EARTHQUAKE ENGINEERING AND STRUCTURAL DYNAMICS
Earthquake Engng Struct. Dyn. 2014; **00**:1–14
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/eqe

Data-driven post-earthquake rapid structural safety assessment

J.A. Goulet[1*], C. Michel[2] and A. Der Kiureghian[1]

[1]*Department of Civil and Environmental Engineering, University of California, Berkeley, USA*
[2]*Swiss Seismological Service, ETH Zurich, Switzerland*

SUMMARY

Earthquake prone cities are exposed to important societal and financial losses. An important part of these losses stems from the inability to use structures as shelters or for generating economic activity after the event of an earthquake. The inability to use structures is not only due to collapse or damage; it is also due to the lack of knowledge about their safety state, which prohibits their normal use. Because a diagnosis is required for thousands of structures, city-scale safety assessment requires solutions that are economically sustainable and scalable. Data-driven algorithms supported by sensing technologies have the potential to solve this challenge. Several ambient vibration monitoring studies of buildings, before and after earthquakes, have shown that the extent of damage in a building is correlated with a decrease in the natural frequency. However, the observed worldwide data may not be representative of specific cities due to factors such as construction type, quality, material, age, etc. In this paper we propose a framework that is able to progressively learn the relationship between frequency shift and damage state as a small number of buildings in a city are inspected after an earthquake, and to use that information to predict the safety state of uninspected but monitored buildings. The capacity of the proposed framework to learn and perform prognosis is validated by applying the methodology to a city with 1000 buildings having simulated frequency shifts and damage states.

Received . . .

KEY WORDS: Resilience; Earthquake; Safety; Statistical learning; Sensors; Condition assessment

1. INTRODUCTION

Earthquake prone cities are exposed to important societal and financial losses. A part of these losses are attributed to the inability to use structures as shelters or for generating economic activity after the event of an earthquake. The inability to use structures is not only due to collapse or damages; it is also due to the lack of knowledge about their safety state, which prohibits their normal use. When relying on qualified inspectors alone, inspecting structures at the scale of a city can take weeks, if not months (e.g. [1]). Not knowing the safety of buildings might force people out of their homes and companies out of business, further delaying the recovery and adding to the societal and economic impact of the earthquake. Several countries have developed procedures for visual inspection for post-earthquake assessment of buildings, e.g. the U.S. [2], Italy [3] and New Zealand [4]. The main challenge for inspectors is to state if a building can withstand aftershocks without performing a detailed seismic analysis.

Because a diagnosis is required for thousands of structures, city-scale safety assessment requires solutions that are economically sustainable and scalable. An alternative to using inspections alone is to employ sensors to monitor characteristic structural responses that are related to damage [5, 6]. In

*Correspondence to: J.A. Goulet, E-mail: james.a.goulet@gmail.com

the earthquake engineering community, the inter-story drift computed from strong-motion monitoring is a common damage parameter [6]. Researchers have proposed alternative approaches to post-earthquake safety assessment using either models-based approaches, e.g. [7–9], or vibration-based condition monitoring [5, 10, 11]. In this paper, in order to obtain an economically sustainable and scalable solution, we choose not to rely on models of the structures studied. Also, we use the shift in the natural frequency computed from ambient vibration measurements as the damage parameter. This choice is motivated by the increasing interest and ease in ambient vibration monitoring and the amount of existing data.

As early as in 1922, Omori [12] showed that the damage caused by an earthquake affected the natural frequencies of a building. Since then, several ambient vibration monitoring (AVM) studies of buildings, before and after an earthquake, have confirmed Omori's observation [13–17]. These experimental studies are further described in Section 2. When taken alone, existing data are not sufficient to support post-earthquake safety assessment applications where lives and billions of dollars worth of infrastructure are involved. This is in part because the relationship between the natural frequency shift and observed damage may not be representative of specific cities due to factors such as construction type, quality, material, age, etc. A solution to this challenge is to use statistical techniques capable of learning the relationship between damage and frequency shift, as structures in the city are inspected after an earthquake.

In this paper, we propose a framework that is able to learn and predict, at the scale of a city, the safety state of structures based on observed shifts in the frequencies and a limited number of inspections. Section 2 reviews a collection of experimental investigations reporting shifts in the frequencies of buildings before and after an earthquake and the corresponding identified damage states. Section 3 presents the statistical learning and prognosis methodologies in two parts. The first part consists in progressively learning, as inspections are performed, the conditional probability of observing a frequency shift given the damage state of a structure. The second part consists in using the learned relationship and observed frequency shifts to predict the safety state of the population of uninspected buildings. Section 4 describes how the proposed methodology can be integrated into a city-scale structural health monitoring (SHM) framework for post-earthquake rapid structural safety assessment. Finally, Section 5 uses simulated data to validate the capacity of the proposed method to learn and predict the safety state of buildings at the scale of a city.

2. COLLECTION AND CHARACTERIZATION OF EXISTING DATA

2.1. Ambient vibration monitoring and safety state classification

AVM consists in recording the natural vibrations of a structure, mostly due to ground ambient vibrations (or seismic noise) or wind (e.g. [18]). Assuming input vibrations are broad band, observed resonance peaks in the Fourier transform of the structural response correspond to its natural frequencies. These frequencies characterize the state of the structure, including those of the load bearing system and stiff non-structural elements, and the effect of soil-structure interaction. In the next subsection we review a collection of existing AVM experimental studies reporting the frequencies of buildings before and after an earthquake. Neither strong-motion recordings nor laboratory tests are selected, since the observed modal properties under these conditions can be substantially different from those recorded for real structures under ambient vibrations. We wish to relate the frequency shift to the observed damage in the corresponding buildings. One challenge is that existing studies did not all use a common metric for damage.

For loss estimation, it is common to use discrete or continuous damage scales that can be related to visual inspections. Hill and Rossetto [19] reviewed several damage scales, highlighting the advantages and disadvantages of each. For post-earthquake safety assessment, a 3-grade scale designated by color coding (green, yellow/orange, red) is generally used. In ATC-20 [2], they are defined as follows:

DOI: 10.1002/eqe

green: "*no restriction on use or occupancy*"

yellow: "*limited entry, entry is permitted only by the owner for emergency purposes, at his or her own risk*"

red: "*unsafe, the structure poses an obvious safety hazard*"

In this study, we employ the EMS98 [20] damage scale to conform with the scale used by most of the available datasets described in the following subsection. This scale has 5 damage indices (1 to 5) plus index 0 for no damage. Up to damage index 1, the building is classified as *safe*. Buildings with damage indices 2 and above are classified as *unsafe*; these buildings do not meet the conditions for normal occupancy. For the purpose of this paper, we also define a sub-class labeled *marginally unsafe* to designate unsafe buildings for which restricted entrance may be allowed; this subclass corresponds to damage indices 2 and 3. In terms of the color designations described above, an approximate mapping is *green* $= \{0, 1\}$, *yellow* $= \{2, 3\}$, and *red* $= \{4, 5\}$.

2.2. *Existing datasets*

The earliest data used is the ambient vibration experiment conducted by F. Omori on the Marunouchi building in Tokyo, Japan, before and after the 1922 earthquake [12]. This building was a steel-brick structure and suffered moderate damage (diagonal cracks in curtain walls) that is classified as damage index (d) 2.

Another building is the Millikan Library on the campus of the California Institute of Technology in Pasadena. Clinton et al. [13] reviewed tests performed on this building, which has experienced several earthquakes. The building is a shear wall reinforced concrete (RC) structure. During the 1970 San Fernando earthquake, slight structural damage was noted [13], corresponding to $d = 2$. The building was subsequently repaired. No evidence of damage was found after the 1987 Whittier Narrows, 1991 Sierra Madre, 1994 Northridge, and 2001 Beverly Hills earthquakes ($d = 0$ assumed), although slight frequency shifts were observed.

In Italy, Mucciarelli et al. [14] recorded ambient vibrations before and after a damaging aftershock of the 2002 Molise earthquake on a RC frame structure in Bonefro. The building was already damaged ($d = 3$) due to the mainshock and suffered additional damage ($d = 4$) during the recorded aftershock.

In Algeria, Dunand et al. [17] and Dunand [21] recorded ambient vibrations in several buildings after the 2003 Boumerdès earthquake. Since the frequencies of similar undamaged buildings were measured, they were considered by the authors to represent the states of the damaged buildings before the earthquake. Dunand et al. [17] reported the damage classification by the local inspectors following the color-coded damage scale presented in Section 2.1. These damage indices were transferred into the EMS98 scale using original pictures provided by Dunand. Six 5-story RC frame structures were moderately to heavily damaged, four 5-story RC shear wall buildings sustained lighter damage, and four 10-story RC shear wall buildings experienced moderate damage. The last three groups of buildings, as well as the building studied by Mucciarelli et al. [14], were already damaged when the earthquake occurred. Therefore, for these buildings, the frequency shifts before and after the earthquake are considered as lower-bound observations for the corresponding total frequency shifts.

The La Trinité Hospital in Martinique (French Lesser Antilles) was studied by Régnier et al. [16] before and after the 2007 moderately damaging earthquake. This structure was an irregular RC frame structure with 9 stories and suffered damage up to $d = 2$. Finally, the most extensive dataset comes from a recent study by Vidal et al. [15]. It consists in 34 RC frame buildings with 2 to 13 stories from Lorca, Spain with ambient vibration recordings performed before and after the 2010 earthquake. Damage indices ranging from 1 to 4 were estimated by Vidal et al. [15].

Figure 1 summarizes the results of the above studies. It shows the frequency shifts x expressed as the ratio of the post and pre-earthquake frequency of the building. Frequencies used are the larger of the two frequency shifts recorded in the longitudinal and transverse directional modes and the corresponding damage index for 45 buildings. Arrowheads indicate lower-bound observations of frequency shifts. Note that a lower-bound observation for a frequency shift corresponds to an upper-bound observation for x.

DOI: 10.1002/eqe

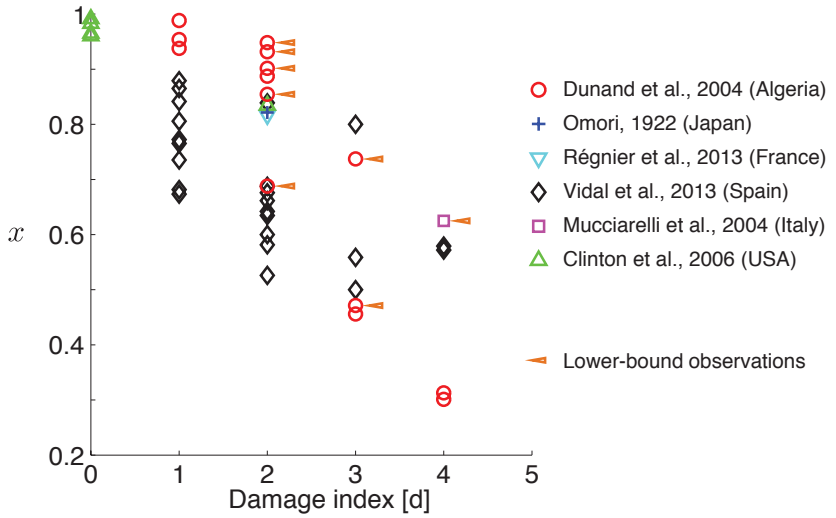


Figure 1. Collection of existing datasets reporting the corresponding frequency shifts x expressed as the ratio of the post and pre-earthquake frequency of the building, and damage index data obtained from AVM experiments and visual inspections. Arrowheads indicate lower-bound observations of frequency shifts.

A clear relationship between the frequency shift and damage index is noticeable within and across the datasets. This collection of data is used in Section 5 to characterize the conditional probability density function (PDF) of observing a frequency shift given the damage index. The PDF is subsequently used to simulate frequency shifts and damage indices for the population of buildings in a city. The simulated data is then used to validate the learning and prognosis capacity of the proposed methodology.

3. DATA-DRIVEN LEARNING AND PROGNOSIS

The methodology proposed in this paper is divided into two parts: First, using data for a collection of n buildings with known damage indices and frequency shifts, we learn about the conditional PDF of the frequency shift given the damage index. Second, for each uninspected building, we compute the probability that the building has a specific damage index given the observed frequency shift after an earthquake.

3.1. Data-driven learning

Let $D \in \{0, \cdots, 5\}$ denote a discrete random variable representing the damage index of a building after an earthquake. The domain of D is based on the EMS98 damage scale presented in §2.1. Also let $X \in (0, 1)$ denote a continuous random variable describing the ratio of the post and pre-earthquake frequencies. The range of possible values of X is determined by assuming that structural frequencies are strictly positive, and that damage can only decrease the frequency of a building. We consider D and X as dependent variables so that $f_{X|D}(x|d)$ is the conditional PDF of X given $D = d$ and $p_{D|X}(d|x)$ is the conditional probability mass function (PMF) of D given $X = x$. For the conditional PDF of X, we employ the Beta distribution,

$$f_{X|D}(x|d) = \frac{x^{a(d)-1}(1-x)^{b(d)-1}}{\mathrm{B}(a(d), b(d))} \tag{1}$$

where $\{a(d), b(d)\}$ are the conditional shape parameters to be estimated and $B(\cdot, \cdot)$ is the Beta function. This choice of distribution is motivated by the physical constraints on the frequency ratio X. The conditional mean, $\mu(d)$, and conditional standard deviation, $\sigma(d)$, of X are related to the

DOI: 10.1002/eqe

shape parameters by

$$\mu(d) = \frac{a(d)}{a(d)+b(d)} \tag{2}$$

and

$$\sigma(d) = \left(\frac{a(d)b(d)}{(a(d)+b(d))^2(a(d)+b(d)+1)}\right)^{1/2} \tag{3}$$

We learn the joint posterior PDF of $\mu(d)$ and $\sigma(d)$ by using data on n buildings for which the frequency shifts and corresponding damage indices, $(\hat{x}_j, d_j), j = 1, \cdots, n$, are observed. Considering that the true frequency shift for the j^{th} building is $X_j \in (0,1)$, we assume $X_j \sim f_{X_j}(x_j)$ has a Beta distribution with mean $\hat{x}_j$ and known standard deviation σ_j. The corresponding parameters of the distribution, a_j and b_j, are obtained in terms of $\hat{x}_j$ and σ_j by inverting relations similar to Eqs. 2 and 3. The damage indices d_j are assumed to consist in error-free observations. The joint posterior PDF of $\mu(d)$ and $\sigma(d)$ is computed from Bayes rule

$$f''(\mu(d),\sigma(d)) \propto \mathcal{L}(\mu(d),\sigma(d)) f'(\mu(d),\sigma(d)) \tag{4}$$

where $\mathcal{L}(\cdot,\cdot)$ is the likelihood function and $f'(\cdot,\cdot)$ is the prior distribution. The likelihood function $\mathcal{L}(\mu(d),\sigma(d))$ is computed form $\mathcal{L}(a(d),b(d))$ using Equations 2-3. The latter is the likelihood function for the parameters $\{a(d),b(d)\}$ and, assuming observations are statistically independent, is given by

$$\mathcal{L}(a(d),b(d)) \propto \prod_{j=1}^{n} \mathcal{L}_j(a(d_j),b(d_j)) \tag{5}$$

where $\mathcal{L}_j(a(d_j),b(d_j))$ is the likelihood for the j^{th} observation. When the frequency shift is directly observed, the likelihood is given by

$$\begin{aligned} \mathcal{L}_j(a(d_j),b(d_j)) &\propto \int_0^1 f_{X|D}(x|d_j) f_{X_j}(x)dx \\ &\propto \int_0^1 \frac{x^{a(d_j)-1}(1-x)^{b(d_j)-1}}{\mathrm{B}(a(d_j),b(d_j))} \cdot \frac{x^{a_j-1}(1-x)^{b_j-1}}{\mathrm{B}(a_j,b_j)} dx \end{aligned} \tag{6}$$

where $f_{X_j}(x)$ is the Beta PDF of X_j. When the observed value is a lower bound of the frequency shift, such as the observations marked by an arrowhead in Figure 1, the conditional PDF $f_{X|D}(x|d)$ in the above expression is replaced by the complement of corresponding conditional CDF, i,e. $1 - F_{X|D}(x|d)$.

We assume no prior information about parameters $\{\mu(d),\sigma(d)\}$ is available, except for the constraint that $\mu(0) > \mu(1) > \cdots > \mu(5)$. These constraints on the mean represent our domain-specific knowledge whereas we expect the mean frequency ratio to be decreasing with increasing values of d. Thus, using diffuse priors for $\mu(d)$ in the applicable domain and non-informative priors proportional to $1/\sigma(d)$ for $\sigma(d)$, the posterior distribution of all twelve parameters $\{\mu(d),\sigma(d), d = 0:5\}$, is given by

$$f''(\mu(0),\sigma(0),...,\mu(5),\sigma(5)) \propto \frac{\mathcal{L}(\mu(0),\sigma(0))}{\sigma(0)} \cdot \ldots \cdot \frac{\mathcal{L}(\mu(5),\sigma(5))}{\sigma(5)}, \quad \mu(0) > \cdots > \mu(5) \tag{7}$$

Note that $\sigma(d)$ terms are not coupled in the above equation. However, the $\mu(d)$ terms are coupled through the boundary of their applicable domain. To find the posterior distribution $f''(\mu(d),\sigma(d))$ for a particular d, we integrate Eq. (7) over the parameters $\{\mu(k),\sigma(k)\}$ for all $k \neq d$:

$$\begin{aligned} f''(\mu(d),\sigma(d)) \propto \tfrac{\mathcal{L}(\mu(d),\sigma(d))}{\sigma(d)} \quad & \prod_{k<d} \int_{\mu(k+1)}^{1} \int_0^\infty \frac{\mathcal{L}(\mu(k),\sigma(k))}{\sigma(k)} d\sigma(k) d\mu(k) \\ & \cdots \prod_{d<k} \int_0^{\mu(k-1)} \int_0^\infty \frac{\mathcal{L}(\mu(k),\sigma(k))}{\sigma(k)} d\sigma(k) d\mu(k) \end{aligned} \tag{8}$$

DOI: 10.1002/eqe

Let $F(\mu(d))$ denote the posterior marginal CDF of $\mu(d)$ if each $\mu(d)$ is assumed to have a diffuse prior over $(0, 1)$. This is equivalent to neglecting the constraint $\mu(0) > \mu(1) > \cdots > \mu(5)$, in which case the pairs of parameters $\{\mu(d), \sigma(d)\}$ are statistically independent for different d values so that,

$$F(\mu(d)) \propto \int_0^{\mu(d)} \int_0^{\infty} \frac{\mathcal{L}(\mu(d), \sigma(d))}{\sigma(d)} d\sigma(d) d\mu(d) \tag{9}$$

It follows that

$$f''(\mu(d), \sigma(d)) \propto \frac{\mathcal{L}(\mu(d), \sigma(d))}{\sigma(d)} \prod_{k<d} (1 - F(\mu(k))) \prod_{d<k} F(\mu(k)) \tag{10}$$

When $d = 0$, the second product term drops out, whereas when $d = 5$, the first product term drops out.

3.2. Data-driven prognosis

Suppose we compute from measurements the ratio of the post and pre-earthquake frequency X for a building after an earthquake. The updated distribution of D for the building is

$$p_{D|X}(d|x) = \frac{\tilde{f}_{X|D}(x|d) p_D(d)}{\tilde{f}_X(x)} = \frac{\tilde{f}_{X|D}(x|d) p_D(d)}{\sum_{d=0}^{5} \tilde{f}_{X|D}(x|d) p_D(d)} \tag{11}$$

where $p_D(d)$ is the prior PMF of d and $\tilde{f}_{X|D}(x|d)$ is the predictive conditional distribution of X given $D = d$, which incorporates the uncertainty in the estimation of the parameters $\mu(d)$ and $\sigma(d)$ and the latter is given by

$$\tilde{f}_{X|D}(x|d) = \int_{\mu(d)} \int_{\sigma(d)} f_{X|D}(x|d) f''(\mu(d), \sigma(d)) d\sigma(d) d\mu(d) \tag{12}$$

Let $S \subset \{0, \cdots, 5\}$ denote a subset of the damage indices. We have

$$\Pr(D \in S | X = x) = \sum_{d \in S} p_{D|X}(d|x) \tag{13}$$

For an uninspected structure, we label the damage state as belonging to set S when

$$\int_0^1 \Pr(D \in S|x) \cdot f_X(x) dx > \phi_{\text{class}} \tag{14}$$

Here, ϕ_{class} is a prescribed threshold probability. To select an optimal value for ϕ_{class}, one should define a utility function

$$U = g(\mathbf{Z}, \phi_{\text{class}}) \tag{15}$$

where $\mathbf{Z}$ is a set of variables defining performance metrics. For post-earthquake structural safety assessment, $\mathbf{Z}$ should contain metrics related to the rate of false positive and false negative classification, as well as a metric capturing the fraction of structures assessed as a function of time. The optimal value ϕ^*_{class} is computed by maximizing the expected utility,

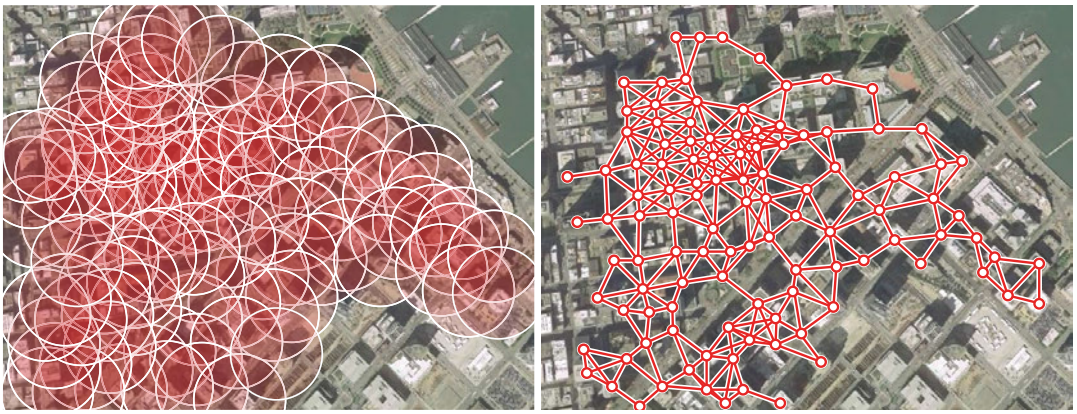
$$\phi^*_{\text{class}} = \arg\max_{\phi_{\text{class}}} E[g(\mathbf{Z}, \phi_{\text{class}})] \tag{16}$$

Equations 2-13 can be extended for cases where frequency shifts are monitored for multiple modes and for cases where other structural characteristic responses, such as inter-story drifts or tilts, are also recorded. In such a case, X is replaced by a vector of random variables $\mathbf{X}$. Furthermore, it is possible to condition X on features such as structural type, height, material or year of construction. In this paper, such a refined categorization is not used because of lack of data. But as data on characteristic structural responses versus damage index are gathered for different types of structures, more refined categorization will be possible. It is also noted that the posterior distribution of $\mu(d)$ and $\sigma(d)$ obtained for a city after an earthquake, can be used as the prior distribution for subsequent earthquakes affecting the city, including aftershocks. In this way, safety assessment can be made prior to conducting any inspections.

DOI: 10.1002/eqe

4. MONITORING FRAMEWORK

The application of the methodology presented in Section 3 requires knowing the natural frequencies for a population of buildings in near-real time after an earthquake. For each building, a minimum requirement is to have a two-axis accelerometer positioned on the rooftop. For the system to survive an earthquake, the power supply and communication channel of each sensor node should not rely on existing infrastructure. One solution would be to use solar panels with a small battery for supplying power, and to use a wireless local area network to enable communication between sensing nodes placed on different buildings. In recent applications, similar approaches have been used to install arrays of sensors on single structures [22–25]. In dense urban areas, using consumer-grade antennas with a range of 100 m would allow having a redundant communication network across a city. Only a few of the sensing nodes need to be connected to internet in order to send a continuous stream of data during the periods before and after an earthquake. In the event where internet connection is not available after an earthquake, the data for the entire population of instrumented buildings can be fetched at any time via a local connection to the wireless network. The wireless network can be accessed in the neighborhood of any instrumented building. Figure 2 illustrates an example of interconnectivity for the city center of San Francisco. In Figure 2a), each circle corresponds to the range of the antenna of a sensing node placed on the rooftop of a building, and links in Figure 2b) represent the interconnectivity of the sensing nodes.



(a) (b)

Figure 2. a) Range of each sensing node antenna, and b) Interconnectivity between sensing nodes.

Continuously monitoring the frequency of structures would allow knowing the frequency shift of buildings, minutes after the event of an earthquake. Then, hours after the event, teams of engineers could start inspecting buildings based on priorities established using observed frequency shifts. Once the knowledge of $p_{D|X}(d|x)$ is available, it becomes possible to perform prognoses for the safety of buildings that have not yet been inspected. In the event of a strong aftershock, it could become necessary to reassess the safety of the buildings. In such a case, the posterior distributions obtained from the main shock can be used immediately to perform prognoses for the entire population of buildings.

5. VALIDATION OF LEARNING AND PROGNOSIS CAPACITY

This section contains three parts. In §5.1 we use the methodology presented in Section 3 to estimate the posterior PDF of $\{\mu(d), \sigma(d)\}$ using the existing data reported in Section 2. In §5.2, we use $f_{X|D}(X|D)$ to generate a city-scale post-earthquake scenario for the purpose of illustrating the utilization and viability of the proposed data-driven safety assessment methodology. Finally, in §5.3

DOI: 10.1002/eqe

we estimate the expected performance of the framework with respect to its capacity to quickly and correctly classify uninspected structures as either safe or marginally unsafe/unsafe.

5.1. *Estimation of* $f''(\mu(d), \sigma(d))$

To estimate the posterior joint distribution of $\mu(d)$ and $\sigma(d)$ for the dataset reported in Section 2, we assume the standard deviation of X_j is $\sigma_j = 0.02$. In order to obtain valid parameter values for the Beta distribution, each observed frequency shift (the mean of the distribution) is assumed to lie within the range $(0.001\text{-}0.999)$. Theoretically, the standard deviation $\sigma(d) \in \mathbb{R}^+$. However, because of lack of data for buildings having a damage index $d = 5$, considering the entire domain of $\sigma(5)$ would lead to an improper PDF. The avoid this problem, we impose the constrain $\sigma(d) \leq 0.25$. Judging from the spread of data in Figure 1, the upper limit of 0.25 for $\sigma(d)$ appears reasonable.

Figure 3 shows contour plots of $f''(\mu(d), \sigma(d))$ computed using Equation 10. As expected $\mu(d)$

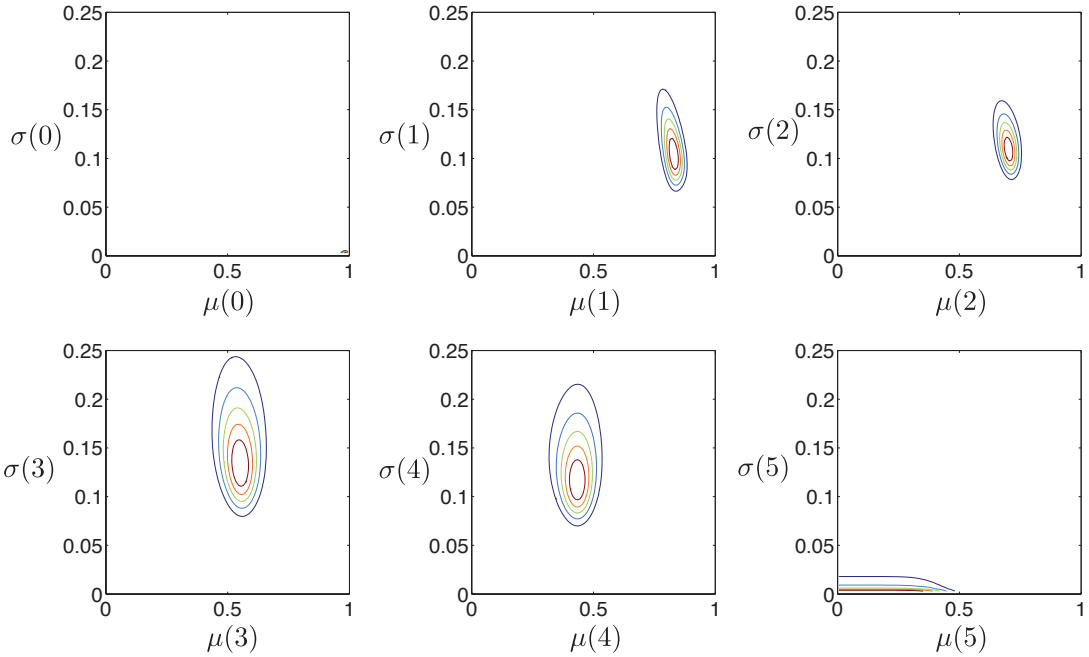


Figure 3. Contour plot of the joint PDFs $f''(\mu(d), \sigma(d))$ computed using Equation 10 for damage indices $d = \{0, 1, \cdots, 5\}$.

is decreasing with the damage index d. Note that despite not having data for $d = 5$, it is possible to obtain an estimate for $\{\mu(5), \sigma(5)\}$ using the knowledge that $\mu(d) < \mu(d-1)$. Without these constraints, the prior knowledge for a damage index could only be modified by data corresponding to that index. Figure 4 shows the predictive conditional PDF of X for each value of d. These predictive

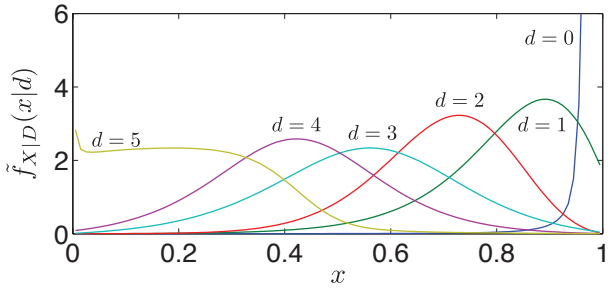


Figure 4. Predictive PDFs of the frequency shifts X given damage indices $d = \{0, 1, \cdots, 5\}$.

DOI: 10.1002/eqe

PDFs are in good agreement with the data presented in Figure 1.

5.2. *Example: Progressive learning and structural condition assessment*

Assume that for a population of 1000 buildings in a city, we know the frequency shifts right after an earthquake. As in the previous section, we assume the standard deviation of X_j is $\sigma_j = 0.02$. The set of damage indices for the population of buildings is simulated using the Binomial distribution

$$p(d) = \frac{5!}{d!(5-d)!} p^d (1-p)^{5-d}, d \in (0:5) \tag{17}$$

with $p = 0.2$. This distribution has a mean of $\mu_D = 1$ and standard deviation of $\sigma_D = 0.894$. Figure 5 shows the apportionment of damage indices among the simulated population of buildings. The number of buildings in each damage category is as follows: 317 in $d = 0$, 403 in $d = 1$, 214 in $d = 2$, 57 in $d = 3$, 8 in $d = 4$ and 1 in $d = 5$. The "observed" ratio of post and pre-earthquake frequency

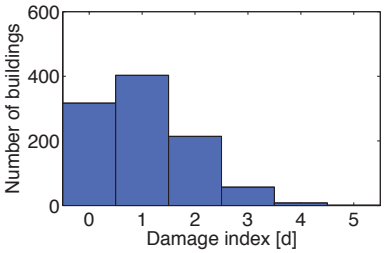


Figure 5. Apportionment of damage indices among the population of buildings.

for each building is obtained by simulating from the conditional distribution $f_{X|D}(x|d)$ using one realization of $f''(\mu(d), \sigma(d))$ of the parameters, as estimated in §5.1 based on the existing data. Figure 6 shows the simulated measured frequency shifts $\hat{x}_j$ and damage indices d_j for the simulated population of buildings. Our purpose is to classify buildings either as *safe*, *marginally unsafe* or

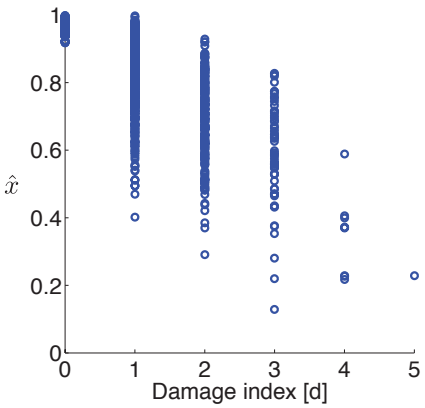


Figure 6. Simulated frequency shifts $\hat{x}_j$ and damage indices d_j for the population of buildings.

{*marginally unsafe, unsafe*}. That is, we wish to categorize buildings in damage classes $S_1 \in \{0, 1\}$ and $S_{2\&3} \in \{2:5\}$. However, we also explore the possibility of considering three damage classes: $S_1 \in \{0, 1\}$, $S_2 \in \{2, 3\}$ and $S_3 \in \{4, 5\}$. For the sake of simplicity, the combined damage state {*marginally unsafe/unsafe*} is hereafter denoted as the *combined unsafe* state.

We assume no prior knowledge about $\mu(d)$ and $\sigma(d)$, except for the constraint described in Subsection 3.1 and the assumed upper bound of 0.25 on $\sigma(d)$. We use the methodology presented in Section 3 to learn $f_{D|X}(d|x)$ as inspections are conducted and damage indices are determined. For the classification probability threshold, we assume $\phi_{\text{class}} = 1$ up to 10 inspections and $\phi_{\text{class}} = 0.97$

DOI: 10.1002/eqe

thereafter. We assumed that once a prognostic is made for a building, it can only be modified by an inspection of the building. This assumption is made so that the classification of buildings is not constantly changing as data is collected. The latency of $\phi_{\text{latency}} = 10$ inspections is introduced to avoid early misclassifications attributed to lack of knowledge. In order to speed up the assessment, buildings inspected are chosen so that their frequency shifts are evenly distributed and inspections are prioritized for buildings for which prognosis is not yet possible. The prior distribution $p_D(d)$ is assumed to be equiprobable for all d.

Figure 7a shows the evolution of the percentages of buildings in three categories as inspections are performed: (1) those whose damage states are identified through inspection (darkly shaded area), (2) those whose damage states are predicted according to the rule following Eq. (14) (lightly shaded area), and (3) the remaining buildings whose damage states cannot be identified with the current state of knowledge and the specified value of the threshold ϕ_{class} (white area). The height of the

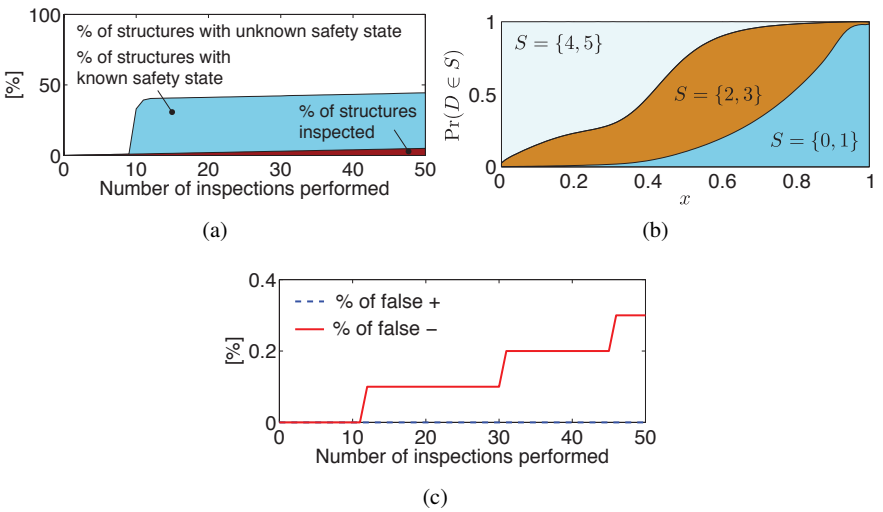


Figure 7. Prognosis of damage states: (a) Percent of buildings with identified damage states with increasing number of inspections. (b) Prognosis tool after 50 inspections. (c) Percent false positive and false negative classifications.

vertical axis represents the entire population of buildings and the height of each area represents the proportion of buildings classified as *inspected*, *not inspected but with known damage state*, or *not inspected with unknown damage state*. Without inspections, 100% of the buildings fall in the last class. Up to the inspection of the tenth building, only buildings that have been inspected have a known damage state. Then, having learned the distribution $f_{X|D}(x|d)$ based on the information gained from the 10 inspections, it is possible to make a prognosis for 320 additional buildings. This number grows with additional inspections. After collecting the results of 50 inspections, it is possible to predict the damage states of 396 buildings, of which 389 are in the safe state and 7 are in the combined unsafe state. Also, no unsafe buildings classified as safe (i.e. false positive) and 3 safe buildings classified as combined unsafe (i.e. false negative). Thus, the proposed data-driven approach is capable of classifying the damage states of buildings much faster than using inspections alone. However, beyond 50 inspections, the percentage of buildings whose damage states can be predicted does not increase significantly. It indicates that our capacity to preform a prognosis is limited by the *aleatory uncertainty* associated with the variability of frequency shifts for each damage state.

Note that the small number of unsafe buildings where a prognosis is possible is due two factors: First, the variability in the conditional distribution of frequency shifts is smaller for safer states. The second is the distribution of damage states that is strongly skewed towards safer buildings. With such a distribution of damage, there are fewer damaged buildings to learn from and fewer on which to perform a prognosis. When using a distribution of damage that is skewed toward unsafe buildings, the effect is reversed and a better prognosis can be made for these buildings. Finally, we note that in

DOI: 10.1002/eqe

this case, one way to address false negatives is to prioritize the inspection of buildings that have been identified as unsafe.

Figure 7b shows the prognosis tool derived from $f_{D|X}(d|x)$ after it is updated with the results of 50 inspections. For a given x, the height of each area represents the probability that the building is in damage classes $S_1 \in \{0,1\}$, $S_2 \in \{2,3\}$ and $S_3 \in \{4,5\}$. Here, we have made a distinction between the classes *marginally unsafe* and *unsafe*. With the assumed $\phi_{\text{class}} = 0.97$, buildings with $\hat{x} > 0.937$ are classified as *safe* and buildings with $\hat{x} < 0.370$ are classified in the *combined unsafe* state. The available information for damage class S_2 (*marginally unsafe*) is too vague to identify buildings in this category. (This is why we combined the marginally unsafe and unsafe classes.) One could use a smaller (larger) value for ϕ_{class} to allow a more (less) refined classification, but that would result in higher (lower) rates of false negatives and false positives. A better option would be to refine the prediction model $f_{X|D}(x|d)$ by developing it for different types of buildings, an effort that can be realized as post-earthquake monitoring data is gathered and analyzed.

Figure 7c shows the evolution of false positive and false negative classifications. A false positive is defined as classifying a building as safer than it is, e.g. classifying an *unsafe* building as *safe*. A false negative is the opposite: classifying a building as less safe than it is. False positives are more critical than false negatives because human casualties may result from occupying an unsafe building. False negatives only have economic or social impact. As Figure 7c shows, for the present example, both rates are small, but the rate of false positives is smaller than that of false negatives. Note that the illustrated example is just one realization of a stochastic process. In the following subsection we study the influence of the classification threshold ϕ_{class} and of the classification latency ϕ_{latency} (the number of inspections before prediction is performed) on the expected value of metrics quantifying the performance of the proposed framework.

5.3. Estimation of the expected performance

We evaluate the expected performance of the proposed framework by estimating, for the above example, the expected values of: (1) the percentage of buildings not yet inspected but with identified damage states, (2) the percentage of false positives, and (3) the percentage false negatives, after 50 simulated inspections. We estimate the expected values of the three metrics using 100 randomly generated populations of buildings and damage scenarios obtained following the procedure presented in §5.2. For each population of buildings, a different Binomial distribution of damage indices is used having parameters $n = 5$ and p, the latter generated from a Beta distribution with mean 0.2 and 20% coefficient of variation. Moreover, for each scenario, we use a different realization of the parameters of the conditional distribution $f_{X|D}(x|d)$ for generating frequency shifts given damage indices. The expected values of the three performance metrics are estimated for the classification threshold values $\phi_{\text{class}} = \{0.90, 0.92, 0.95, 0.97, 0.99\}$ and the latency threshold values $\phi_{\text{latency}} = \{0, 5, 10, 15, 20\}$.

Figure 8 shows the expected performance of the proposed methodology in terms of the three metrics as a function of parameters ϕ_{class} and ϕ_{latency}. Each point corresponds to the expected result obtained from 100 simulations and the meshed grid corresponds to the sample mean plus one standard deviation. The surface is a second-degree polynomial function fitted to the simulation results by minimizing the mean-square error. The results indicate that it is possible to control the safety assessment capacity and the rate of false classifications by adjusting the two parameters. Optimal values ϕ^*_{class} and ϕ^*_{latency} could be obtained from

$$\{\phi^*_{\text{class}}, \phi^*_{\text{latency}}\} = \underset{\{\phi_{\text{class}}, \phi_{\text{latency}}\}}{\arg\max} \; E[g(\mathbf{Z}, \phi_{\text{class}}, \phi_{\text{latency}})] \tag{18}$$

This optimization can be performed from a trivial search operation on the results presented in Figure 8.

Note that the low values of ϕ_{class} and ϕ_{latency} produce large percentages of structures assessed, however with unacceptable rates of false negatives. For more reasonable parameter values such as $\phi_{\text{class}} = 0.97$ and $\phi_{\text{latency}} = 10$, the expected percentage of structures assed after 50 inspections is 50% with an expected percentage of false positive and false negative respectively equal to 1.3% and 4.7%. Results indicate that the expected performance reaches levels sufficient to justify consideration of the proposed framework as a candidate for improving the resilience of earthquake-prone cities.

DOI: 10.1002/eqe

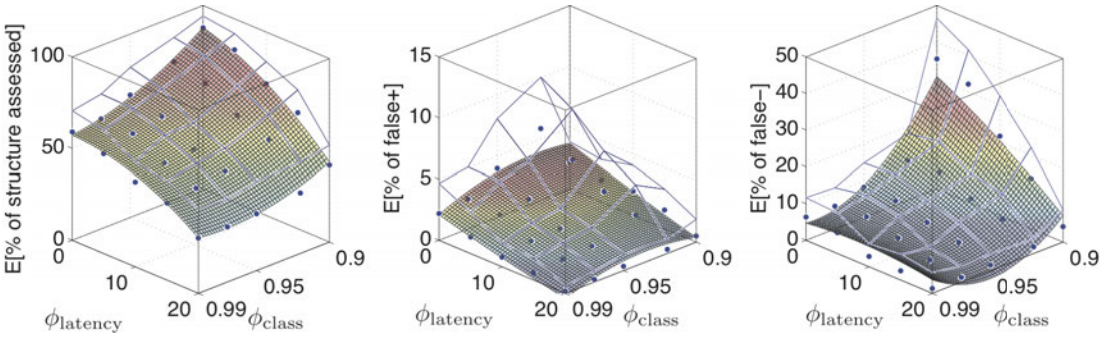


Figure 8. Expected performance as a function of classification threshold ϕ_{class} and classification latency ϕ_{latency}. Each point corresponds to the expected result obtained from 100 simulations and the meshed grid corresponds to the sample mean plus one standard deviation. The surface is a second-degree polynomial function fitted to the simulation results by minimizing the mean-square error.

6. DISCUSSION AND LIMITATIONS

With the framework proposed in this paper, we intend to take advantage of the frequency data collected on a large number of buildings across a city. Despite requiring a minimal number of sensors, such instrumentation can nowadays still be expensive because of the small amplitudes of building oscillations under ambient vibration, which determines the required sensitivity of the sensor. Notwithstanding recent developments in wireless sensor nodes for SHM applications [22–26], the cost of instrumenting one building is currently of the order of thousands of US dollars. However, recent advances in open-source hardware [27, 28] such as Raspberry Pi, Ardurinos and MinnoBoard, as well as in microelectromechanical-based sensor technologies [29, 30] allows us to expect significant cost reductions. As cost should decrease, public agencies in earthquake-prone areas might follow the example of California who requires the instrumentation of new buildings having 10 stories or more [31]. In addition to monitoring buildings for post-earthquake safety assessment, recording characteristic responses of structures during their service life would contribute to understanding their long-term behavior as well as guarantee their integrity in-between routine inspections. Moreover, it should be possible to improve the performance of the methodology by further categorizing buildings according to materials, usages and structural systems.

This paper focuses on the technical aspects allowing learning the relationship between post-earthquake frequency shifts and a damage index. The details of what occupants should do with the information provided by the system remain to be established in collaboration with emergency response authorities.

7. SUMMARY AND CONCLUSIONS

A collection of data available in the literature demonstrates the potential of using the frequency shift to infer the damage state of a structure. Using this information, this paper proposes a data-driven framework for predicting the damage states of structures in a city in the hours following an earthquake. In this framework, Bayesian updating allows learning, as inspection data is collected, the relationship between the shift in the structure frequency and the corresponding damage index obtained from each inspection. This information is then used to classify the damage states of uninspected structures, for which frequency shifts are measured by monitoring devices. The illustrative application shows that, when generating city-scale simulated scenarios using existing data, we expect to be able to predict the safety states of 50% of the monitored structures after having performed inspections only for 5%. The corresponding expected percentages of false positives and false negatives is respectively 1.3% and 4.7%. The proposed method leverages the information provided by visual inspections and

DOI: 10.1002/eqe

frequency shifts to speed up the safety assessment process, valorize the assessment performed by qualified personnel, and support the prioritization of building inspections.

ACKNOWLEDGEMENT

The authors thank Prof. Manuel Navarro and Dr. François Dunand for sharing part of the data used in this project. The first author thanks the Swiss National Science Foundation and the Quebec Research Fund for Nature and Technology for funding this research. Additional support was provided by the U.S. National Science Foundation under Grant No. CMMI-1130061.

REFERENCES

[1] Rapporto attività di sopralluogo effettuate al: 26/06/2009. *Technical Report*, Dipartemento della Protezione Civile, L'Aquila, Italy 2009.

[2] Procedures for Post-Earthquake Safety Evaluation of Buildings, ATC-20. *Technical Report*, Applied Technology Council (ATC), Redwood City, CA. 1989.

[3] Baggio C, Bernardini A, Colozza R, Corazza L, Bella MD, Pasquale GDI, Dolce M, Goretti A, Martinelli A, Orsini G, *et al.*. Field Manual for post-earthquake damage and safety assessment and short term countermeasures (AeDES). *Technical Report*, European Commission, Joint Reasearch Centre 2007.

[4] Marshall JD, Jaiswal K, Gould N, Turner F, Lizundia B, Barnes JC. Post-Earthquake Building Safety Inspection: Lessons from the Canterbury, New Zealand, Earthquakes. *Earthquake Spectra* Aug 2013; **29**(3):1091–1107, doi: 10.1193/1.4000151.

[5] Porter KA, Beck JL, Ching J, Mitrani-Reiser J, Miyamura M, Kusaka A, Kudo T, Ikkatai K, Hyodo Y. Real-time loss estimation for instrumented buildings. *Technical Report EERL 2004-08*, Earthquake Engineering Research Laboratory, California Institute of Technology, Kajima Corporation, Pasadena (California) 2004.

[6] Çelebi M, Sinclair ASM, Gallant S, Radulescu D. Real-Time Seismic Monitoring needs of a Building Owner and the Solution - A Cooperative Effort. *13 th World Conference on Earthquake Engineering*, vol. 3104, Vancouver, B.C., Canada, 2004.

[7] Mosquera V, Smyth AW, Betti R. Rapid evaluation and damage assessment of instrumented highway bridges. *Earthquake Engineering & Structural Dynamics* 2012; **41**(4):755–774.

[8] Iervolino I, Giorgio M, Chioccarelli E. Closed-form aftershock reliability of damage-cumulating elastic-perfectly-plastic systems. *Earthquake Engineering & Structural Dynamics* Sep 2013; **43**(4):613–625, doi:10.1002/eqe.2363.

[9] Wu S, Beck JL. Synergistic combination of systems for structural health monitoring and earthquake early warning for structural health prognosis and diagnosis. *SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring*, vol. 8348, International Society for Optics and Photonics, 2012; 83 481Z–83 481Z–10.

[10] Carden EP, Fanning P. Vibration Based Condition Monitoring: A Review. *Structural Health Monitoring* 2004; **3**(4):355–377, doi:10.1177/1475921704047500.

[11] Rodríguez R, Escobar JA, Gómez R. Damage detection in instrumented structures without baseline modal parameters. *Engineering Structures* Jun 2010; **32**(6):1715–1722, doi:10.1016/j.engstruct.2010.02.021.

[12] Omori F. The semi-destructive earthquake of April 26. *Seismological Notes (Imperial Earthquake Investigation Committee)* 1922; **3**:1–30.

[13] Clinton JF, Bradford SC, Heaton TH, Favela J. The observed wander of the natural frequencies in a structure. *Bulletin of the Seismological Society of America* 2006; **96**(1):237–257.

[14] Mucciarelli M, Masi A, Gallipoli MR, Harabaglia P, Vona M, Ponzo F, Dolce M. Analysis of RC Building Dynamic Response and Soil-Building Resonance Based on Data Recorded during a Damaging Earthquake. *Bulletin of the Seismological Society of America* 2004; **94**(5):1943–1953.

[15] Vidal F, Navarro M, Aranda C, Enomoto T. Changes in dynamic characteristics of lorca rc buildings from pre-and post-earthquake ambient vibration data. *Bulletin of Earthquake Engineering* 2013; :1–16.

[16] Régnier J, Michel C, Bertrand E, Guéguen P. Contribution of ambient vibration recordings (free-field and buildings) for post-seismic analysis: The case of the Mw 7.3 Martinique (French Lesser Antilles) earthquake, November 29, 2007. *Soil Dynamics and Earthquake Engineering* 2013; **50**:162–167.

[17] Dunand F, Meziane YA, Guéguen P, Chatelain JL, Guillier B, Salem RB, Hadid M, Hellel M, Kiboua A, Laouami N, *et al.*. Utilisation du bruit de fond pour l'analyse des dommages des bâtiments de Boumerdes suite au séisme du 21 mai 2003. *Mém. Serv. Géol. Alg.* 2004; **12**:177–191.

[18] Goulet JA, Michel C, Smith IF. Hybrid probabilities and error-domain structural identification using ambient vibration monitoring. *Mechanical Systems and Signal Processing* May 2013; **37**(1-2):199–212, doi:10.1016/j.ymssp.2012.05.017.

[19] Hill M, Rossetto T. Comparison of building damage scales and damage descriptions for use in earthquake loss modelling in Europe. *Bulletin of Earthquake Engineering* Jan 2008; **6**(2):335–365, doi:10.1007/s10518-007-9057-y.

[20] Grünthal G, Musson RMW, Schwartz J, Stucchi M. *European Macroseismic Scale 1998*, vol. 15. Cahiers du Centre Européen de Géodynamique et de Séismologie: Luxembourg, 1998.

[21] Dunand F. Pertinence du bruit de fond sismique pour la caractérisation dynamique et l'aide au diagnostic sismique des structures de génie civil. PhD Thesis, Université Joseph-Fourier-Grenoble I 2005.

[22] Kurata M, Kim J, Zhang Y, Lynch J, Van der Linden G, Jacob V, Thometz E, Hipley P, Sheng L. Long-term assessment of an autonomous wireless structural health monitoring system at the new Carquinez suspension bridge. *Proceedings of SPIE*, vol. 7983, 2011.

[23] Lynch J, Loh K. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest* 2006; **38**(2):91–130.

[24] Cho S, Yun CB, Lynch JP, Zimmerman AT, Spencer Jr BF, Nagayama T. Smart wireless sensor technology for structural health monitoring of civil structures. *International Journal of Steel Structures* 2008; **8**(4):267–275.

[25] Kurata N, Spencer BF, Ruiz-Sandoval M. Risk monitoring of buildings with wireless sensor networks. *Structural Control and Health Monitoring* 2005; **12**(3-4):315–327.

[26] Clayton RW, Heaton T, Chandy M, Krause A, Kohler M, Bunn J, Guy R, Olson M, Faulkner M, Cheng M, *et al.*. Community seismic network. *Annals of Geophysics* 2011; **54**(6):728–747.

[27] Pearce JM. Building research equipment with free, open-source hardware. *Science* 2012; **337**(6100):1303–1304.

[28] Pearce JM. *Open-Source Lab: How to Build Your Own Hardware and Reduce Research Costs*. Newnes, 2013.

[29] Milligan DJ, Homeijer B, Walmsley R. An ultra-low noise mems accelerometer for seismic imaging. *Sensors, 2011 IEEE*, IEEE, 2011; 1281–1284.

[30] D'Alessandro A, D'Anna G. Suitability of low-cost three-axis mems accelerometers in strong-motion seismology: Tests on the lis331dlh (iphone) accelerometer. *Bulletin of the Seismological Society of America* 2013; **103**(5):2906–2913.

[31] CBSC. *California Building Code, Appendix L*. California Building Standards Commision, Sacramento, CA 2013.