet losses." *Automatica*, 43(4), 598–607.

James, G. H., III, Carne, T. G., and Lauffer, J. P. (1993). "The natural excitation technique (NExT) for modal parameter extraction from operating wind turbines." *Technical Rep. No. SAND-92-1666*, Sandia National Laboratories, Albuquerque, NM.

Jones, R. H. (1962). "Spectral analysis with regularly missed observations." *Ann. Math. Stat.*, 33(2), 455–461.

Jones, R. H. (1971). "Spectrum estimation with missing observations." *Ann. Inst. Stat. Math.*, 23(1), 387–398.

Jones, R. H. (1980). "Maximum likelihood fitting of time ARMA models to time series with missing observations." *Technometrics*, 22(3), 389–395.

Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems." *Trans. ASME J. Basic Eng.*, 82(1), 35–45.

King, G. (1989). *Unifying political methodology: The likelihood theory of statistical inference*, University of Michigan Press, Ann Arbor, MI.

Kluge, S., Reif, K., and Brokate, M. (2010). "Stochastic stability of the extended Kalman filter with intermittent observations." *IEEE Trans. Autom. Control*, 55(2), 514–518.

Lin, C. W., and Yang, Y. B. (2005). "Use of a passing vehicle to scan the fundamental bridge frequencies: An experimental verification." *Eng. Struct.*, 27(13), 1865–1878.

Little, R. J. A., and Rubin, D. B. (2002). *Statistical analysis with missing data*, Wiley, Hoboken, NJ.

Lynch, J. P. (2007). "An overview of wireless structural health monitoring for civil structures." *Phil. Trans. Ser. A Math. Phys. Eng. Sci.*, 365(1851), 345–372.

Lynch, J. P., and Loh, K. J. (2006). "A summary review of wireless sensors and sensor networks for structural health monitoring." *Shock Vib. Digest*, 38(2), 91–128.

Matarazzo, T., and Pakzad, S. (2016). "STRIDE for structural identification using expectation maximization: Iterative output-only method for modal identification." *J. Eng. Mech.*, 10.1061/(ASCE)EM.1943-7889.0000951, 04015109.

Matarazzo, T. J., and Pakzad, S. N. (2013). "Mobile sensors in bridge health monitoring." *Structural Health Monitoring 2013: A Roadmap to Intelligent Structures: Proc., 9th Int. Workshop on Structural Health Monitoring*, DEStech Publications, 8.

Matarazzo, T. J., and Pakzad, S. N. (2014). "Modal identification of Golden Gate Bridge using pseudo mobile sensing data with STRIDE." *Dynamics of civil structures*, Vol. 4, Springer, Berlin, 293–298.

Matarazzo, T. J., and Pakzad, S. N. (2015). "Sensitivity metrics for maximum likelihood system identification." *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.*, B4015002.

Mendelssohn, R., and Roy, C. (1986). "Environmental influences on the French, Ivory-Coast, Senegalese and Moroccan tuna catches in the Gulf of Guinea." *Proc., ICCAT Conf. on the Int. Skipjack Year Program*, E. K. Symons, P. M. Miyake, and G. T. Sakagawa, ICCAT, Madrid, 170–188.

Mo, Y., and Sinopoli, B. (2011). "Kalman filtering with intermittent observations: Critical value for second order system." *Proc., 18th Int. Federation of Automatic Control*, International Federation of Automatic Control, 6592–6597.

Mo, Y., and Sinopoli, B. (2012). "Kalman filtering with intermittent observations: Tail distribution and critical value." *IEEE Trans. Autom. Control*, 57(3), 677–689.

Nagayama, T., Sim, S. H., Miyamori, Y., and Spencer, B. F. (2007). "Issues in structural health monitoring employing smart sensors." *Smart Struct. Syst.*, 3(3), 299–320.

Nigro, M. B., Pakzad, S. N., and Dorvash, S. (2014). "Localized structural damage detection: A change point analysis." *Comput. Aided Civ. Infrastruct. Eng.*, 29(6), 416–432.

Pakzad, S. N. (2010). "Development and deployment of large scale wireless sensor network on a long-span bridge." *Smart Struct. Syst.*, 6(5–6), 525–543.

Pakzad, S. N., and Fenves, G. L. (2009). "Statistical analysis of vibration modes of a suspension bridge using spatially dense wireless sensor network." *J. Struct. Eng.*, 10.1061/(ASCE)ST.1943-541X.0000033, 863–872.

Pakzad, S. N., Fenves, G. L., Kim, S., and Culler, D. E. (2008). "Design and implementation of scalable wireless sensor network for structural monitoring." *J. Infastruct. Syst.*, 10.1061/(ASCE)1076-0342(2008)14:1(89), 89–101.

Parzen, E. (1961). "Spectral analysis of asymptotically stationary time series." *Bulletin De L'institut International De Statistique*, 18.

Peeters, B., and De Roeck, G. (1999). "Reference-based stochastic subspace identification for output-only modal analysis." *Mech. Syst. Sig. Process.*, 13(6), 855–878.

Plarre, K., and Bullo, F. (2009). "On Kalman filtering for detectable systems with intermittent observations." *IEEE*, 54(2), 386–390.

Rauch, H. E., Striebel, C. T., and Tung, F. (1965). "Maximum likelihood estimates of linear dynamic systems." *AIAA J.*, 3(8), 1445–1450.

Rubin, D. B. (1976). "Inference and missing data." *Biometrika*, 63(3), 581–592.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*, Wiley, New York.
Shahidi, G., Nigro, M. B., Pakzad, S. N., and Pan, Y. (2015). "Structural damage detection and localisation using multivariate regression models and two-sample control statistics." *Struct. Infrastruct. Eng.*, 11(10), 1277–1293.
Shumway, R. H., and Stoffer, D. S. (1981). "An approach to time series smoothing and forecasting using the EM algorithm." *Technical Rep. No. 27*, Univ. of California, Davis.
Shumway, R. H., and Stoffer, D. S. (1982). "An approach to time series smoothing and forecasting using the EM algorithm." *J. Time Ser. Anal.*, 3(4), 253–264.
Shumway, R. H., and Stoffer, D. S. (2011). *Time series analysis and its applications with R examples*, Springer, New York.
Sibley, G. T., Rahimi, M. H., and Sukhatme, G. S. (2002). "Robomote: A tiny mobile robot platform for large-scale ad-hoc sensor networks." *Proc., 2002 IEEE Int. Conf. on Robotics and Automation (Cat. No. 02CH37292)*, IEEE, Piscataway, NJ, 1143–1148.
Singhvi, V., Krause, A., Guestrin, C., Jr, J. H. G., and Matthews, H. S. (2005). "Intelligent light control using sensor networks." *Proc., 3rd Int. Conf. on Embedded Networked Sensor Systems*, ACM, New York.
Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., and Sastry, S. S. (2004). "Kalman Filtering with Intermittent Observations." *IEEE Trans. Autom. Control*, 49(9), 1453–1464.
Stoffer, D. S. (1982). "Estimation of parameters in a linear dynamic system with missing observations." Univ. of California, Davis.
Stoffer, D. S. (1986). "Estimation and identification of space-time ARMAX models in the presence of missing data." *J. Am. Stat. Assoc.*, 81(395), 762–772.
Unnikrishnan, J., and Vetterli, M. (2012). "Sampling and reconstructing spatial fields using mobile sensors." *2012 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Berkeley, CA.
Van Overschee, P., and De Moor, B. (1992). "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems." *Automatica*, 30(1), 75–93.
Wu, C. F. J. (1983). "On the convergence properties of the EM algorithm." *Ann. Stat.*, 11(1), 95–103.
Xu, N., et al. (2004). "A wireless sensor network for structural monitoring." *Proc., 2nd Int. Conf. on Embedded Networked Sensor Systems*, ACM, New York, 13–24.
Zhao, J., and Ramesh, G. (2003). "Understanding packet delivery performance in dense wireless sensor networks." *SenSys '03: Proc. 1st International Conf. on Embedded Networked Sensor Systems*, ACM, New York, 1–13.
Zhu, D., Guo, J., Cho, C., Wang, Y., and Lee, K. (2012). "Wireless mobile sensor network for the system identification of a space frame bridge." *IEEE/ASME Trans. Mechatron.*, 17(3), 499–507.
Zhu, D., Yi, X., Wang, Y., Lee, K.-M., and Guo, J. (2010). "A mobile sensing system for structural health monitoring: design and validation." *Smart Mater. Struct.*, 19(5), 1–11.