

Disaster Intensity-Based Selection of Training Samples for Remote Sensing Building Damage Classification

Luis Moya , Christian Geiß , *Member, IEEE*, Masakazu Hashimoto, Erick Mas ,
Shunichi Koshimura , and Günter Strunz

Abstract—Previous applications of machine learning in remote sensing for the identification of damaged buildings in the aftermath of a large-scale disaster have been successful. However, standard methods do not consider the complexity and costs of compiling a training data set after a large-scale disaster. In this article, we study disaster events in which the intensity can be modeled via numerical simulation and/or instrumentation. For such cases, two fully automatic procedures for the detection of severely damaged buildings are introduced. The fundamental assumption is that samples that are located in areas with low disaster intensity mainly represent nondamaged buildings. Furthermore, areas with moderate to strong disaster intensities likely contain damaged and nondamaged buildings. Under this assumption, a procedure that is based on the automatic selection of training samples for learning and calibrating the standard support vector machine classifier is utilized. The second procedure is based on the use of two regularization parameters to define the support vectors. These frameworks avoid the collection of labeled building samples via field surveys and/or visual inspection of optical images, which requires a significant amount of time. The performance of the proposed method is evaluated via application to three real cases: the 2011 Tohoku-Oki earthquake–tsunami, the 2016 Kumamoto earthquake, and the 2018 Okayama floods. The resulted accuracy ranges between 0.85 and 0.89, and thus, it shows that the result can be used for the rapid allocation of affected buildings.

Index Terms—Automatic labeling, building damage, multiregularization parameters, support vector machine (SVM).

I. INTRODUCTION

MACHINE learning has become a dominant data processing paradigm for the extraction of information from remote sensing data. The underlying strategy is to establish a model from limited but properly encoded prior knowledge (i.e., training samples) to assign a thematic label (e.g., a damage state in the application context of this article) to an instance under analysis (e.g., a building). Such methods are especially useful if explicit modeling based on, e.g., mechanical models, is too complex. At the same time, such approaches require both a sufficient amount of prior knowledge and viable descriptors to characterize the instances under analysis in order to achieve high predictive accuracy. However, many applications suffer from the unavailability of a sufficient number of training samples. In numerous cases, gathering training samples can become immensely expensive and time-consuming. Under these circumstances, various approaches have been proposed for alleviating the scarceness of training samples. For instance, multiobjective-based sparse representation classifiers [1], generation of virtual samples [2], and active learning methods [3]–[5] have been adopted in previous studies. Other approaches aim at compiling a training set in a fully automated manner from specified input data. The use of top-of-atmosphere reflectance to identify samples of forest areas [6] and fusion of multisource geodata [7] are such examples.

In the explicit application context of this work, studies that aim to extract natural hazard-induced damage levels of the built environment have extensively deployed machine learning algorithms in recent years. Supervised machine learning classifiers have shown high performance in terms of accuracy in disaster events, such as the 2010 Haiti earthquake [8]–[10], the 2011 Tohoku-Oki earthquake–tsunami [11]–[14], the 2016 Kumamoto earthquake [15], and the 2018 Sulawesi, Indonesia earthquake–tsunami [16], [17]. However, a careful reader may realize that the training data were provided after more than a month for the 2010 event [18], after four months for the 2011 event [19], after two months for the 2016 event [20], and after one week for the 2018 event [21]. Note that the

Manuscript received March 30, 2020; revised August 4, 2020 and December 4, 2020; accepted December 10, 2020. This work was supported in part by the National Fund for Scientific, Technological, and Technological Innovation Development (Fondecyt-Peru) within the framework of the “Project for the Improvement and Extension of the Services of the National System of Science, Technology and Technological Innovation” under Contract 038-2019, in part by the Japan Science and Technology Agency (JST) CREST Project under Grant JP-MJCR1411, in part by the Japan Society for the Promotion of Science (JSPS) Kakenhi under Grant 17H06108, in part by the Core Research Cluster of Disaster Science at Tohoku University (a Designated National University), and in part by the Helmholtz Association under Grant “pre_DICT” (PD-305). (*Corresponding author: Luis Moya.*)

Luis Moya is with the Japan-Peru Center for Earthquake Engineering Research and Disaster Mitigation (CISMID), National University of Engineering, Lima 15333, Peru, and also with the International Research Institute of Disaster Science (IRIDeS), Tohoku University, Sendai 980-8579, Japan (e-mail: lmoyah@uni.pe).

Christian Geiß and Günter Strunz are with the German Remote Sensing Data Center, German Aerospace Center (DLR), 82234 Weßling, Germany.

Masakazu Hashimoto, Erick Mas, and Shunichi Koshimura are with the International Research Institute of Disaster Science (IRIDeS), Tohoku University, Sendai 980-8579, Japan.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TGRS.2020.3046004>.

Digital Object Identifier 10.1109/TGRS.2020.3046004

last event's training data were sooner than others because it was based on visual interpretation of high-resolution optical imagery. However, it was later confirmed it contained several misclassifications [17], [22].

The necessary logistics for conducting a field survey directly after a major natural disaster and the subsequent digitization of the data are expensive and time-consuming. Furthermore, in most cases, avoidance of human exposure to hazardous areas is recommended. In the context of disaster mitigation, damage mapping is a race against the clock. The faster a satisfactory estimate is provided, the faster the first aid can be sent and the higher the chances that people who are trapped in collapsed buildings will survive [23]. As described earlier, a critical issue of the application of machine learning for damage mapping using remote sensing data is the lack of training data. Among the potential solutions is the development of a global network by building upon crowdsourcing for rapid damage assessment [24]. Another potential solution is to transfer training data that have been collected from one disaster event to another disaster event. To realize this objective, the database must be sufficiently large to consider various sensors, seasonal variations, various building types and infrastructural typologies, and heterogeneous types of disasters. Furthermore, not all disasters are recorded by remote sensing data, and training data are available for even fewer events. There are, however, recent studies for a specific type of disaster [25], [26].

There is another alternative for exploiting the experience that has been gathered from previous disasters. For decades, researchers have been collecting data to correlate the amount of damage with a metric of the disaster intensity, namely, the demand parameter at a specified location [27]–[32]. Intensity denotes the level of severity produced by a disaster in a particular location. The demand parameter refers to a quantitative measure of the intensity to which an asset is subjected. For the case of earthquakes, the peak ground acceleration (PGA), the peak ground velocity (PGV), the Modified Mercalli Intensity, and the spectral response are often used as demand parameters [28], [33], [34]. For tsunamis or floods, the inundation depth has commonly been used as a demand parameter because it can be measured from postdisaster field surveys [22], [31], [32], [35]. However, other demand parameters have been proposed [30], [36]. A fragility function, often idealized by a sigmoid function, is defined as the relationship between the probability that an asset reaches or exceeds a damaged state and the demand it experiences [28], [30], [36], [37]. Fragility functions, together with instrumentation and numerical simulations of the demand parameter, is often used to approximate, in real time, the number of assets that have been damaged within a specified area [38]–[42]. Recently, whether this aggregated damage information can replace training samples for the establishment of a damage map with a higher spatial resolution (building units instead of uniform spatial grids) from remote sensing data was investigated. A simple experiment from the 2011 Tohoku-Oki earthquake–tsunami is reported in [43], from which a linear discriminant function is calibrated over a bidimensional feature space via exhaustive search. The calibration implied to find a linear

discriminant function that yields a damage scenario that is consistent with the aggregates that are computed from the demand parameter and the fragility function. The accuracy of the results was on the same level as those of previous studies in which standard supervised machine learning was applied. Following the referenced study, a modification of the supervised logistic regression method was proposed in [44]. Here, the training data are replaced with probabilistic information that is computed from demand parameters and fragility functions. This enabled the use of an n -dimensional feature space and optimization methods to calibrate the discriminant function.

There is, however, a substantial pitfall in relying on fragility functions: fragility functions are available only for limited types of disasters, such as earthquakes and tsunamis. There is also a controversy regarding the transferability of fragility functions that have been constructed from empirical data, for instance, whether fragility functions for wooden buildings that were constructed in Japan can be used in other countries. To establish a solution that is independent of the availability of fragility functions, we uniquely deploy the estimated demand parameter directly after a hazardous event for automatic rule-based training sample selection. The spatial distribution of the affected buildings is expected to be consistent with the spatial distribution of the demand parameter. Namely, areas that are assigned a low demand parameter should contain mainly nondamaged buildings. In contrast, areas with a medium and large demand parameter likely include buildings with different damage levels. Using these assumptions, our objective is to learn a model that can solve a dichotomous classification problem and distinguish between two thematic classes: “severely damaged buildings” and “nonseverely damaged buildings.” The most common approach, termed change detection, aims to identify changes between a pair of images recorded before and after a disaster, from which changed samples are associated with severely damaged buildings and nonchanged samples are associated with nonseverely damaged buildings. It is assumed that, given that the images' recording time is close, the changes between the images are associated with the effects of the disaster. We provide two novel methods for calibrating a support-vector-machine-based discriminant function. As feature space, we use hand-engineered features computed from remote sensing data. The demand parameter is used to collect the training data automatically. Using automatic sample selection for change detection is not a new idea. Previous studies have first used unsupervised classification to collect reliable samples of changed and nonchanged samples and then improve the classifier using supervised/semisupervised classification algorithms [45], [46]. Unfortunately, such approaches to collect training samples do not provide a complete representation of the classes in the feature space. Furthermore, to the best of our knowledge, unsupervised techniques perform poorly when the disaster-affected area is much smaller than the area covered by the remote sensing data. Our contributions can be highlighted as follows.

- 1) The demand parameter allows reducing the search for changes to solely areas with medium/large demand parameters.

- 2) We use a threshold on the demand parameter to collect nonchanged samples. The demand parameter has a clear physical meaning, and thus, the selection of the demand threshold is very intuitive and does not require preliminary processing, such as unsupervised classification algorithms.
- 3) Because the demand parameter information is independent of remote sensing data, the collected nonchanged samples provide a better representation of the class nonchanged in the feature space.
- 4) We integrate information from in-place sensors (i.e., ground motion sensor, tidal gauges), numerical simulation of a natural phenomenon, and remote sensing.

The remainder of this article is organized as follows. Section II introduces the proposed approaches to calibrate a classifier model. Empirical evaluations are conducted in Section III. Additional comments regarding the proposed methods and the studied cases are provided in Section IV. Finally, the conclusions are drawn in Section V.

II. PROPOSED BUILDING DAMAGE DETECTION APPROACH

Consider a set of samples $\{(x_i, d_i)\}_{i=1}^M$, where each $x_i \in \mathbb{R}^n$ is a feature vector that was computed from remote sensing data for each building footprint and $d_i \in \mathbb{R}$ is a demand parameter for building i during the disaster event. The components of a feature vector are hand-engineered features associated with the degree of similarity between a pair of images. The main objective is to determine whether a building i has suffered damage via the identification of changes in images that were captured before and after the disaster. The proposed approach consists of two steps: 1) exploit the spatial distribution of the disaster intensity to collected training samples using a demand parameter threshold and 2) calibration of a discriminant function by exploiting the distribution of the training samples in the feature space. Two different strategies are reported in this step. In the following, the two steps are described in detail.

A. Demand Parameter Thresholding

The first step of the proposed approach for change detection aims to collect reliable training samples. Heuristic reasoning is used to divide the samples into two subsets. The key to the proposed approach is the prior assumption that it is expected for the ratio of severely damaged buildings to increase as the demand parameter increases. This assumption has been supported empirically by the studies on fragility functions discussed in Section I. Accordingly, it is possible to set a demand parameter threshold D such that the ratio of severely damaged buildings with $d_i \leq D$ is negligible. In contrast, samples with $d_i > D$ can be severely or nonseverely damaged buildings. For instance, during a tsunami, buildings affected by an inundation depth of less than, for example, 30 cm will be mostly nondamaged buildings (see Fig. 1). Let us define the corresponding subsets B_1 and B_{-1} as

$$\begin{aligned} B_1 &= \{x_i | d_i \leq D\} \\ B_{-1} &= \{x_i | d_i > D\}. \end{aligned} \quad (1)$$

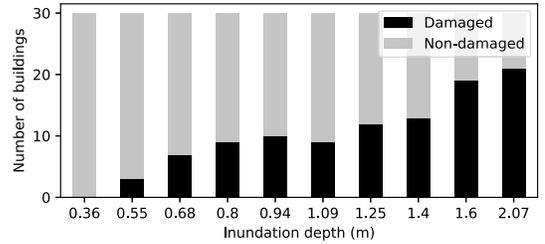


Fig. 1. Number of damaged and nondamaged buildings inspected in the coastal area of the city of Palu, Indonesia, during the 2018 Sulawesi-Indonesia tsunami. Note that buildings with an inundation depth of less than 36 cm were not damaged. The figure was modified from [22].

From the assumption mentioned earlier, B_1 can be used as the training data of nonchanged samples. Note that although samples with $d \leq D$ belong to the nonchanged class, it does not necessarily mean that samples with $d > D$ belong to the changed class. For instance, in the case of earthquakes, additional information, such as the strength, stiffness, and dynamic properties of the structure, is required to infer if a given building suffered damaged when experience a demand d . Thus, the set B_{-1} contains both changed and nonchanged samples.

In some cases, it is observed that $\#B_1 \gg \#B_{-1}$, where $\#$ denotes the cardinality of a set. In other cases, the opposite is observed, $\#B_1 \ll \#B_{-1}$. To avoid imbalance in the sizes of the data sets, either B_{-1} or B_1 is subsampled to the same size as the other. Thus, the following rule is adopted:

- 1) If $\#B_{-1} > \#B_1$, recall that the larger the demand parameter, the larger the percentage of damaged buildings (see Fig. 1). Consequently, in order to induce more changed samples in B_{-1} , samples with the largest demand parameter in B_{-1} are kept, and the rest are filtered.
- 2) If $\#B_{-1} < \#B_1$, then B_1 is randomly subsampled.

The relevance of the demand parameter in the first step is twofold. First, it provides criteria to select samples that are entirely independent of the sample distribution in the feature space. The second advantage of using the demand parameter for thresholding is the straightforward selection of D . The demand parameter has a clear physical meaning. In most cases, it is possible to select a very conservative value near zero, without compromising the distribution of the training data in the feature space. In some cases, it will require expertise in the disaster in question; however, it does not represent a pitfall if we consider that these experts are always involved in disaster management and support the decision-makers. Furthermore, we show later that small variations of D do not affect the accuracy.

B. Calibration of Discriminant Function

The second stage considers the problem of learning from labeled and unlabeled data. Given the label set $\mathcal{Z} = \{1, -1\}$, samples $x_i \in B_1$ are labeled as $z_i = 1$, whereas samples $x_i \in B_{-1}$ are unlabeled. The goal is to define a discriminant function that accurately separates changed ($z_i = -1$) from

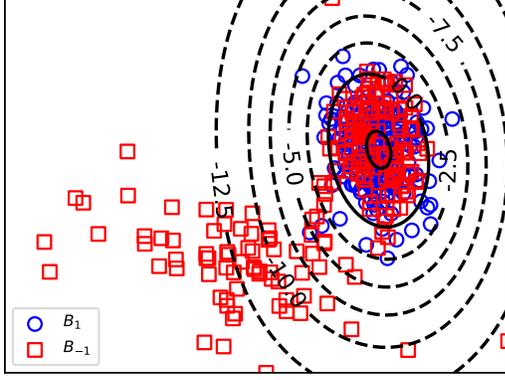


Fig. 2. Contour lines of the one-class SVM decision function, $f(x)$, at intervals of 1.5. The blue circle marks denote the samples from B_1 and were used to calibrate $f(x)$. The red square marks are samples from B_{-1} . The parameter $\nu = 0.1$ is used, that is, the closed boundary, $f(x) = 0$, contains at least 90% of samples from B_1 . The radial-basis function (RBF) kernel function with $\gamma = 0.1$ was used. The dashed isolines denote the negative values of $f(x)$.

unchanged ($z_i = 1$) samples. In the following, we report two procedures for realizing this objective.

1) Distance-Based Sample Selection (DSS) Approach:

An iterative approach to identify the elements of a subset $\hat{B}_{-1} \subseteq B_{-1}$ that can be used as training of changed samples is adopted. Then, with training samples from both classes, a standard classifier (here, SVM) for binary classification can be learned. It is expected that \hat{B}_{-1} should contain the patterns associated with the change class in the feature space and should not conflict with the patterns associated with the set B_1 . Accordingly, we propose to identify \hat{B}_{-1} based on the coordinates of its elements in the feature space with respect to the set B_1 . It is, thus, necessary to characterize B_1 in a functional form. There are several methods to define a closed boundary that contains B_1 [46]–[48]. Here, we use the support vector of high-dimensional distribution [48], which is referred to, in this study, as one-class SVM.

The one-class SVM consists of mapping B_1 to a feature space that belongs to a kernel and fitting a hyperplane to separate these samples from the origin with the maximum margin. To separate B_1 from the origin, the following quadratic problem needs to be solved:

$$\begin{aligned} \min_{w, \xi, \rho} & \left\{ \frac{1}{2} \|w\|^2 + \frac{1}{\nu l} \sum_i \xi_i - \rho \right\} \\ \text{s.t. } & (w \cdot \Phi(x_i)) \geq \rho - \xi_i, \quad \xi_i \geq 0 \end{aligned} \quad (2)$$

where ξ_i is a slack variable associated with a sample $x_i \in B_1$, $\nu \in (0, 1]$ is a parameter that denotes the upper bound on the fraction of outliers in B_1 if $\rho \neq 0$, $\Phi(x)$ is a function that maps x_i to the kernel feature space, $\mathbb{R}^n \rightarrow F$, $w \in F$ is a vector that is perpendicular to the hyperplane, and $\rho \in \mathbb{R}$ is a constant offset. The function that identifies B_1 in the feature space is expressed as

$$f(x_i) = (w \cdot \Phi(x_i)) - \rho \quad (3)$$

where $f(x)$ is positive for most of the samples from B_1 , the closed boundary is defined as $f(x) = 0$, and $f(x)$ is negative for samples outside the closed boundary (see Fig. 2).

The selection of samples for \hat{B}_{-1} is based on their distances to the hyperplane in the F -space calibrated via the one-class SVM [see ((2))]. Recall that sorting the samples according to their distances to the hyperplane is equivalent to sorting the samples by their corresponding values of $f(x)$ [see ((3))]. That is, sorting the samples that are located outside the closed boundary $f(x) = 0$ in ascending order of their values of $f(x)$ is equivalent to sorting the samples by their distances to the hyperplane in descending order. Therefore, \hat{B}_{-1} is defined such that, $\forall \hat{x}_i \in \hat{B}_{-1}$ and $\forall x_i \in B_{-1} \setminus \hat{B}_{-1}$, the following inequality holds:

$$f(\hat{x}_i) \leq f(x_i). \quad (4)$$

Simply, \hat{B}_{-1} contains the samples from B_{-1} with the largest distance to the closed boundary $f(x) = 0$. Note that $f(x)$ was already calibrated using the one-class SVM; thus, sorting samples from B_{-1} does not require significant computational effort. Using \hat{B}_{-1} and B_1 , the final discriminant function, $g(x) = (w \cdot \Phi(x_i)) + \rho$, is calibrated by solving the quadratic programming problem associated with the one-norm soft-margin SVM

$$\begin{aligned} \min_{w, \xi} & \left\{ \frac{1}{2} w \cdot w + \lambda \sum_{i|x_i \in B_1 \cup \hat{B}_{-1}} \xi_i \right\} \\ \text{s.t. } & z_i (w \cdot \Phi(x_i) + \rho) - 1 + \xi_i \geq 0 \\ & \xi_i \geq 0 \quad \forall i|x_i \in B_1 \cup \hat{B}_{-1} \end{aligned} \quad (5)$$

$$\xi_i \geq 0 \quad \forall i|x_i \in B_1 \cup \hat{B}_{-1} \quad (6)$$

where $z_i = 1$ if $x_i \in B_1$ and $z_i = -1$ if $x_i \in \hat{B}_{-1}$. The proposed method consists of determining $\# \hat{B}_{-1}$. An iterative strategy that gradually increases $\# \hat{B}_{-1}$ from an initial low value, for example, $\# \hat{B}_{-1} = 1$, is adopted. As result, the boundary decision function (i.e., $g(x) = 0$) is gradually approaching B_1 (see Fig. 3). When $\# \hat{B}_{-1}$ reaches $\# B_{-1}$, and because B_{-1} may contain several nonchange samples, it is very likely that, after calibration, the region in the feature space with $g(x) < 0$ (i.e., region where samples will be classified as changed) will overlap greatly the region where $f(x) > 0$ (i.e., the region where most samples from B_1 are located). From these considerations, it follows that the optimal value for $\# \hat{B}_{-1}$ should be lower than $\# B_{-1}$.

The application of the distance-based sample selection (DSS) approach requires adjustment of $\# \hat{B}_{-1}$, the regularization parameter λ , and the parameter associated with the kernel function γ . Recall that the realization of high accuracy on samples from B_1 is of high priority due to the assumption that most of them belong to the nondamaged class. B_{-1} is used as an auxiliary data set to improve the classifier. Thus, we evaluate the free parameters using the following expressions:

$$\max_{\gamma, \lambda, \# \hat{B}_{-1}} \{s\} \quad (7)$$

where

$$s = \frac{r \cdot R_{B_1} + R_{B_{-1}}}{r + 1} \quad (8)$$

where R_{B_1} denotes the ratio of samples classified as non-changed from set B_1 , $R_{B_{-1}}$ is the ratio of samples classified as changed from set B_{-1} , and $r > 1$ denotes a weight

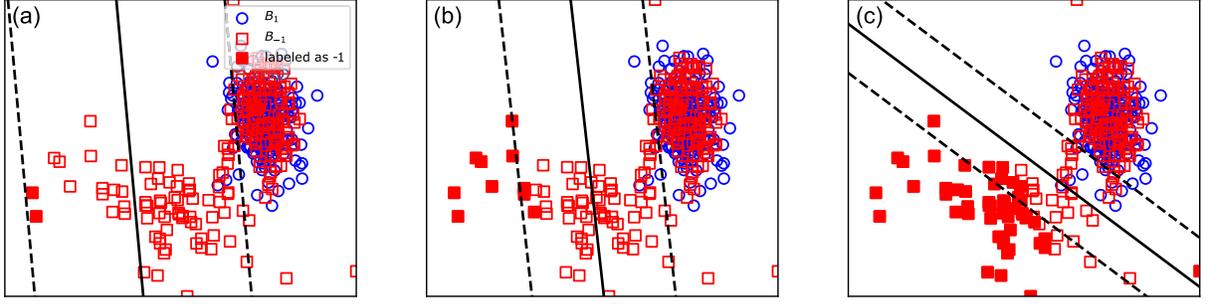


Fig. 3. Discriminant function (solid line) and margin bounds (dashed lines) at different stages of the DSS approach. Blue circles: samples from B_1 . Red squares: samples from B_{-1} ($\#B_1 = \#B_{-1} = 200$). The filled squares are the samples that belong to \hat{B}_{-1} when (a) $\#\hat{B}_{-1} = 2$, (b) $\#\hat{B}_{-1} = 10$, and (c) $\#\hat{B}_{-1} = 50$.

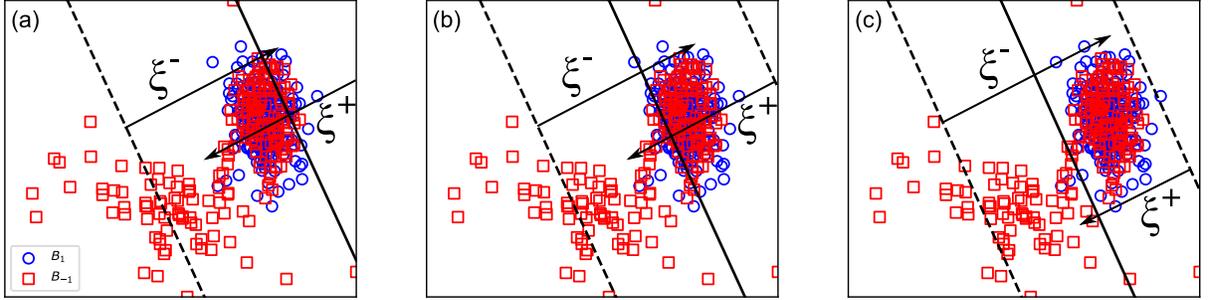


Fig. 4. MRP approach: effect of λ_p and λ_n in the calibration. Boundary decision function computed for (a) $\lambda_p/\lambda_n = 0.5$, (b) $\lambda_p/\lambda_n = 1.0$, and (c) $\lambda_p/\lambda_n = 2.0$. The solid line denotes the boundary decision function $h(x) = 0$. The dashed lines denote the bounds of the margin. The arrows denote the largest slack variable for B_1 (ζ^+) and B_{-1} (ζ^-). Note that the shortest value of ζ^- occurs when $\lambda_p/\lambda_n = 0.5$, whereas the shortest value for ζ^+ occurs when $\lambda_p/\lambda_n = 2.0$.

factor that is defined in advance by the user. Recall that this iterative strategy is only possible under the assumption that B_1 is a good representation of the nonchange class. Different approaches were adopted when the amount of training data is limited for all classes, or the information regarding the patterns of the classes in the feature space is incomplete [45], [49].

2) *Multiple Regularization Parameter (MRP) Approach:* The second alternative to calibrate the discriminant function $g(x)$ is based on a modification of (5) and (6) regarding how outlier samples are addressed. The second term of (5) is known as a regularization term, and it is the reason for the term soft margin in the method's name. It relaxes the assumption that the hyperplane, $g(x) = 0$, separates the two classes. The slack variable, ζ_i , is the amount of discrepancy when a support vector x_i does not lie in the bounds of the margin (i.e., $g(x) = 1$ for the set B_1 and $g(x) = -1$ for the set B_{-1}). For samples that are not support vectors, $\zeta = 0$. The slack variables allow misclassifications in the training data. The regularization term, $\lambda > 0$, is critical to the performance of the classifier. A small value will result in a flexible discriminant function that allows many outliers in the training data. A large value of λ will result in a rigid discriminant function with few outliers. The election of λ is a tradeoff issue between the generalization of the discriminant function and the accuracy of the training data. Advice regarding its selection can be found in [50].

For the problem that this study addresses, if we regard B_1 and B_{-1} as training samples for nonchanged and changed

classes, respectively, B_{-1} would be contaminated by many more outliers than B_1 . Therefore, a proper discriminant function should allow many outliers from B_{-1} while being stricter/more rigid with samples from B_1 . This desirable function can be obtained via the use of two regularization parameters, instead of one, in the quadratic programming problem

$$\min_{w, \zeta} \left\{ \frac{1}{2} |w|^2 + \lambda_p \sum_{i|x_i \in B_1} \zeta_i + \lambda_n \sum_{i|x_i \in B_{-1}} \zeta_i \right\} \quad (9)$$

$$\text{s.t. } z_i(w \cdot \Phi(x_i) + \rho) - 1 + \zeta_i \geq 0 \\ \zeta_i \geq 0 \quad (10)$$

where $z_i = 1$ for all $x_i \in B_1$ and $z_i = -1$ for all $x_i \in B_{-1}$. The regularization parameters λ_p and λ_n are associated with the samples from B_1 and B_{-1} , respectively. Consequently, the dual-problem of (9) and (10) is defined as follows:

$$\min_{\alpha} \left\{ \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j z_i z_j k(x_i, x_j) - \sum_i \alpha_i \right\} \quad (11)$$

$$\text{s.t. } 0 \leq \alpha_i \leq \lambda_p \text{ if } z_i = 1 \\ 0 \leq \alpha_i \leq \lambda_n \text{ if } z_i = -1 \quad (12)$$

where α_i denotes the Lagrange multipliers and $k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ is a kernel function. Fig. 4 depicts the decision boundary $g(x) = 0$ calibrated under different values for λ_p and λ_n . When $\lambda_n > \lambda_p$, the resulted decision boundary prioritizes accuracy in the set B_{-1} at the expense of errors in B_1 ,

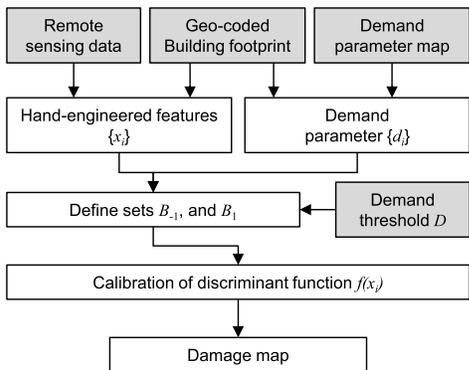


Fig. 5. Processing chain for damage mapping.

which is a nondesired situation. In order to have a decision boundary that prioritizes accuracy in the set B_1 , we expect that $\lambda_n < \lambda_p$. Recall that the idea of multiregularization parameters for change detection is not new. In [45], the samples that are located close to the decision boundary have different regularization parameters than those located far from the decision boundary. Unlike the work in [45], we do not impose a restriction on the potential value for λ_p and λ_n . However, we expect $\lambda_p \geq \lambda_n$, which is confirmed in the empirical evaluations.

Like the DSS method, the multiple regularization parameter (MRP) method has three free parameters that require adjustment: the regularization parameters λ_p and λ_n and the parameter associated with the kernel function Φ . These parameters will be evaluated from (8)

$$\max_{\gamma, \lambda_p, \lambda_n} \{s\}. \quad (13)$$

III. EMPIRICAL EVALUATIONS

We evaluate the performances of the proposed methodologies on three disaster events, each of which recorded with different types of remote sensing data. The block diagram in Fig. 5 illustrates the damage mapping processing chain that is going to be followed in every case study. The process requires four inputs: remote sensing data, a geocoded building inventory, a demand parameter map, and D . First, the remote sensing data and the geocoded building footprint are used to compute the hand-engineered features for each building, and the demand parameter map is used to extract the demand at each building sample. This approach focuses on built environment areas because hand-engineered features may exhibit patterns that differ among land use areas, such as agriculture targets [51]. Then, using D , the sets B_1 and B_{-1} are defined, and the discriminant function, $g(x_i)$, is calibrated either with the DSS or the MRP method. Finally, the discriminant function is used for binary classification of changed and nonchanged samples of the study area, where changed samples are associated with severely damaged buildings and nonchanged samples are associated with nonseverely damaged buildings. For all the experiments, the score [see ((8))] is computed with $r = 2$. Note that it is expected that the building footprint is available before the occurrence of a disaster, and

the demand parameter can be computed in real or near real time. Therefore, the implementation at an operational level lies in the availability of remote sensing data.

The accuracy of the DSS and MRP methods is computed with truth data collected by a third party. In addition, the results are contrasted with those obtained from the one-class SVM method and previous studies that used the same data sets.

A. First Case Study (CS1): The 2011 Tohoku Earthquake–Tsunami

The March 11, 2011, Tohoku earthquake, with Mw 9.0, is one of the largest well-recorded earthquakes ever. The maximum recorded strong-motion acceleration was 2.7 g, and the largest coseismic deformation that was recorded by Global Navigation Satellite System (GNSS) was 5 m. The earthquake triggered a tsunami, which caused extensive damage in the coastal area of Tohoku. The maximum measured tsunami height was 40 m. The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) [19] conducted field surveys and provided a building damage inventory, as shown in Fig. 6(a), for the coast of Miyagi prefecture. Seven levels of damage, which range from no damage to washed away, were defined in the survey. In this study, we focus on the buildings that are located within the inundated areas on the coast of Miyagi prefecture for which TerraSAR-X images were available.

1) *Demand Parameter*: Most of the damages were produced by the tsunami. Thus, the tsunami inundation depth is used here as the demand parameter. In this empirical evaluation, the actual inundation depth, which was provided by the MLIT, is used [see Fig. 6(b)].

2) *Feature Space*: The feature space was prepared from two microwave images of the coastal area of Miyagi prefecture, which were recorded by TerraSAR-X [see Fig. 6(c) and (d)]. The images were captured on October 12, 2010, and March 13, 2011, namely, before and after the earthquake–tsunami. As preprocessing, coregistration, radiometric calibration, speckle filter, and terrain correction were applied to the synthetic aperture radar (SAR) images. The resolution of the images is 1.25 m. Two features are computed from the imagery: the averaged difference in backscattering intensity, x_{i1} , and the correlation coefficient between the images, x_{i2} . The features were computed at the location of each building that was surveyed by the MLIT. Both features were computed with the pixels within a rectangular box that contains the building footprint [see Fig. 6(a)]. A total of 31262 samples constitute the data set. As preprocessing, the feature vectors were standardized such that, for any $j \in \{1, 2\}$, $(1/M) \sum_i x_{ij} = 0$, and $(1/M) \sum_i x_{ij}^2 = 1$.

3) *Results*: The subsets were defined using $D = 0.15$ m. From the MLIT’s inventory, 498 buildings experienced an inundation depth of less than or equal to 0.15 m, and 30764 buildings experienced an inundation depth that exceeded 0.15 m. Then, $\#B_1 = \#B_{-1} = 498$. Fig. 7 shows the sets B_1 and B_{-1} in the bidimensional feature space. The resulted predictions are shown in Fig. 8, from which nonchanged and changed samples are colored blue

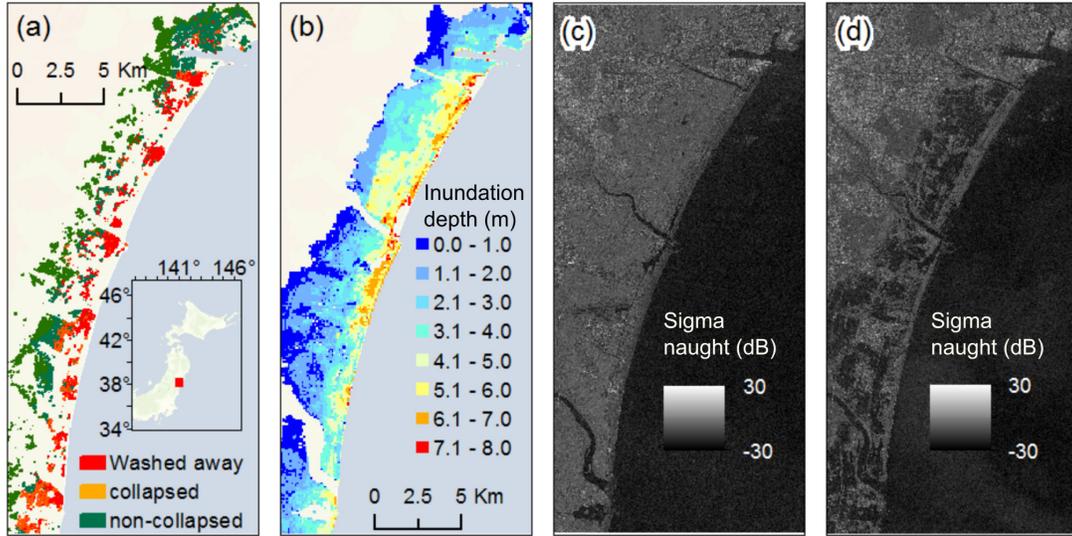


Fig. 6. CS1: the 2011 Tohoku-Oki earthquake. (a) Study area: coast of Miyagi prefecture, Japan. The mark denotes the building’s geolocation, and its color denotes the damage that was surveyed by the MLIT. The inset presents the location of the study area within northern Japan. (b) Tsunami inundation depth, as provided by MLIT. (c) TerraSAR-X intensity image that was recorded on October 21, 2010. (d) TerraSAR-X intensity image that was recorded on March 13, 2011.

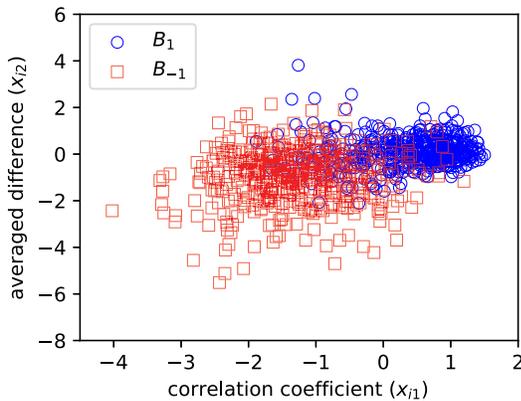


Fig. 7. Sample sets B_1 (blue marks) and B_{-1} (red marks) selected using $D = 0.15$ m for the CS1.

and red, respectively. The one-class SVM classifications were computed from B_1 using $\nu = 0.1$ and the RBF kernel function, $k(x_i, x_j) = \exp(\gamma \|x_i - x_j\|^2)$, with $\gamma = 0.1$. It is worth noting that samples with the largest correlation coefficient, x_{i1} , are classified as changed, which is a misclassification. The correlation coefficient is a hand-engineered feature associated with the degree of change between the pair of images. The largest value, before standardization, is one, which is associated with nonchange samples. However, in practical applications, a correlation coefficient of one is not frequent. Such samples appear to be outliers; thus, they might be incorrectly classified as damaged by the one-class SVM. Therefore, the use of a classifier calibrated solely from B_1 is not completely satisfactory.

A total of $\#B_{-1} = 374$ samples from B_{-1} were used as changed samples for the DSS approach. The regularization parameter was set to $\lambda = 0.05$, and the RBF kernel parameter was set to $\gamma = 0.07$. Regarding the MRP approach

[see Fig. 8(c)], the regularization parameters were set to $\lambda_p = 5.46$ and $\lambda_n = 3.36$. The Gaussian kernel parameter was set to $\gamma = 0.043$. Fig. 9 presents the largest values of the score s under fixed values of the kernel γ . The numbers that are specified at each mark are the optimal values of the other two parameters, namely, S and λ for the DSS method and λ_p and λ_n for the MRP method. When γ is less than approximately 1, s does not vary substantially. Regarding the DSS method, the percentage of samples from B_{-1} that is used to calibrate the discriminant function fluctuates between 75% and 95%. For the MRP method, for any arbitrary value of γ , the selected regularization parameters always satisfy the inequality $\lambda_p \geq \lambda_n$, namely, the classifier is tolerating more outliers from B_{-1} than from B_1 . Note that the same trend observed in Fig. 9 is also seen in the other experimental studies.

The damage level that was surveyed by MLIT and our predictions are compared in Table I. Most buildings that were labeled as noncollapsed were predicted as nonchanged. Likewise, most buildings that were labeled as washed-away were predicted as changed. Surprisingly, most of the samples that were labeled as collapsed were predicted as nonchanged. This controversy has been discussed in our previous studies [14], [43], [44]. The classification by MLIT was conducted in the context of the building’s structural system. Therefore, it is highly likely that by the time the postevent image was recorded, many of the buildings that were labeled as collapsed were still standing. Nevertheless, DSS and MRP clearly outperform the one-class SVM. In addition to the results of our study, the results that were reported in [44] are also included here as a baseline (BL) performance for our predictions. Recall that [44] used exactly the same data but with additional information from fragility curves.

The performances of the classifiers are reported in Table II, where the user accuracy (UA), producer accuracy (PA), and

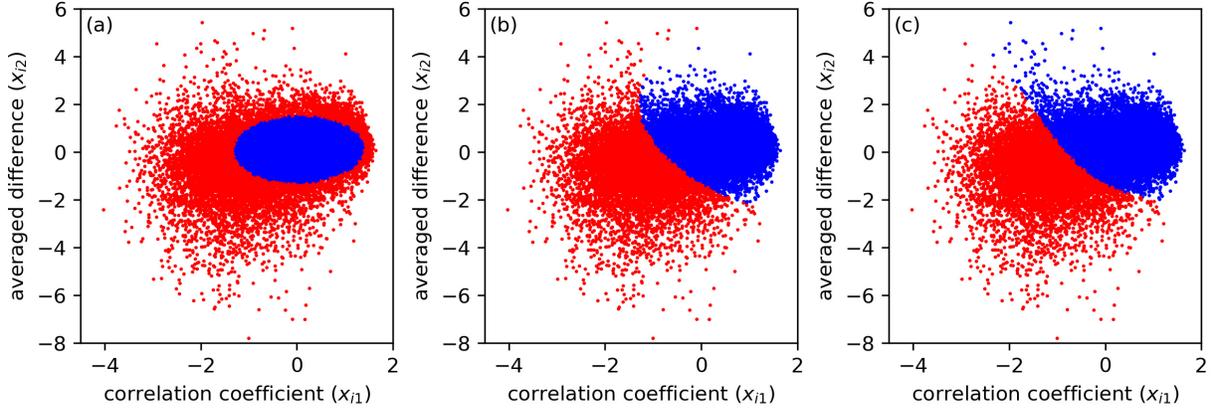


Fig. 8. Classification of changed (red marks) and nonchanged (blue marks) samples using (a) one-class SVM, (b) DSS, and (c) MRP methods for the CS1.

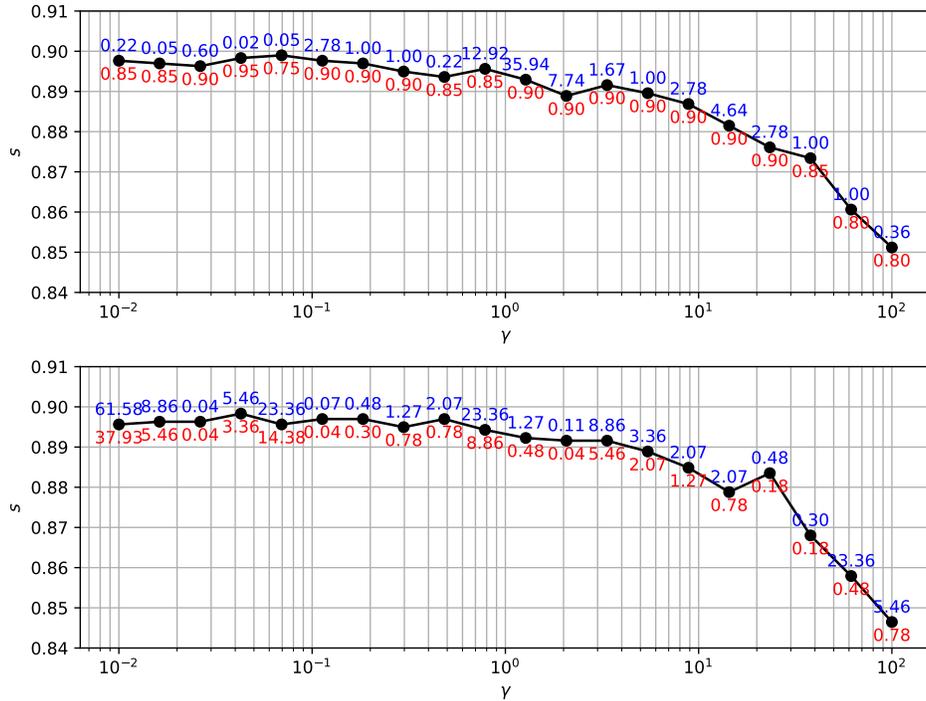


Fig. 9. Calibration of parameters for the CS1. Top: DSS method. The numbers in red and blue are the $\# \hat{B}_{-1}$ and λ values that yield the highest score under a fixed value of γ . The highest score is realized when $\# \hat{B}_{-1} = 0.75 \# B_{-1}$, $\lambda = 0.05$, and $\gamma = 0.07$. Bottom: MRP method. The numbers in blue and red are the values of λ_p and λ_n that yield the highest score under a fixed value of γ . The highest score is realized for $\lambda_p = 5.46$, $\lambda_n = 3.36$, and $\gamma = 0.04$. Note that, in all cases, $\lambda_p > \lambda_n$.

TABLE I

COMPARISON OF THE MLIT'S SURVEY WITH THE PREDICTIONS FROM THE ONE-CLASS SVM (1SVM), THE DSS APPROACH, AND THE MRP APPROACH AND THOSE THAT WERE PUBLISHED IN [44] (BL) FOR THE CS1. NC: NONCHANGE. C: DETECTED CHANGE

| | 1SVM | | DSS | | MRP | | BL | | Total |
|---------------|-------|------|-------|------|-------|------|-------|------|-------|
| | NC | C | NC | C | NC | C | NC | C | |
| Non-collapsed | 16520 | 2279 | 17816 | 983 | 17849 | 950 | 17348 | 1451 | 18799 |
| Collapsed | 2966 | 666 | 2977 | 655 | 2996 | 636 | 2738 | 894 | 3632 |
| Washed-away | 3230 | 5574 | 2249 | 6555 | 2319 | 6485 | 1612 | 7192 | 8804 |
| Total | 22716 | 8519 | 23042 | 8193 | 23164 | 8071 | 21698 | 9537 | 31235 |

F1 scores are presented. Here, the UA represents the percentage of samples that were surveyed as noncollapsed (washed-away) and were predicted as nonchanged (changed). Likewise, the PA is the percentage of samples that were predicted as

nonchanged (changed) and were surveyed as noncollapsed (washed away). The F1 score is computed via the following expression: $2(UA^{-1} + PA^{-1})^{-1}$. Again, DSS and MRP outperform the one-class SVM in terms of all the scores.

TABLE II
 UA, PA, AND F1 SCORE FOR THE PREDICTIONS FROM THE ONE-CLASS SVM (1SVM), THE DSS APPROACH,
 AND THE MRP APPROACH AND THOSE THAT WERE PUBLISHED IN [44] FOR THE CS1

| | UA | | | | PA | | | | F1 | | | |
|----|------|------|------|------|------|------|------|------|------|------|------|------|
| | 1SVM | DSS | MRP | BL | 1SVM | DSS | MRP | BL | 1SVM | DSS | MRP | BL |
| NC | 0.88 | 0.95 | 0.95 | 0.92 | 0.84 | 0.89 | 0.89 | 0.91 | 0.86 | 0.92 | 0.92 | 0.91 |
| C | 0.63 | 0.74 | 0.74 | 0.82 | 0.71 | 0.87 | 0.87 | 0.83 | 0.67 | 0.80 | 0.80 | 0.82 |
| Av | 0.76 | 0.85 | 0.85 | 0.87 | 0.78 | 0.88 | 0.88 | 0.87 | 0.77 | 0.86 | 0.86 | 0.87 |

Furthermore, the results from DSS and MRP demonstrate almost the same level of accuracy as those that were reported in [44].

B. Second Case Study (CS2): The 2016 Kumamoto Earthquake

On April 14, 2016, an Mw 6.2 earthquake struck the Kumamoto prefecture, Japan. Then, approximately 28 h later, another earthquake with Mw 7.0 occurred. The first event was designated as the foreshock and the second as the mainshock. In both events, the largest recorded shaking intensity was 7, which is the highest score in the Japanese Meteorological Agency (JMA) intensity rating system. Both events occurred in the town of Mashiki. Substantial damages to buildings, lifeline systems, and transportation infrastructure were reported [15], [53], [54].

1) *Demand Parameter*: In this case study, the demand parameter represents the PGV. Other potential demand parameters can be used for earthquake events, such as the PGA and the seismic intensity. However, field observations have shown that damages to buildings that were made of wood in Japan, which is the most common construction material, are well correlated with the PGV. The PGVs at the locations of the samples are provided by QuiQuake [see Fig. 10(c)], which is a platform that provides several demand parameter maps in near real time [52], [55]. QuiQuake uses the strong motion records from the K-NET and KiK-net networks to compute the PGA, the PGV, and the JMA seismic intensity scale. The K-NET and KiK-net together consist of approximately 2000 accelerometers that cover Japan's entire territory uniformly. Fig. 11 illustrates a scheme that is followed to compute the demand parameter maps. First, QuiQuake computes the demand parameter at each accelerometer's location. Then, using soil amplification factors, the demand parameter values at the bedrock level are estimated. Using an interpolation method, the demand parameter values are estimated approximately every 250 m within a grid format. The interpolation considers attenuation laws of the demand parameter with respect to its distance to the earthquake source. The interpolation is conducted at the bedrock to avoid the soil effects of the subsurface layers. Finally, the soil amplification factors are used again to estimate the demand parameter map at the ground surface.

2) *Feature Space*: Immediately after the foreshock, a mission to record the area with Lidar was conducted by the Asia Air Survey Co., Ltd. Later, a second mission was conducted seven days after the mainshock. The features were computed from two Lidar-based digital surface model (DSM) products of these missions [see Fig. 10(a) and (b)]. The resolution of

the DSMs is 50 cm. Using a geocoded building footprint [see Fig. 10(b)], three features were computed at each building: the average (x_{i1}) and the standard deviation (x_{i2}) of the differences between the DSMs and the correlation coefficient (x_{i3}) between the DSMs. A total of 27428 samples were extracted. The same features were used in [15]; thus, we will be able to use their results as a BL. Similar to the first case study, the feature spaces were standardized.

3) *Results*: A threshold of $D = 80$ cm/s was selected. Of the samples, 1012 experienced a demand that was lower than D and 26416 experienced a demand that exceeded D . Therefore, $\#B_1 = \#B_{-1} = 1012$ [see Fig. 12]. After calibration, the following values for the DSS approach are selected: $\gamma = 0.01$, $\hat{B}_{-1} = 810$ (80% of $\#B_{-1}$), and $\lambda = 1.67$. For the MRP method, the following values are selected: $\gamma = 0.18$, $\lambda_p = 5.46$, and $\lambda_n = 3.36$. Fig. 13 presents the resulted predictions. Table III presents our predictions on 903, from which its actual damage level is known [20]. Four levels of damage to buildings were reported in [20]: no damage (DL1), partial/moderate damage (DL2), inclined (DL3), and collapsed (DL4). It is observed that the changes that were detected in the predictions are more strongly associated with collapsed buildings. Moreover, the collapsed buildings are better predicted by DSS and MRP than by the one-class SVM method; however, part of this improvement is at the expense of lower accuracy in the prediction of the noncollapsed class. In [15], a standard SVM was used with the same surveyed samples to calibrate a discriminant plane. The prediction of the referenced study is also reported in Table III, from which the samples with damage levels from DL1 to DL3 were merged. These samples were used as training samples and not for testing. Thus, this prediction represents the best separation of samples that can be realized with a plane. Table IV reports the accuracy scores UA, PA, and F1. Here, the UA represents the percentage of samples that were surveyed as DL1-3 (DL4) and were predicted as nonchanged (changed). Similarly, the PA is the percentage of samples that were predicted as nonchanged (changed) and were surveyed as DL1-3 (DL4). These scores demonstrate that the DSS and MRP methods outperform the one-class SVM method.

C. Third Case Study (CS3): The 2018 Okayama Floods

Beginning from July 5, 2018, heavy rainfall occurred in western Japan. As of August 2, 2018, 220 casualties, nine missing persons, and 381 injuries had been reported due to the floods, mainly in Hiroshima and Okayama prefectures. Furthermore, 9663 houses were partially or completely collapsed, and 2579 were partially damaged [56].

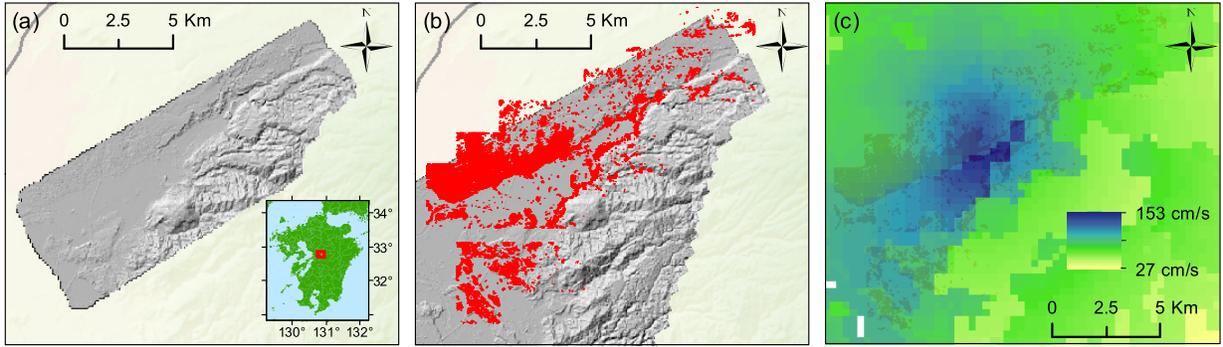


Fig. 10. CS2. (a) Lidar-based DSM that was recorded on April 23, 2016, over the study area. The inset shows the location of the study area within Kyushu Island, Japan. (b) Lidar-based DSM that was recorded on April 15, 2016, over the same area. The red marks denote the buildings' geolocations. (c) PGV that was provided by QuiQuake [52].

TABLE III

COMPARISON OF [20]'S SURVEY WITH THE PREDICTIONS FROM THE ONE-CLASS SVM (1SVM), THE DSS APPROACH, AND THE MRP APPROACH AND THOSE PUBLISHED IN [15] (BL) FOR THE CS2. NC: NONCHANGE. C: DETECTED CHANGE

| | 1SVM | | DSS | | MRP | | BL | | Total |
|-------|------|-----|-----|-----|-----|-----|-----|-----|-------|
| | NC | C | NC | C | NC | C | NC | C | |
| DL1 | 323 | 9 | 321 | 11 | 313 | 19 | | | 332 |
| DL2 | 216 | 5 | 208 | 13 | 205 | 16 | 674 | 24 | 221 |
| DL3 | 130 | 15 | 130 | 15 | 125 | 20 | | | 145 |
| DL4 | 62 | 143 | 48 | 157 | 41 | 164 | 39 | 166 | 205 |
| Total | 731 | 172 | 707 | 196 | 684 | 219 | 713 | 190 | 903 |

TABLE IV

UA, PA, AND F1 SCORE THAT WERE COMPUTED OVER THE PREDICTIONS FROM THE ONE-CLASS SVM (1SVM), THE DSS APPROACH, AND THE MRP APPROACH AND THOSE THAT WERE PUBLISHED IN [15] FOR THE CS2

| | UA | | | | PA | | | | F1 | | | |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|
| | 1SVM | DSS | MRP | BL | 1SVM | DSS | MRP | BL | 1SVM | DSS | MRP | BL |
| NC | 0.96 | 0.94 | 0.92 | 0.97 | 0.92 | 0.98 | 0.94 | 0.94 | 0.94 | 0.96 | 0.93 | 0.95 |
| C | 0.70 | 0.77 | 0.80 | 0.81 | 0.83 | 0.80 | 0.75 | 0.87 | 0.76 | 0.78 | 0.77 | 0.84 |
| Av. | 0.83 | 0.87 | 0.86 | 0.89 | 0.88 | 0.89 | 0.85 | 0.91 | 0.85 | 0.87 | 0.85 | 0.90 |

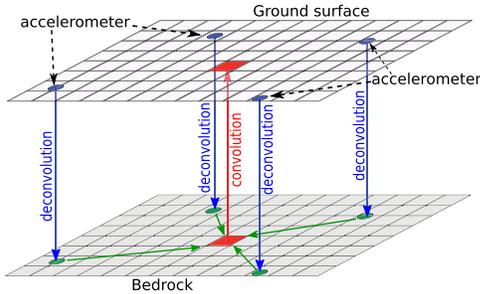


Fig. 11. Schematic of the estimation of the demand on a uniform grid using the Japanese strong motion network. It consists of three main steps. First, the demand parameter is computed at the accelerometer's station, and its value at the base rock is estimated via a deconvolution process (blue arrows). Second, the demand parameter is estimated at an arbitrary location at the bedrock via an interpolation method (green arrows). Third, the demand parameter is estimated at the ground surface via a convolution process (red arrow).

1) *Demand Parameter*: In this case study, the demand parameter is the flood inundation depth, which had been estimated from the results of a numerical flood analysis. Fig. 14(a) denotes the location of Mabi town, which is the area of interest that experienced the largest flood. It is located in the city of Kurashiki, Okayama prefecture. The simulation

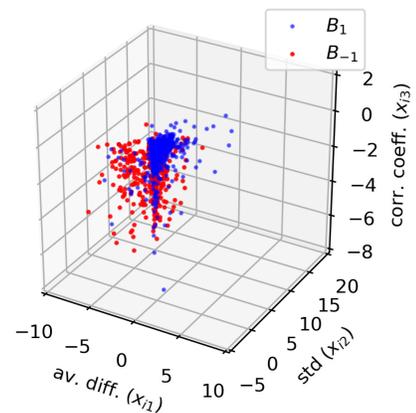


Fig. 12. Sample sets B_1 (blue marks) and B_{-1} (red marks) selected using $D = 80$ cm/s for the CS2.

area is 156 km^2 , and 62 400 grids were prepared. Topographic data were obtained from the Geospatial Information Authority (GSI). We upscaled the grids from the original 5- to 50-m grids for the analysis to reduce the calculation load. The water level data were obtained from the Public Works Research Institute (PWRI). Three observation sites, namely, Higashiminari, Hiwa, and Sakadu, were selected as boundaries. The river area

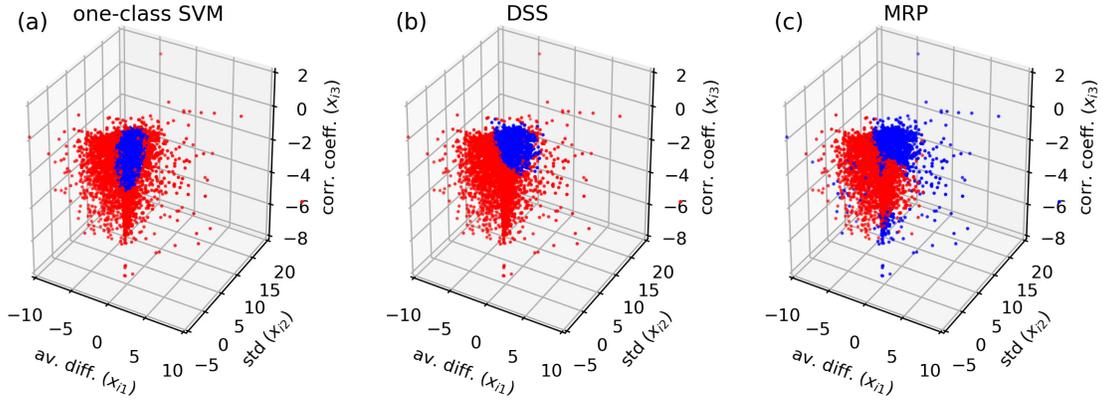


Fig. 13. Changes that were detected in the DSMs that were due to CS2 using (a) one-class SVM, (b) DSS, and (c) MRP methods. The samples in blue correspond to nonchanged buildings, whereas the samples in red are detected changes.

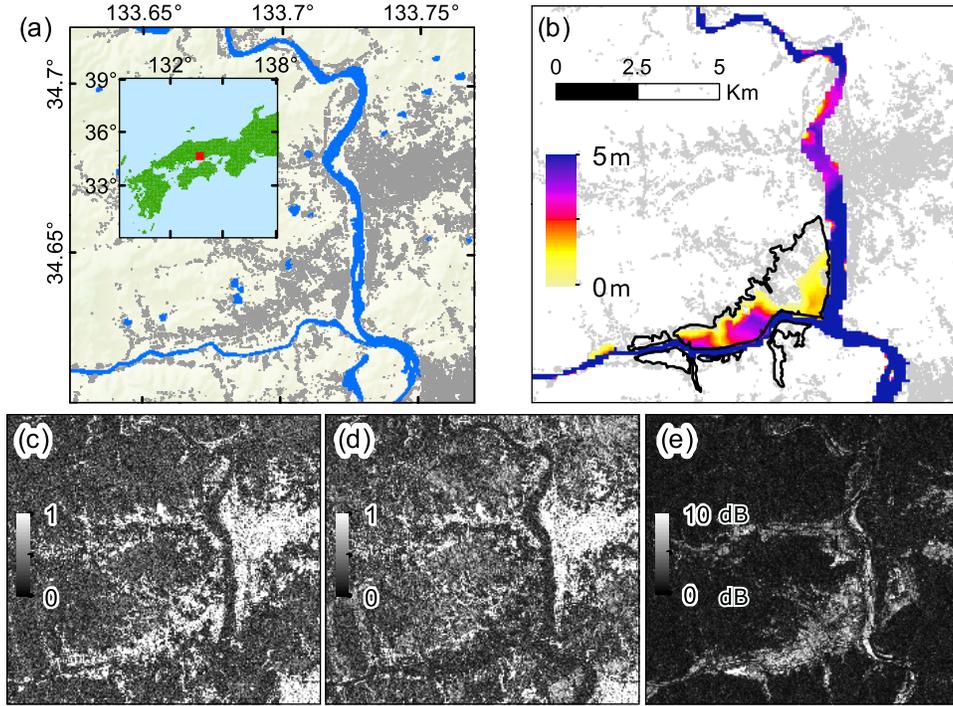


Fig. 14. CS3: the 2018 Western Floods in Japan. (a) Study area: Okayama city. The gray marks denote the buildings' geolocations. The inset shows the location of the study area within western Japan. (b) Numerical-simulation-based inundation depth. (c) Preevent coherence that was computed from a pair of ALOS2-PALSAR2 images that were acquired on March 3, 2017, and April 14, 2018. (d) Coevent coherence that was computed from a pair of ALOS2-PALSAR2 images that were acquired on April 14, 2018, and July 7, 2018. (e) Dissimilarity that was computed from the same imagery as the coevent coherence.

was identified using data on the land use that was provided by GSI, and the river bed elevation was assumed to be -4 m from the original elevation data. The flood at the area of interest was simulated using the following nonlinear shallow water equations:

$$\frac{\partial h}{\partial t} + \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} = 0 \quad (14)$$

$$\frac{\partial M}{\partial t} + \frac{\partial(uM)}{\partial x} + \frac{\partial(vM)}{\partial y} = -gh \frac{\partial H}{\partial x} - \frac{gn^2 u \sqrt{u^2 + v^2}}{h^{1/3}} \quad (15)$$

$$\frac{\partial N}{\partial t} + \frac{\partial(uN)}{\partial x} + \frac{\partial(vN)}{\partial y} = -gh \frac{\partial H}{\partial y} - \frac{gn^2 v \sqrt{u^2 + v^2}}{h^{1/3}} \quad (16)$$

where h is the water depth (m); M and N are the fluxes in the x - and y -directions (m^2/s), respectively; u and v are the velocities in the x - and y -directions (m/s), respectively; H is the water level (m); and g is the acceleration that is caused by gravity (m/s^2). A leap-frog difference scheme was used to solve this system of equations, and structured meshes were used for the calculations. The Manning roughness coefficient, n , was set as $0.025 \text{ s}/\text{m}^{1/3}$ for the river area and $0.03 \text{ s}/\text{m}^{1/3}$ for the flood plain. Rainfall, infiltration, and evaporation were not considered in this model. The flood inundation depth that was obtained via the numerical simulation is presented in Fig. 14(b). Also included is an estimate of the inundated area, which was based on images and videos, as provided by the GSI [57]. According to GSI results, our results underestimate the affected area.

2) *Feature Space*: In the aftermath of the 2018 western floods in Japan, the ALOS-2 satellite conducted observations of the affected areas. Thus, a postdisaster SAR image (POS) that was recorded on July 7, 2018, was available. In addition, two preevent SAR images, which were recorded on March 3, 2017 (PRE1) and April 14, 2018 (PRE2), were used to construct the feature space. All images were in the single look complex (SLC) standard product. As preprocessing, coregistration, radiometric calibration, speckle filter, and terrain correction were applied to the SAR images. The resolution of the images is 3 m. In contrast to the CS1, both phase information and amplitude information are available here; thus, richer information can be used. A 3-D feature space is set in this case study. The first feature (x_{i1}) is the coherence that was computed between the PRE1 and PRE2 images, which is, hereafter, referred to as the preevent coherence. The second feature (x_{i2}) is the coherence that was computed between the PRE2 and the POS images, which is, hereafter, referred to as the coevent coherence. In an ordinary scenario, it is typical to expect large coherence values in urban areas, unless the areas have suffered perturbations, such as the effect of the flood. Thus, it is expected that most of the samples will exhibit large preevent coherence. Furthermore, the samples that are affected by the flood will exhibit lower values of coevent coherence. The third feature (x_{i3}) is the averaged absolute difference in intensity that is computed over a window size of 5×5 between the PRE2 and POS images. Within the context of texture analysis in a 3-D domain, this feature is referred to as the dissimilarity [14]. Depending on the flood inundation depth relative to the building height, the backscattering intensity in the affected area can either decrease or increase. If the inundation depth is much lower than the building height, an increment of the backscattering intensity is expected due to the joint effect of the specular reflection and double-bounding backscattering mechanisms; in contrast, if the inundation depth is larger than the building height, only the specular reflection mechanism occurs, and, thus, the backscattering intensity decreases significantly. Despite negative or positive changes, the dissimilarity will be always positive, which is the main reason for its use as a feature. A total of 41 130 samples were extracted. As in previous cases, the samples were standardized to zero mean and unit standard deviation.

3) *Results*: A threshold $D = 0$ was selected. According to the numerical simulation, 37 947 samples did not experience any inundation ($D = 0$), and 3183 samples were inundated ($D > 0$). Thus, $\#B_1 = \#B_{-1} = 3183$. Fig. 15 shows the sets B_1 and B_{-1} in the feature space, and Fig. 16 depicts the resulted predictions. The final parameters for the DSS method are $\#\hat{B}_{-1} = 2705$ (85% of $\#B_{-1}$), $\lambda = 1.67$, and $\gamma = 0.016$. The final parameters for the MRP method are $\lambda_p = 23.76$, $\lambda_n = 14.38$, and $\gamma = 0.01$. Considering the meanings of the features, it is clear that the results from both the DSS and the MRP methods outperform those from the one-class SVM.

It is unfortunate that we were unable to rigorously evaluate the predictions. As of today, the official report provides only aggregated values of the affected buildings per city and/or town. However, we evaluate the consistency of our predictions

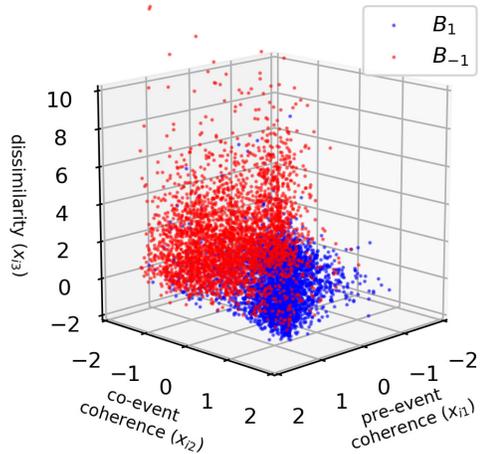


Fig. 15. Sample sets B_1 (blue marks) and B_{-1} (red marks) selected using $D = 0.0$ m for CS3.

TABLE V
CS3: NUMBERS OF SAMPLES PREDICTED AS NC AND CHANGED (C)
WITHIN THE AREA THAT WAS DELINEATED BY GSI

| | One-class SVM | DSS | MRP |
|-------|------------------|------|------|
| NC | 3394 | 1622 | 1615 |
| C | 2534 | 4306 | 4313 |
| Total | 5928 | 5928 | 5928 |

with those from GSI [see Fig. 14(b)]. The area that is delineated by GSI is based on photographs and videos. The field survey was not conducted; thus, we are uncertain regarding the accuracy of the boundary area and whether there were other affected areas nearby. Fig. 17 depicts the building change map product of the predictions from the one-class SVM, DSS, and MRP methods. Again, a significant discrepancy is observed between the results from the one-class SVM and the GSI map. In contrast, the results from both the DSS and MRP methods are less noisy and much more consistent with the GSI map. Table V reports the numbers of buildings that are classified as changed and nonchanged inside the area that is delineated by GSI. The results from one-class SVM are very poor, with about 43% of buildings predicted as changed. In contrast, the percentages of buildings that are classified as changed by DSS and MRP are 73% and 73%, respectively.

IV. DISCUSSION

This manuscript reports the application of SVM-based frameworks for the automatic detection of severely damaged buildings without gathering training samples in the traditional form. While the method is fully automatic, its implementation may require expertise on the disaster that is under consideration. Consider the threshold value, D , for instance. For CS1 and CS3, an inundation depth of $D = 0$, or close to zero, is a safe choice. However, in some cases, it may not be a straightforward decision, such as in the CS2. Within the area of interest, the demand parameter (PGV) exceeded 50 cm/s. Thus, knowledge of earthquake engineering was required for setting a suitable threshold such that B_1 is mainly composed

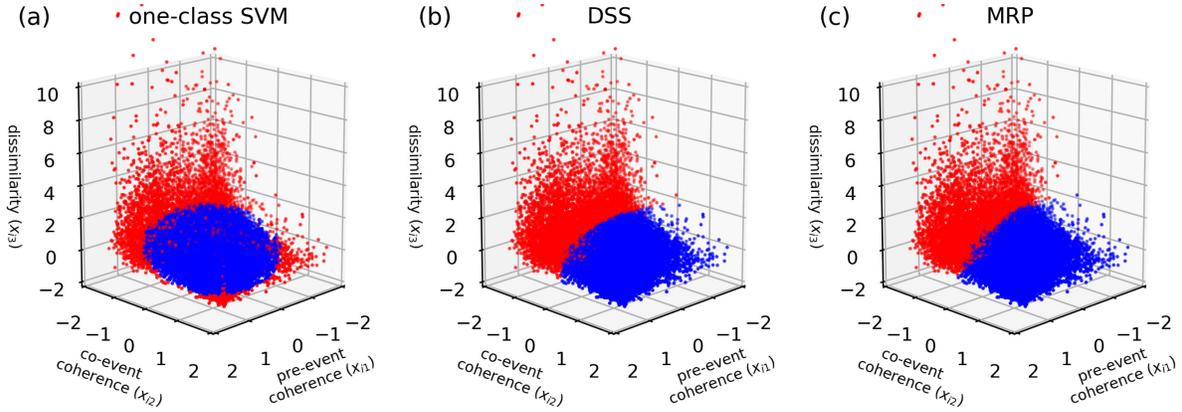


Fig. 16. Changes that were detected in the ALOS-2 PALSAR-2 imagery that were due to CS3 using (a) one-class SVM, (b) DSS, and (c) MRP methods. The samples in blue are unchanged samples, whereas samples in red are samples in which changes have been detected.

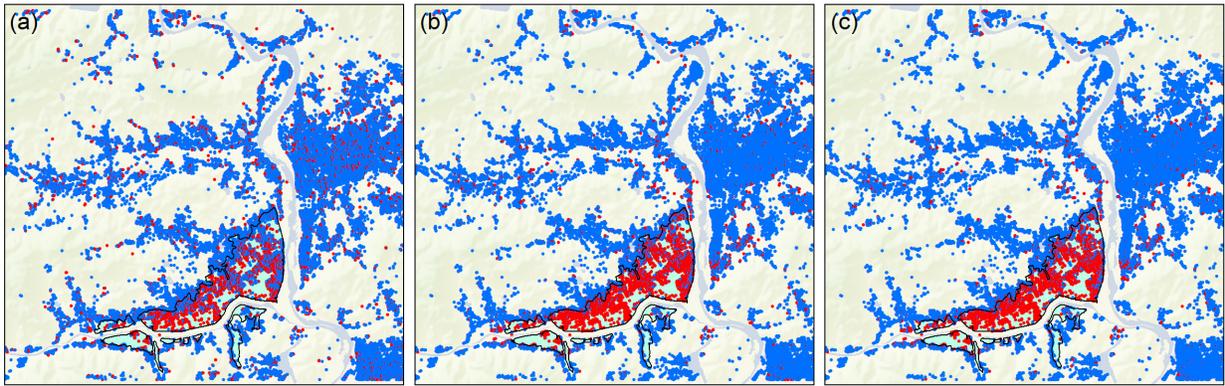


Fig. 17. CS3: spatial distribution of the predicted changes by (a) one-class SVM, (b) DSS, and (c) MRP methods. The black polygon denotes the flooded boundary that was provided by the Geospatial Information Authority of Japan.

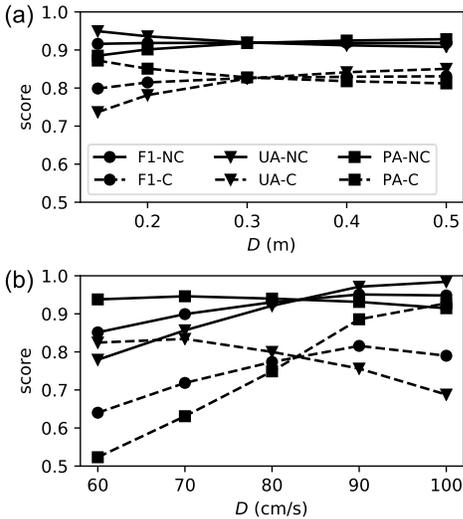


Fig. 18. Effect of the threshold value D on the accuracy of the results for (a) CS1 and (b) CS2.

of nonseverely damaged samples. Another issue that requires consideration is the effect of D on the sizes of the subsets B_1 and B_{-1} . A small value of D may produce a very small B_1 . Fig. 18 shows the effects of D on the accuracy scores (UA, PA, and F1) for CS1 and CS2, for which ground-truth data are available. For CS1, the effect of D is almost negligible.

In contrast, the effect of D is distinguished for CS2. We posit that, for values that are lower than 80 cm/s, the size of B_1 affects the results substantially. Upon closer inspection, the size of B_1 decreased significantly for $D = 60$ cm/s (see Table VI). A tradeoff is observed between the UA and PA scores for both CS1 and CS2. If a change of D increases (decreases) the UA, then the PA will decrease (increase). Hence, the decision boundary is always moving toward either of the classes if D is increasing or decreasing only. For CS1, the decision boundary is approaching the non-changed class as D increases. In contrast, for CS2, the decision boundary is moving toward the changed class as D increases.

Note that this study has focused on the detection of severely damaged buildings. Tables I and III show that the detected changes are mainly associated with the highest damage level, that is, washed away buildings for CS1 and collapsed buildings for the CS2. Intermediate levels of damage were classified as nonchanged. We believe that the detection of intermediate damage levels is limited by the spatial resolution of the remote sensing data. According to FEMA-356 [58], intermediate damage is associated with cracks of 1/8 inches in the structural elements. Besides, it is expected a story-drift of less than 3%, which translates to a roof deformation of 24 cm for two-story buildings. Such patterns cannot be observed with the spatial resolution of the remote sensing data used in this study. As a future study, assuming that remote sensing data with

TABLE VI
EFFECT OF D ON THE SIZE OF $\#B_1$ FOR CS1 AND CS2

| CS1 | | CS2 | |
|------|---------|-----|---------|
| D | $\#B_1$ | D | $\#B_1$ |
| 0.15 | 498 | 60 | 178 |
| 0.20 | 897 | 70 | 539 |
| 0.30 | 1453 | 80 | 1012 |
| 0.40 | 2053 | 90 | 2009 |
| 0.50 | 2553 | 100 | 12,431 |

higher resolution are available, further modifications on the approaches will be carried out for a multiclass classification problem.

We have reported two approaches to calibrate a discriminant function that identifies changed and nonchanged samples. As we observed in Section III, their performance in terms of accuracy is practically the same. There are, however, some contrasts in their implementation. It can be noted that the DSS approach uses the coordinates of the feature vectors from B_{-1} with respect to the set B_1 in the feature space F . Thus, it is not restricted to only the SVM method. Recent developments focus on deep neural networks, which can identify the intrinsic features behind the satellite imagery. For instance, in [59], a cycle-consistent generative adversarial network to calibrate a generator network for transcoding SAR images into the optical image domain is implemented. The referred network is then used to extract bitemporal deep features. In [60], variational autoencoders were used to align the distribution of deep features from different domains, from which one domain consisted of labeled samples, whereas the other domain contained nonlabeled samples. It is our belief that such networks can be adapted to follow the DSS approach as a constraint in the calibration process (i.e., a term in the loss function). On the other hand, the MRP approach uses the quadratic programming problem associated with the one-norm soft-margin SVM; thus, it cannot be adapted to other classifiers.

Another issue to be considered is the implications of the DSS and MRP approaches in their numerical implementation. For instance, the sequential minimal optimization algorithm (SMO) [61], a popular algorithm to solve the SVM optimization problem, breaks the quadratic programming problem [see (11)] into a series of smaller quadratic problems consisting of only two Lagrange multipliers. In unusual cases, when the two chosen Lagrange multipliers are associated with two samples with equal feature vectors but different labels, no improvement in the calibration is achieved. The use of B_{-1} as training samples of class $z = -1$, as in the MRP approach, strongly induces the occurrence of such cases because of the number of nonchanged samples that it might contain. As a result, the search for a suitable pair of Lagrange multipliers may compromise the classifier's performance in terms of runtime. Conversely, the DSS approach uses only samples from B_{-1} that is distant from the region where the nonchanged samples are clustered. Thus, the runtime of the DSS is lower than that of the MRP approach.

In Section II, we stated that the sizes of B_1 and B_{-1} must be balanced, and throughout the empirical evaluations, we used

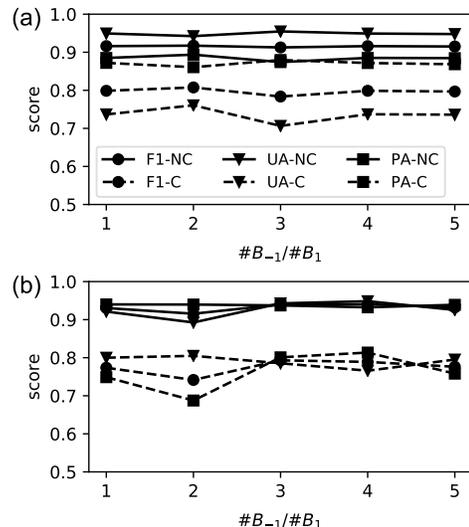


Fig. 19. Effect of the ratio of size on the accuracy of the results for (a) CS1 and (b) CS2.

the same size for both subsets. However, this is not strictly necessary. Intuitively, $\#B_{-1}$ should be larger than $\#B_1$ since B_{-1} is not entirely composed of one class. Fig. 19 reports that the effect of the ratio $\#B_{-1}/\#B_1$ does not change the accuracy scores significantly. The reason lies in the approach that was used to define B_{-1} : B_{-1} must contain the samples that correspond to the largest demand parameter values in the study area.

Regarding the calibration process, the two proposed methods, namely, DSS and MRP, require the tuning of two parameters each. However, for a nonlinear discriminant function, at least one additional parameter (kernel γ) must be tuned. The optimal set of parameter values is selected from a range of discrete values using (8). For the regularizations terms (λ for the DSS and λ_p and λ_n for the MRP) and γ , which is associated with the kernel function, a logarithmic grid from 10^{-2} to 10^2 was employed. For the parameter S from the DSS method, a linear grid from 5% to 100% of the size of B_{-1} was used. If a grid array of N discrete values is used for each parameter, then the discriminant function will be calibrated N^3 times. Thus, significant computational resources may be necessary. Practical strategies are available for mitigating this. For instance, the use of a linear discriminant function will reduce the number of parameters to two, which will significantly reduce the number of calibrations. In our experience, the linear discriminator performs well for binary classification. Another option could be to restrict the sizes of the subsets B_1 and B_{-1} to a few thousand. For the particular case of the MRP approach, if the following restriction is imposed, $\lambda_p > \lambda_n$, then the number of calibrations reduces to about half.

Regarding CS3, namely, the case study of the 2018 Okayama floods, the numerical simulation of the floods was conducted under the assumption of real-time application, namely, only information that was available at the time of the event was used. In addition, the grid resolution was reduced to reduce the runtime. Thus, the few constraints and the lower resolution were two of the reasons

for the differences between our estimates and the flooded area that was delimited by GSI. However, DSS and MRP are robust under low-accuracy estimates of the demand parameter. This is because the proposed methods are based on the soft-margin SVM, which considers outliers that are produced by the inaccuracy in the estimation of the demand parameter, among other factors. Although not presented here, our predictions could be used as additional constraints to improve the numerical flood simulation. However, this is a topic for another study.

V. CONCLUSION

This manuscript introduces the use of the demand parameter, which quantifies the disaster intensity, to systematically extract samples from remote sensing imagery and use them to calibrate a change detection classifier. The demand parameter of each sample is estimated via instrumentation and/or numerical simulation, which can be computed in real or near real time. We propose the use of a demand parameter map to group the samples into two subsets, where one subset is composed of samples for which the geolocations experienced low demand and the other subset is composed of samples with medium and/or large demand. We assumed that the first subset was mainly composed of nonchanged samples and that the second subset was composed of both changed and nonchanged samples. Under these constraints, two methods are reported for calibrating a discriminant function. The first method is composed of two main steps. First, the subset with low demand is used to calibrate the discriminant function using the one-class support vector machine (SVM). Second, the discriminant function is improved using the other subset. The second method uses a soft margin SVM with two regularization parameters. In contrast to the standard SVM, which employs one regularization parameter, the SVM with two regularization terms can have different levels of tolerance for the subsets, namely, the discriminant function will accept few outliers from the subset that is composed of samples with low demand while being highly flexible and accepting many outliers from the subset that is composed of samples with large demand.

The proposed methods were evaluated on three disasters: the 2011 Tohoku earthquake–tsunami, the 2016 Kumamoto earthquake, and the 2018 western Japan floods. In addition, the feature space of each case study was constructed from different types of remote sensing data. Backscattering intensities from microwave imagery were used for the first case study, Lidar-based DSMs were employed in the second case study, and backscattering complex values from microwave imagery were used for the third case study. The results were of approximately the same level of accuracy as the results that were reported in previous studies in which traditional machine learning methods were employed. However, in contrast to the other studies, our methods can be used in near real time.

In the aftermath of a large-scale disaster, the traditional procedure for extracting training samples represents the bottleneck in the creation of a machine-learning-based damage map. The automatic extraction of training samples is an open problem in the use of machine learning for early disaster response.

Therefore, the relevance of our study is that it contributes to solutions to events from which the disaster intensity can be estimated.

ACKNOWLEDGMENT

The satellite images were preprocessed with ArcGIS 10.6 and ENVI 5.5, and the other processing and analysis steps were implemented in Python using the Geospatial Data Abstraction Library (GDAL) and the NumPy Library.

REFERENCES

- [1] B. Pan, Z. Shi, and X. Xu, "Multiobjective-based sparse representation classifier for hyperspectral imagery using limited samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 239–249, Jan. 2019.
- [2] C. Geiß, P. A. Pelizari, L. Blickensdörfer, and H. Taubenböck, "Virtual support vector machines with self-learning strategy for classification of multispectral remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 42–58, May 2019.
- [3] C. Geiß, M. Thoma, M. Pittore, M. Wieland, S. W. Dech, and H. Taubenböck, "Multitask active learning for characterization of built environments with multisensor Earth observation data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5583–5597, Dec. 2017.
- [4] C. Persello, A. Boularias, M. Dalponte, T. Gobakken, E. Næsset, and B. Schölkopf, "Cost-sensitive active learning with lookahead: Optimizing field surveys for remote sensing data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6652–6664, Oct. 2014.
- [5] C. Geib, M. Thoma, and H. Taubenböck, "Cost-sensitive multitask active learning for characterization of urban environments with remote sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 6, pp. 922–926, Jun. 2018.
- [6] C. Huang *et al.*, "Use of a dark object concept and support vector machines to automate forest cover change analysis," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 970–985, Mar. 2008.
- [7] C. Geiß, P. A. Pelizari, S. Bauer, A. Schmitt, and H. Taubenböck, "Automatic training set compilation with multisource geodata for DTM generation from the TanDEM-X DSM," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 456–460, Mar. 2019.
- [8] E. Booth, K. Saito, R. Spence, G. Madabhushi, and R. T. Eguchi, "Validating assessments of seismic damage made from remote sensing," *Earthq. Spectra*, vol. 27, no. 1, pp. 157–177, Oct. 2011.
- [9] A. Cooner, Y. Shao, and J. Campbell, "Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 haiti earthquake," *Remote Sens.*, vol. 8, no. 10, p. 868, Oct. 2016.
- [10] M. Janalipour and A. Mohammadzadeh, "A fuzzy-GA based decision making system for detecting damaged buildings from high-spatial resolution optical images," *Remote Sens.*, vol. 9, no. 4, p. 349, Apr. 2017.
- [11] H. Gokon *et al.*, "A method for detecting buildings destroyed by the 2011 Tohoku earthquake and tsunami using multitemporal TerraSAR-X data," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1277–1281, Jun. 2015.
- [12] M. Wieland, W. Liu, and F. Yamazaki, "Learning change from synthetic aperture radar images: Performance evaluation of a support vector machine to detect earthquake and tsunami-induced changes," *Remote Sens.*, vol. 8, no. 10, p. 792, Sep. 2016.
- [13] Y. Bai *et al.*, "A framework of rapid regional tsunami damage recognition from post-event TerraSAR-X imagery using deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 43–47, Jan. 2018.
- [14] L. Moya, H. Zakeri, F. Yamazaki, W. Liu, E. Mas, and S. Koshimura, "3D gray level co-occurrence matrix and its application to identifying collapsed buildings," *ISPRS J. Photogramm. Remote Sens.*, vol. 149, pp. 14–28, Mar. 2019.
- [15] L. Moya, F. Yamazaki, W. Liu, and M. Yamada, "Detection of collapsed buildings from lidar data due to the 2016 kumamoto earthquake in Japan," *Natural Hazards Earth Syst. Sci.*, vol. 18, no. 1, pp. 65–78, Jan. 2018.
- [16] B. Adriano, J. Xia, G. Baier, N. Yokoya, and S. Koshimura, "Multi-source data fusion based on ensemble learning for rapid building damage mapping during the 2018 sulawesi earthquake and tsunami in Palu, Indonesia," *Remote Sens.*, vol. 11, no. 7, p. 886, Apr. 2019.

- [17] L. Moya *et al.*, "Detecting urban changes using phase correlation and 1-based sparse model for early disaster response: A case study of the 2018 sulawesi Indonesia earthquake-tsunami," *Remote Sens. Environ.*, vol. 242, Jun. 2020, Art. no. 111743.
- [18] UNITAR/UNOSAT, EC JRC, and World Bank. (2010). *Haiti Earthquake 2010: Remote Sensing based Building Damage Assessment Data*. Accessed: Apr. 2, 2019. [Online]. Available: <https://unitar.org/unosat/haiti-earthquake-2010-remote-sensing-based-building-damage-assessment-data>
- [19] Ministry of Land, Infrastructure, Transport and Tourism (MLIT). (2011). *Results of the Survey on Disaster Caused by the Great East Japan Earthquake (First Report)*. Accessed: Apr. 2, 2019. [Online]. Available: http://www.mlit.go.jp/report/press/city07_hh_000053.html
- [20] M. Yamada, J. Ohmura, and H. Goto, "Wooden building damage analysis in Mashiki Town for the 2016 Kumamoto earthquakes on April 14 and 16," *Earthq. Spectra*, vol. 33, no. 4, pp. 1555–1572, Nov. 2017.
- [21] Copernicus, Emergency Management Service. (2018). *EMSR317: Earthquake in Indonesia*. Accessed: Apr. 2, 2019. [Online]. Available: <https://emergency.copernicus.eu/mapping/list-of-components/EMSR317>
- [22] E. Mas *et al.*, "Characteristics of tsunami fragility functions developed using different sources of damage data from the 2018 sulawesi earthquake and tsunami," *Pure Appl. Geophys.*, vol. 177, no. 6, pp. 2437–2455, Jun. 2020.
- [23] Y. Ohta, H. Murakami, Y. Watoh, and M. Koyama, "A model for evaluating life span characteristics of entrapped occupants by an earthquake," in *Proc. 13th World Conf. Earthq. Eng.*, 2004, p. 9.
- [24] S. Ghosh *et al.*, "Crowdsourcing for rapid damage assessment: The global Earth observation catastrophe assessment network (GEO-CAN)," *Earthq. Spectra*, vol. 27, no. S1, pp. 179–198, Oct. 2011.
- [25] M. Wieland and O. Murao, "A modular processing chain for automated flood monitoring from multi-spectral satellite data," *Remote Sens.*, vol. 11, no. 19, p. 2330, Oct. 2019.
- [26] L. Moya, E. Mas, and S. Koshimura, "Learning from the 2018 Western Japan heavy rains to detect floods during the 2019 Hagibis typhoon," *Remote Sens.*, vol. 12, no. 14, p. 2244, Jul. 2020.
- [27] C. A. Cornell, "Engineering seismic risk analysis," *Bull. Seismol. Soc. Amer.*, vol. 58, no. 5, pp. 1583–1606, 1968.
- [28] F. Yamazaki and O. Murao, *Vulnerability Functions for Japanese Buildings Based on Damage Data From the 1995 Kobe Earthquake (Implications of Recent Earthquakes on Seismic Risk)*, vol. 2. London, U.K.: Imperial College Press, 2000.
- [29] K. Porter, R. Kennedy, and R. Bachman, "Creating fragility functions for performance-based earthquake engineering," *Earthq. Spectra*, vol. 23, no. 2, pp. 471–489, May 2007.
- [30] S. Koshimura, T. Oie, H. Yanagisawa, and F. Imamura, "Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from banda aceh, Indonesia," *Coastal Eng. J.*, vol. 51, no. 3, pp. 243–273, Sep. 2009.
- [31] E. Mas *et al.*, "Developing tsunami fragility curves using remote sensing and survey data of the 2010 Chilean tsunami in dichato," *Natural Hazards Earth Syst. Sci.*, vol. 12, no. 8, pp. 2689–2697, Aug. 2012.
- [32] A. Suppasri *et al.*, "Building damage characteristics based on surveyed data and fragility curves of the 2011 Great East Japan tsunami," *Natural Hazards*, vol. 66, no. 2, pp. 319–341, Mar. 2013.
- [33] M. Villar-Vega *et al.*, "Development of a fragility model for the residential building stock in South America," *Earthq. Spectra*, vol. 33, no. 2, pp. 581–604, May 2017.
- [34] H. Lovon, N. Tarque, V. Silva, and C. Yepes-Estrada, "Development of fragility curves for confined masonry buildings in Lima, Peru," *Earthq. Spectra*, vol. 34, no. 3, pp. 1339–1361, Aug. 2018.
- [35] S. Medina, J. Lizarazo-Marriaga, M. Estrada, S. Koshimura, E. Mas, and B. Adriano, "Tsunami analytical fragility curves for the Colombian Pacific coast: A reinforced concrete building example," *Eng. Struct.*, vol. 196, Oct. 2019, Art. no. 109309.
- [36] I. Charvet, A. Suppasri, H. Kimura, D. Sugawara, and F. Imamura, "A multivariate generalized linear tsunami fragility model for Kesenuma city based on maximum flow depths, velocities and debris impact, with evaluation of predictive accuracy," *Natural Hazards*, vol. 79, no. 3, pp. 2073–2099, Dec. 2015.
- [37] Y. Maruyama, F. Yamazaki, K. Mizuno, Y. Tsuchiya, and H. Yogai, "Fragility curves for expressway embankments based on damage datasets after recent earthquakes in Japan," *Soil Dyn. Earthq. Eng.*, vol. 30, no. 11, pp. 1158–1167, Nov. 2010.
- [38] A. Martinelli, G. Cifani, G. Cialone, L. Corazza, A. Petracca, and G. Petrucci, "Building vulnerability assessment and damage scenarios in Celano (Italy) using a quick survey data-based methodology," *Soil Dyn. Earthq. Eng.*, vol. 28, nos. 10–11, pp. 875–889, 2008.
- [39] C. Geiß *et al.*, "Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques," *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 175–188, Jun. 2015.
- [40] A. C. Zülfiqar, N. Ö. Z. Fercan, S. Tunç, and M. Erdik, "Real-time earthquake shake, damage, and loss mapping for Istanbul metropolitan area," *Earth, Planets Space*, vol. 69, no. 1, p. 9, Jan. 2017.
- [41] S. Karimzadeh, B. Feizizadeh, and M. Matsuoka, "From a GIS-based hybrid site condition map to an earthquake damage assessment in Iran: Methods and trends," *Int. J. Disaster Risk Reduction*, vol. 22, pp. 23–36, Jun. 2017.
- [42] N. I. Frolova, V. I. Larionov, J. Bonnin, S. P. Sushchev, A. N. Ugarov, and M. A. Kozlov, "Seismic risk assessment and mapping at different levels," *Natural Hazards*, vol. 88, no. S1, pp. 43–62, Aug. 2017.
- [43] L. Moya, E. Mas, B. Adriano, S. Koshimura, F. Yamazaki, and W. Liu, "An integrated method to extract collapsed buildings from satellite imagery, hazard distribution and fragility curves," *Int. J. Disaster Risk Reduction*, vol. 31, pp. 1374–1384, Oct. 2018.
- [44] L. Moya, L. M. Perez, E. Mas, B. Adriano, S. Koshimura, and F. Yamazaki, "Novel unsupervised classification of collapsed buildings using satellite imagery, hazard scenarios and fragility functions," *Remote Sens.*, vol. 10, no. 2, p. 296, Feb. 2018.
- [45] F. Bovolo, L. Bruzzone, and M. Marconcini, "A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2070–2082, Jul. 2008.
- [46] F. Bovolo, G. Camps-Valls, and L. Bruzzone, "A support vector domain method for change detection in multitemporal images," *Pattern Recognit. Lett.*, vol. 31, no. 10, pp. 1148–1154, Jul. 2010.
- [47] D. M. J. Tax and R. P. W. Duin, "Data domain description using support vectors," in *Proc. Eur. Symp. Artif. Neural Netw.*, 1999, pp. 251–256.
- [48] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443–1471, Jul. 2001.
- [49] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.
- [50] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York, NY, USA: Cambridge Univ. Press, 2007.
- [51] L. Moya, Y. Endo, G. Okada, S. Koshimura, and E. Mas, "Drawback in the change detection approach: False detection during the 2018 western Japan floods," *Remote Sens.*, vol. 11, no. 19, p. 2320, Oct. 2019.
- [52] QuiQuake. (2016). *Quick Estimation System for Earthquake Maps Triggered by Observation Records*. Accessed: May 7, 2019. [Online]. Available: <http://gbank.gsj.jp/QuiQuake/QuickMap/index.en.html>
- [53] Cabinet Office of Japan. (2016). *Summary of Damage Situation in the Kumamoto Earthquake Sequence*. Accessed: Mar. 3, 2020. [Online]. Available: <http://www.bousai.go.jp/updates/h280414jishin/index.html>
- [54] L. Moya, E. Mas, F. Yamazaki, W. Liu, and S. Koshimura, "Statistical analysis of earthquake debris extent from wood-frame buildings and its use in road networks in Japan," *Earthq. Spectra*, vol. 36, no. 1, pp. 209–231, Feb. 2020.
- [55] M. Matsuoka and N. Yamamoto, "Web-based quick estimation system of strong ground motion maps using engineering geomorphologic classification map and observed seismic records," in *Proc. 15th World Conf. Earthq. Eng.*, 2012, pp. 24–28.
- [56] Cabinet Office of Japan. (2018). *Summary of Damage Situation Caused by the Heavy Rainfall in July 2018*. Accessed: Aug. 7, 2018. [Online]. Available: <http://www.bousai.go.jp/updates/h30typhoon7/index.html>
- [57] Geospatial Information Authority of Japan. (2018). *Information About the July 2017 Heavy Rain*. Jul. 19, 2019. [Online]. Available: <https://www.gsi.go.jp/BOUSAI/H30.taihuu7gou.html>
- [58] *Prestandard and Commentary for the Seismic Rehabilitation of Buildings*, Prepared Amer. Soc. Civil Eng. Federal Emergency Manage. Agency, Washington, DC, USA, 2000.
- [59] S. Saha, F. Bovolo, and L. Bruzzone, "Building change detection in VHR SAR images via unsupervised deep transcoding," *IEEE Trans. Geosci. Remote Sens.*, early access, Jun. 18, 2020, doi: 10.1109/TGRS.2020.3000296.
- [60] Y. Li, W. Hu, H. Li, H. Dong, B. Zhang, and Q. Tian, "Aligning discriminative and representative features: An unsupervised domain adaptation method for building damage assessment," *IEEE Trans. Image Process.*, vol. 29, pp. 6110–6122, Apr. 2020.
- [61] J. Platt, "Fast training of support vector machines using sequential minimal optimization," in *Advances in Kernel Methods—Support Vector Learning*. Cambridge, MA, USA: MIT Press, Jan. 1998.



Luis Moya was born in Lima, Peru, in 1985. He received the B.S. and M.S. degrees in civil engineering from the National University of Engineering, Lima, in 2009 and 2013, respectively, and the Ph.D. degree in engineering from Chiba University, Chiba, Japan, in 2016.

In 2012, he was a Visiting Student with the Building Research Institute, Tsukuba, Japan. From 2016 to 2020, he was a Researcher with the International Research Institute of Disaster Science (IRIDeS), Tohoku University, Sendai, Japan. Since 2020, he has been a Research Fellow with the Japan-Peru for Earthquake Engineering Research and Disaster Mitigation, National University of Engineering. He is a Principal Investigator of a project that aims to integrate remote sensing, in-place sensors, and numerical simulations for early disaster response. His research interests include earthquake engineering, remote sensing technologies, intelligent evacuation systems, applied machine learning for disaster mitigation.

Dr. Moya received the Alexander von Humboldt Fellowship in 2019.



Christian Geiß (Member, IEEE) received the M.Sc. degree in applied geoinformatics from the Paris Lodron University of Salzburg, Salzburg, Austria, in 2010, and the Ph.D. degree (Dr. rer. nat.) from the Humboldt University of Berlin, Berlin, Germany, in 2014. He is pursuing the Habilitation degree in geography with the Julius-Maximilians University of Würzburg, Würzburg, Germany, with the title “Collective Sensing Techniques and Artificial Intelligence for Natural Hazard Risk and Impact Assessment.”

Since 2010, he has been with the German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Weßling, Germany. In 2017, he was also with the Cambridge University Center for Risk in the Built Environment (CURBE), University of Cambridge, Cambridge, U.K., as a Visiting Scholar. Consequently, his research interests include the development of machine learning methods for the interpretation of Earth observation data, multimodal remote sensing of the built environment, exposure and vulnerability assessment in the context of natural hazards, as well as techniques for automated damage assessment after natural disasters.



Masakazu Hashimoto was born in Kamogawa, Japan. He graduated from the University of Yamanashi, Kofu, Japan. He received the Ph.D. degree from the Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, in 2014.

He is an Assistant Professor with the International Research Institute of Disaster Science, Tohoku University, Sendai, Japan. His research interests are related to flood numerical simulation, hydro-epidemiology, and unmanned aerial vehicle (UAV) photogrammetry.



Erick Mas was born in Lima, Peru. He graduated from the National University of Engineering, Lima. He received the Ph.D. degree from the Graduate School of Engineering, Tohoku University, Sendai, Japan, in 2012.

He has professional experience in disaster risk management at regional and local governments in Peru. He is an Associate Professor with the International Research Institute of Disaster Science and also appointed at the Tough Cyber Physical AI Research Center, both in Tohoku University. His research interests are related to agent-based modeling, tsunami numerical simulation, evacuation simulation, unmanned aerial vehicle (UAV) and remote sensing for tsunami damage, and risk assessment.



Shunichi Koshimura received the Ph.D. degree from the Graduate School of Engineering, Tohoku University, Sendai, Japan, in 2000.

He is a Professor with the International Research Institute of Disaster Science, Tohoku University. His research interests include developing a real-time natural hazard forecast system and estimating social impacts by integrating numerical modeling, Earth observation, and geoinformatics. He is also a Co-Founder of RTi-cast, a technology firm to offer real-time tsunami inundation damage forecast services to government organizations and commercial clients.



Günter Strunz received the Diploma degree in geodesy and the Ph.D. degree in photogrammetry and remote sensing from Technical University Munich (TUM), Munich, Germany, in 1984 and 1992, respectively.

Since 1993, he has been with the German Aerospace Center (DLR), Weßling, Germany, where he is the Head of Department of Geo Risks and Civil Security. His main research fields are satellite-based Earth observation and its applications in environmental monitoring, emergency response and disaster risk research, as well as the development of early warning systems for natural hazards.