Bulletin of Earthquake Engineering
https://doi.org/10.1007/s10518-019-00648-7

ORIGINAL RESEARCH

# A transferable remote sensing approach to classify building structural types for seismic risk analyses: the case of Val d'Agri area (Italy)

**Mariangela Liuzzi[1] · Patrick Aravena Pelizari[2] · Christian Geiß[2] · Angelo Masi[1] · Valerio Tramutoli[1] · Hannes Taubenböck[2]**

Received: 19 June 2018 / Accepted: 23 May 2019


**Abstract**
This study proposes a methodology based on machine learning (ML) algorithms for rapid and robust classification of building structural types (STs) in multispectral remote sensing imagery aiming to assess buildings' seismic vulnerability. The seismic behavior of buildings is strongly affected by the ST, including material, age, height, and other main structural features. Previous works deployed in situ data integrated with remote sensing information to statistically infer STs through supervised ML methods. We propose a transferable methodology with specific focus on situations with imbalanced in situ data (i.e., the number of available labeled samples for model learning differs largely between different STs). We learn a transferable model by selecting features from an exhaustive set. The transferability relies on deploying geometric features characterizing individual buildings; thus, the model is less sensitive to domain adaption problems frequently induced by e.g., changes in acquisition parameters of remotely sensed imagery. Thereby, we show that few geometry features enable generalization capabilities similar to models learned with a large number of features describing spectral, geometrical or contextual building properties. We rely on an extensive geodatabase containing almost 18,000 building footprints. We follow a Random Forest (RF)-based feature selection strategy to objectively identify most valuable features for prediction. Furthermore, the problem of unbalanced classes is addressed by adopting two approaches: downsampling the majority class and modifying the classifier internally (weighted RF). The implemented model is transferred on the challenging urban morphology of the Val d'Agri area (Italy). Results confirm the statistical robustness of the model and the importance of the geometry features, allowing for reliable identification of STs.

**Keywords** Machine learning · Remote sensing · Seismic vulnerability · Building structural type · Building inventory · Class imbalance

✉ Mariangela Liuzzi
mariangelaliuzzi88@gmail.com

[1] School of Engineering, Università degli Studi della Basilicata, Via Nazario Sauro 85, 85100 Potenza, Italy

[2] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchener Straße 20, 82234 Weßling, Germany

Published online: 15 June 2019

## 1 Introduction

The disastrous seismic events which struck Italy in the last 50 years caused more than 5000 casualties and an economic loss of more than 120 billion €. These numbers stress the high vulnerability of urban systems with respect to earthquakes. In order to adequately mitigate risk, the vulnerability of structures must be decreased both by means of earthquake-resistant design for new structures and, especially, through strengthening actions on the existing ones. The destructive events in Italy in 2016 emphasize the need for an urgent area-wide vulnerability assessment of existing buildings to identify hot spots of damage and consequently carry out appropriate protection measures. Risk analyses and earthquake damage scenarios exploit quantity, arrangement, and value of the exposed structures and their expected seismic behavior to quantify physical effects due to earthquakes of given intensity (Dolce et al. 2003; Chiauzzi et al. 2012).

The expected seismic behavior of buildings due to an earthquake having fixed intensity is assessed by "Damage Estimation Models" (DEMs). Several DEMs have been developed following different approaches based on specific physical characteristics of the structures as input data (Calvi et al. 2006). In this study, the term "building stock" is confined to residential buildings and does not involve other exposed elements such as public buildings and lifelines. Physical variables characterizing the building stock such as building height, size, spatial configuration, symmetry in elevation and plan, disposal of resisting elements in plan, arrangement of buildings, foundations type, etc. are crucial for the application of DEMs (Dolce et al. 2003). These properties enable the classification of the structural type (ST) which has a significant influence on seismic behavior (Taubenböck et al. 2009): buildings that pertain to the same ST are likely to show similar behavior with respect to a given ground shaking. The calculation of expected losses in terms of number of destroyed and unusable buildings following earthquakes of given intensity can allow estimation of homeless people, fatalities, needs for shelters, and economic losses. Consequently, this information supports the development of strategies aimed at decreasing the actual seismic risk; in a pre-event phase it enables the intelligent and efficient preparedness of emergency efforts relevant to the post-event phase. Therefore, it is crucial to assess the vulnerability of the exposed building stock. At the same time, compiling this information is very demanding in terms of required time as well as economic and human resources.

Countries highly threatened by earthquake risk most frequently lack reliable building datasets containing important information and attributes (Geiß et al. 2016), due to format problems or to restricted access for privacy concerns. In the perspective of efficient risk assessment, the scientific community is operating towards the preparation of building inventories involving the calculation of physical parameters. Large areas can be investigated thanks to the spatial resolution and coverage of remote sensing imagery (Mueller et al. 2006; Taubenböck et al. 2008). Thereby, the usability of remote sensing for contributing to the determination of seismic risk in different areas of the world is investigated jointly by civil engineering and remote sensing communities. Numerous recent studies emphasize the feasibility of remote sensing in supporting pre-event vulnerability assessments and post-event structural damage evaluation of built-up structures in a spatially contiguous manner. Sarabandi et al. (2008) developed a set of methodologies for updating spatial and geometric information of buildings from single and multiple high-resolution optical satellite images. Balkaya et al. (2015) proposed a GIS-based real time monitoring system deploying satellite imagery for disaster management in the aftermath of damaging events. Casciati et al. (2016) designed and implemented

a satellite-based asset tracking system to sustain emergency preparedness in the context of a critical event. However, a globally applicable methodology enabling the fully automated extraction of physical parameters affecting the seismic vulnerability of the building stock does not exist. More specifically, not all the parameters that are crucial for a vulnerability assessment (e.g., the ST) can be directly calculated but have to be inferred in an indirect manner.

The first aim of this study is to explore the potential of the combined use of remote sensing data and a large in situ building data base to classify the building stock by ST in order to provide large area information for the implementation of DEMs. Various studies proved that physical characteristics describing the building stock derived from multi-sensor remote sensing data allow for an assessment of building STs. Furthermore, the fragility curves based on such data have been applied for the risk assessment of exposed buildings (Taubenböck et al. 2009; Geiß et al. 2015). However, most studies focus on larger metropolitan areas where a multitude of clearly distinguishable STs are present which allows the correlation between building STs and fragility curves. In contrast, this study deals with a dominantly rural and peri-urban area with less STs: only two dominating STs are present. The two STs exhibit features that appear very similar from a remote sensing perspective. Thus, the identification of the two different classes is challenging and requires meaningful and robust features. At the same time, the morphology and spatial composition of these two STs is very complex. This complex morphology, representative of the Italian small urban centers, will allow us to show that even in such difficult morphologic environments remote sensing can provide data useful for the large scale vulnerability assessment of buildings with viable accuracies.

However, in many real-world situations, the availability of in situ data on the building stock is limited and the representability of the samples deployed to train the classification algorithm is frequently restricted to the source domain (i.e., the area which the samples were collected from) (e.g. Geiß et al. 2015). In this context, transfer-learning problems can affect the analysis of remote sensing data. Thereby, the challenge is to generate a robust model which is learnt on a source domain and also applicable on a target domain without a significant decrease in accuracy. Especially a shift of the covariates (i.e., features), as frequently induced by changes during the data acquisition (e.g., atmospheric condition, illumination angle etc.), can hamper the unadapted transfer of models (Tuia et al. 2016). We propose a transferable methodology with specific focus on situations with imbalanced in situ data (i.e., the number of available labeled samples for model learning differs largely between different STs). We learn a transferable model by selecting features from an exhaustive set. The transferability relies on deploying solely geometric features characterizing individual buildings; thus, the model is less sensitive to domain adaption problems frequently induced by a covariate shift. Thereby, we show that those few geometry features enable generalization capabilities similar to models learned with a large number of features describing spectral, geometrical or contextual building properties. The developed methodology is supported by an in situ database of almost 18,000 buildings, supplying substantial data for validation and being a particular novelty of this work; it provides detailed physical knowledge on STs (Masi et al. 2014).

The following organization is used for the rest of the document. Section 2 describes the study area and the seismic characteristics of the two building classes; furthermore, in situ data, the derived geo-information and remote sensing data, are also explained. Section 3 outlines the methodology, introduces the features used for the classification and explains the algorithms as well as the metrics used to assess the accuracy of the classification maps. Section 4 illustrates the experimental setup. Section 5 contains the analyses of the obtained

results. Section 6 provides a discussion of the results in the perspective of future developments in engineering applications; in Sect. 7 final conclusions are remarked.

## 2 Materials

### 2.1 Study area: Val d'Agri, Italy

The Val d'Agri area (1405.45 $km^2$), located in the Basilicata region in Southern Italy, comprises 18 villages. This area, situated in the seismically active axial sector of Southern Apennines, is characterized by high hazard: among the most destructive events which have stricken the Italian peninsula, it is worth mentioning the 16 December 1857 earthquake (X–XI intensity of Mercalli scale, 7.0 magnitude) having its epicenter exactly in Val d'Agri. This event concerned a considerable portion of the Southern Apennines, and in the Val d'Agri area it caused heavy damage and many casualties in the Potenza and Salerno provinces (INGV 2018).

From a remote sensing point of view, the classification of the dwelling buildings into STs influencing their seismic behavior is challenging due to the organic urban structure of the villages (Fig. 1). Specifically, the arrangement of the buildings is very diverse and complex and distinctive spatial patterns cannot be clearly identified: generally, the villages present a very dense historical center, where buildings are mostly irregularly arranged and very close to each other, often sharing the same walls, and a more recent surrounding area where buildings show lower density.

Two main building types are present in the area of interest: masonry (Fig. 2a) and reinforced concrete (RC; Fig. 2b). The classification of buildings in masonry and RC is a useful element to preliminarily distinguish seismic performances. Specifically, a first difference is that in masonry buildings the resistance to ground shaking is distributed almost in all the vertical structural elements, while in RC buildings this resistance is concentrated only in some specific elements. Masonry is a compound material consisting of large size inert elements, having natural or artificial origin, and a binder. The performance of masonry structures is mainly determined by: nature, geometry and quality of ceilings and wall surfaces; kind and effectiveness of the connection elements, as well as construction details (Penazzi et al. 2000). RC buildings show seismic performance, which is mainly due to strength

**(a)** **(b)**

**Fig. 1** Example of the arrangement of the building stock in the studied area (Viggiano village), **a** from aerial picture; **b** from 2013 orthophoto

(a) (b)

**Fig. 2** **a** Example of ensemble of masonry buildings; **b** example of ensemble of RC buildings in the study area in their typical arrangement

and ductility of structural elements. Strength and ductility in the individual elements are strongly dependent on mechanical and quantitative characteristics of steel and concrete (longitudinal and transversal rebar ratio, grip between the two materials, dimensions of the reinforcement bars, resistance), disposal of the two elements (mooring in concrete and overlap, disposition in the section of tension and compression reinforcement, constructional details in beam-column joints, elongation of longitudinal and transversal reinforcement bars), quality of the executed work (Fardis 2009; Masi et al. 2013). Ultimately, geometrical attributes of RC buildings contribute in characterizing their performance under seismic load (Masi 2003; Masi and Vona 2012).

## 2.2 Derived geoinformation and in situ data

A thorough building-by-building inventory is available in GIS format for the 18 villages of the Val d'Agri area for pre-event vulnerability analyses (Masi et al. 2014). This inventory was collected by purposely-trained technicians under the supervision of researchers from the University of Basilicata School of Engineering during in situ inspections between 2001 and 2006. The data set consists of a digital cadastral map of 17,462 geolocated building polygons (Table 1) affiliated with physical parameters which affect their seismic vulnerability. These include the total number of storeys, average storey height, average floor area, building age and use, regularity, ST, roof type and building use, among others.

With regard to the ST, the data base consists of 13,819 masonry and 3643 RC buildings (79.1% and 20.9% of the total inventory respectively). For a detailed description of the

**Table 1** Numbers of buildings and ST distributions of the whole inventory, the village Paterno deployed for the learning of ST classification models and the remaining 17 villages used for evaluating the model transfer

| Structural type | Buildings [no. (%)] | | |
|---|---|---|---|
| | All 18 villages | Paterno (model learning) | 17 remaining villages (model transfer) |
| Masonry | 13,819 (79.1) | 1338 (72.6) | 12,481 (79.9) |
| RC | 3643 (20.9) | 506 (27.4) | 3137 (20.1) |
| Total | 17,462 | 1844 | 15,618 |

survey procedure, the collected building attributes as well as the deployed survey form the reader is referred to Masi et al. (2014).

Mapping inconsistencies within the inventory building footprint geometries were corrected by manually editing inaccurate polygons taking the remote sensing data (Sect. 2.3) as reference.

### 2.3 Remote sensing data

With concern to the remote sensing data we use an airborne orthophoto with a spatial resolution of 50 cm and 4 spectral bands (Red, Green, Blue and Near Infrared) acquired in 2013. This data set was chosen on the one hand because it is comparable to current multispectral satellite images with very high spatial resolution (VHR; e.g. GeoEye, Pléiades or WoldView II–IV) and could therefore easily be substituted by them. On the other hand, these data are easily available for the whole of Italy via the Italian National Directory of Territorial Data (RNDT 2018).

## 3 Methodology

Two main building STs, i.e. masonry and RC, are dominating the study area. In consequence we address a binary classification problem to assign the buildings their respective ST by the use of remote sensing data automatically. To this end we implemented the following sequential procedure (Fig. 3).

On the basis of the building polygons and the remote sensing images (Sects. 2.1 and 2.2), relevant spatial and spectral image features determined in previous studies (e.g. Taubenböck et al. 2009; Geiß et al. 2014, 2015) have been calculated based on a multi-level concept (Sect. 3.1). These descriptive image features aim at capturing the individual buildings' physical characteristics as well as their spatial context properties and form the basis for modelling the relationship between the remote sensing data and in situ collected ST information using supervised machine learning. In a next step, the resulting classification models are used to estimate the STs for unseen buildings based on their associated feature manifestations.

Two different Machine Learning algorithms were tested for ST classification, RF and K-Nearest Neighbors (KNN). A RF-based feature selection procedure has been integrated

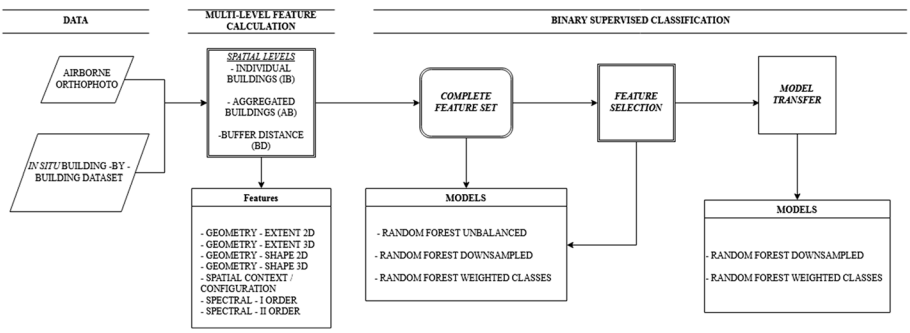


**Fig. 3** Overview on the framework and calculation steps carried out in this study

in the classification workflow in order to create a robust and considerably reduced subset of features (Sect. 3.2.1) which is convenient with regard to the computation time. The models were then re-learned using the selected features to provide a higher level of model transferability and rapid applicability. In addition, we use techniques to address misclassification problems related to class imbalance (Sect. 3.2.2). Finally, the resulting classifiers have been transferred to the entire Val d'Agri area. The classification maps have been evaluated in terms of confusion matrices and corresponding accuracy measures (Sect. 3.3).

### 3.1 Multi-level feature calculation

For an exhaustive and discriminative characterization of the individual buildings, a comprehensive state-of-the-art feature set including geometric, spatial context as well as spectral measures has been calculated based on three different spatial levels. Please refer to the "Appendix" for a detailed list of all deployed features, respective definitions and references.

#### 3.1.1 Spatial levels

To comprehensively take the buildings' urban environment into account, three different spatial levels (Fig. 4) have been considered each of them encoding distinctive spatio-contextual aspects: (1) the individual building (IB) represents the actual footprint of every single dwelling structure. Derived from the first, the other two levels describe the morphological setting in which the buildings are embedded in: (2) the aggregated buildings (AB) level involves the adjacent building footprints forming additional polygons. (3) The "buffer distance" (BD) level is the largest considered entity: a buffer of 10 m has been calculated around each building, then overlapping buffer polygons have been merged together and cut by using a layer of the main road network (available from OpenStreetMap). The distance of 10 m has been chosen since it has been shown appropriate to reflect the spatial arrangement of neighboring buildings in the considered area. In this manner the non-regular settlement structure of the villages where common spatial entities such as building blocks are not applicable was captured and used for encoding contextual information. The three information layers served as basis for the calculation of features.

#### 3.1.2 Geometry features

Geometry features relate to the two- and three-dimensional extent of buildings as well as to the description of their shape characteristics. The building heights can be derived from LIDAR or stereoscopic remote sensing data with high accuracies (e.g. Rottensteiner and Briese 2002; Wurm et al. 2011; Geiß et al. 2015; Leichtle et al. 2017). However, since such data was lacking in this study the analysis has been based on the in situ height information of the building inventory data base for an evaluation of the capability of these features. The height value information of each building contained in the database enabled the computation of 3D features such as volume, shape index 3D, Inverted Floor Area Ratio, average building volume contained in AB and BD polygons.

#### 3.1.3 Spectral features

The spectral characteristics are mainly used as indications of the material and layout of the roof area (Mueller et al. 2006). The mean and standard deviation values of the different

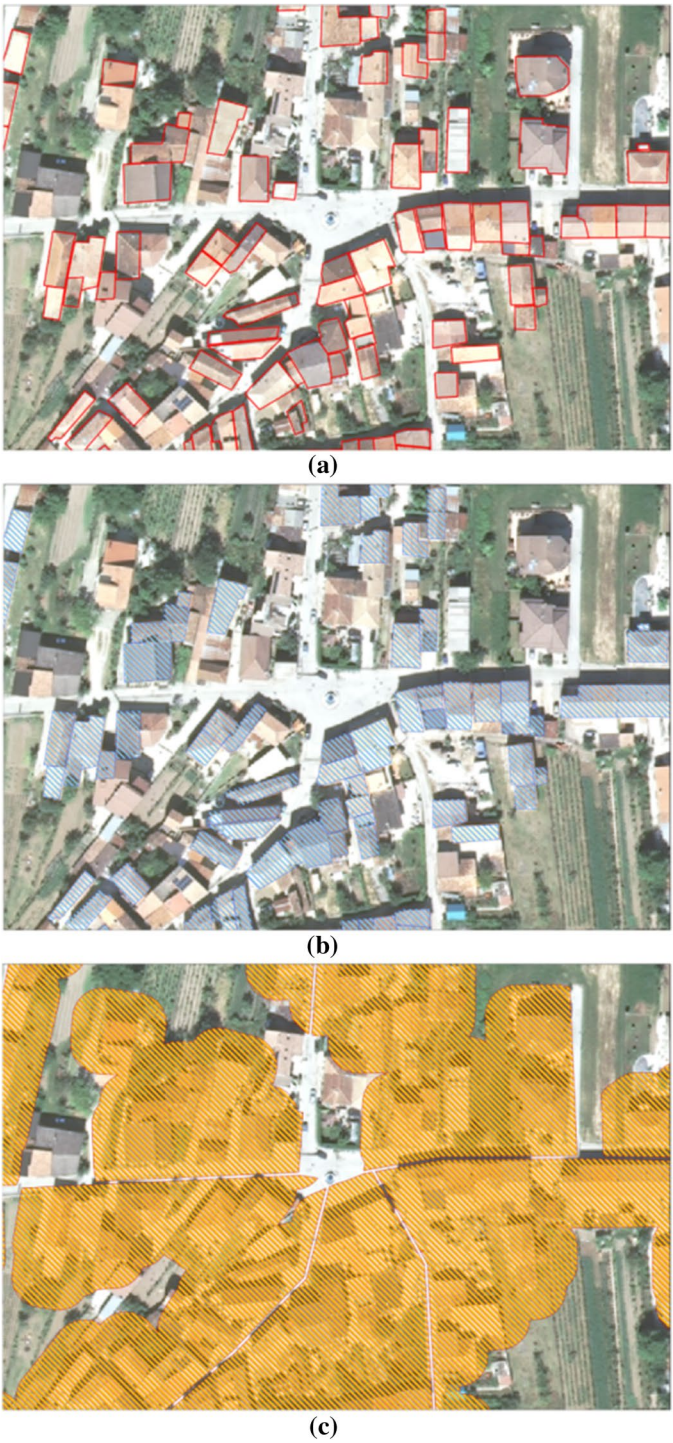


**Fig. 4** Spatial levels used as basis for the calculation of features. **a** Individual building (IB) level; **b** aggregated buildings (AB) level; **c** buffer distance (BD) level

image bands have been calculated in addition to the band ratios and index characteristics, which aim to highlight spectral variation (Geiß et al. 2014, 2015; Leichtle et al. 2017). Finally, rotation-invariant texture measures relying on the grey-level co-occurrence matrix (GLCM) (Haralick and Shanmugam 1973) have been derived since it was shown that such features can provide supplementary information if the spectral resolution is limited (Pacifici et al. 2009).

#### 3.1.4 Spatial features

In addition to spectral information (e.g. average values per band, median, minimum and maximum, average ratio, variance, etc.), building objects store spatial information related to location, neighborhood, typology, etc. (Blaschke 2010). To account for spatial context and configuration of IB and AB, orientation and geographical distance measures are introduced, among others. Other features whose purpose is to depict the spatial environment are calculated at BD level and consist of the area and average size of buildings.

In total, every building object is described by a 123-dimensional feature vector summarized in Table 2 (also compare "Appendix" for details), whereby 51 features are calculated based on the individual building footprints, 53 are calculated based on the aggregated buildings and 19 are calculated based on the buffer distance. Since the approach is flexible for considering more features, there is the opportunity for further investigating the contribution of more remote sensing features.

### 3.2 Supervised classification

RF (Breiman 2001) is an ensemble method combining a number of N decision trees to solve a single prediction problem by aggregating their results. This nonparametric classification approach was chosen to account for a considerable redundancy shown by the features, which can be critical for the estimation of statistics in parametric approaches (Geiß et al. 2015). RF is a robust and effective tool in prediction since it is based on the Law of Large Numbers (Feller 1968): features are used at each node to grow a tree and the forest chooses a class having the most out of N votes, for that case. In addition, RF is extremely advantageous to obtain a timely grasp of the relevance of the feature, especially if a feature selection is required (Géron 2017). In order to generate a RF classifier, two

**Table 2** List of features from different categories with number of features according to spatial scales

| Features | Number of features per spatial level | | |
|---|---|---|---|
| | IB | AB | BD |
| Geometry-extent 2D | 6 | 5 | 5 |
| Geometry-shape 2D | 15 | 15 | – |
| Geometry-extent 3D | 3 | 5 | 7 |
| Geometry-shape 3D | 1 | 2 | – |
| Spatial context/configuration | 3 | 3 | 7 |
| Spectral-1st order | 20 | 20 | – |
| Spectral-2nd order | 3 | 3 | – |
| $\sum$ | 51 | 53 | 19 |

*IB* individual building, *AB* aggregated buildings, *BD* buffer distance

hyperparameters are required: the number of features used at each node to generate a tree and the number of trees. In this study case, the first parameter has been set as the squared root of the number of features. The number of trees has been set to 500 since this was the value which provided the best accuracy in a set of experiments. Furthermore, numerous studies (e.g. Genuer et al. 2008) have shown that this amount of trees is enough to provide a reliable error estimate, while keeping computation times appropriate. The sorted attribute selection method is "entropy" rather than the typical Gini impurity, since this parameter has provided the greatest accuracy for our data set; in addition, Gini impurity is likely to separate the most populated class in its own tree branch, while entropy is prone to generating relatively more balanced trees (Géron 2017). For details on both measures we refer to Hastie et al. 2009.

K-Nearest Neighbors is a non-parametric algorithm, which considers all the available labeled data and classifies new data based on similarity measure. The key principle of this method is to assign an unclassified sample to the class involving the majority of the K nearest neighbors in the training set. During the training phase the feature vectors and the labels of the training samples are acquired. In the classification phase, the algorithm computes the K nearest neighbors of the query point (i.e. unlabeled data) and the class is assigned by voting among these neighbors (e.g. Hastie et al. 2009). This classifier has been chosen due to its simplicity and its robustness to noisy training data. The KNN algorithm requires only two parameters: K and the distance metric for obtaining significant accuracy. In this study, the KNN algorithm has been empirically tested for various numbers of neighbors and it has been found that $K=5$ produces the best result. The commonly employed Euclidean distance has been considered as distance metric.

### 3.2.1 Feature selection

In supervised learning problems based on high dimensional data as in this case, it is useful to decrease the number of features used by the model. When the best features are chosen, a good accuracy is achieved and less training data are required. Reducing the dimensionality of the feature vector is beneficial because it implies a reduction in time and data complexity to implement the model. Confronting high dimensionality typically implies the occurrence of negligible, redundant and noisy features that do not increase the accuracy of a predictive model or may effectively reduce it (Guyon and Elisseeff 2003; Aravena Pelizari et al. 2018).

In this study, feature selection is an essential element to implement a model based on few features enabling robust transferability. It is worth specifying that in this procedure we do not take into account any information from the target domain besides the output labels used for validation. This step of the work has been performed by ranking the features, according to their importance in the RF classification. This was done by using a function (i.e., "Feature importance") supplied by Scikit-learn, a free machine learning library for the Python programming language (Pedregosa et al. 2011). The feature importance function is based on the assumption that, looking at a single Decision Tree, important features supporting the ultimate choice of a broader portion of the input samples are likely to appear closer to the root of the tree, whereas less important features often turn out to be closer to the leaf. Therefore, it is possible to extract a features' importance by computing the average depth at which it shows up across all the trees in the forest. Scikit-learn calculates this for each feature after training (Géron 2017). The ten most important features have been chosen based on the results of this calculation.

### 3.2.2 Class imbalance

Machine learning applications are often affected by class imbalance, which emerges when a dataset contains many more samples from one class than from the other class(es). The uneven class distribution has a negative effect on the classifier performance, since if a class is represented by a small number of samples, it is difficult to uncover regularities and thus construct decision boundaries. Additionally, the classifiers are prone to prefer the most frequent class (Nguyen et al. 2008). The reason of these effects on classification is the fact that the algorithms are accuracy driven, therefore they try to minimize the overall error that is not much influenced by the minority class. In addition, they suppose a smooth distribution of data for all classes and errors to have similar cost for all classes (KrishnaVeni and Sobha Rani 2011; Ganganwar 2012). In order to reduce the incorrect classification, researchers have established several methods (Chawla et al. 2004; He and Garcia 2009). In this study, two different approaches have been implemented: the first is the downsampling of the largest class during the training step (Khalilia et al. 2011). Therefore, the frequency of masonry buildings has been decreased and conformed to that of the RC buildings considered to learn the model. The second approach envisages the change of a hyper-parameter in the RF model, in order to edit the weights related to errors from the two classes and to penalize more the erroneous classification on the less frequent class (Chen et al. 2004; Winham et al. 2013). If this parameter is not given, by default all classes are assumed to have weight 1.

## 3.3 Evaluation of the accuracy

For the evaluation of the accuracy we derive confusion matrices from the experiments described in Sect. 2. Confusion matrices oppose instances in a predicted class against instances in an actual class. Given "masonry" as the positive class and "RC" as the negative and considering True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN), measures of classification accuracy can be calculated from the confusion matrices (Table 3) as follows.

The first score is the overall accuracy (OA i.e. the total classification accuracy):

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Other metrics include recall (RE), related to commission errors (i.e. objects tagged as being part of the positive class but actually belonging to the negative class) and precision (PR, relevant to omission errors (i.e. objects labeled as negative but belonging to the positive class):

**Table 3** Confusion matrix for the binary classification problem of this study case

| | | *Reference data* | |
|---|---|---|---|
| | | 1 (positive = masonry) | 2 (negative = RC) |
| *Classification* | 1 (positive = masonry) | TP | FP |
| | 2 (negative = RC) | FN | TN |

$$RE = \frac{TP}{TP + FN} \tag{2}$$

$$PR = \frac{TP}{TP + FP} \tag{3}$$

The F1-score (F1) is calculated as the harmonic average of precision and recall, but does not consider the true negatives; it is higher for classifiers that have similar precision and recall:

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

Cohen´s κ is suitable to assess the performance of the classifier, although it is sensitive to imbalanced data (Jeni et al. 2013) (5); it varies between −1 and +1 and, where +1 indicates absolute agreement. The values of κ can be qualitatively interpreted according to the following categories (McHugh 2012):

- $\kappa \leq 0$ indicates no agreement;
- $0.01 < \kappa < 0.20$ indicates no to slight agreement;
- $0.21 < \kappa < 0.40$ indicates fair agreement;
- $0.41 < \kappa < 0.60$ indicates moderate agreement;
- $0.61 < \kappa < 0.80$ indicates substantial agreement;
- $\kappa > 0.80$ indicates almost perfect agreement.

$$\kappa \equiv \frac{OA - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+FN+TN)^2}}{1 - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+FN+TN)^2}} \tag{5}$$

*True Skills Statistics* (TSS, Allouche et al. 2006) has been also considered since it is widely insensitive to imbalanced classes (Klotz et al. 2016); it ranges from −1 to +1, where +1 indicates perfect agreement and values of zero or less indicate a performance no better than random:

$$TSS = \frac{TN}{FP + TN} + \frac{TP}{FN + TP} - 1 \tag{6}$$

## 4 Experimental setup

The experimental setting comprises the following steps: *first*, the classifier has been built based on one sample area (the village of Paterno). There, the building stock has been divided into two portions, one for training and the second for test purposes. *Second*, both classifiers—RF and KNN—have been learned using the complete set of 123 features. *Third*, the RF-based feature selection has been applied to identify the most important features and run the models again. *Fourth*, approaches involving downsampling and different weights for the two classes have been tested to overcome the class imbalance problem.

*Fifth*, the most accurate method has been transferred to all other villages for analyzing the capability for seismic building vulnerability assessment over large areas using earth observation data.

### 4.1 Splitting into training and testing areas

The 1844 buildings in the Paterno village, whose reference data are depicted in Fig. 5, have been spatially split into training and testing areas. Specifically, 20% of the buildings have been selected for testing the model. Each BD polygon, containing a different amount of buildings, has been randomly selected in order to prevent issues arising from sample selection bias and covariate shift (Geiß et al. 2017) resulting when the training set is not a casual population portion (Tuia et al. 2016). Furthermore, this has been done since all the elements contained in each BD have the same feature value for the BD-related features. In each realization of classification, a different sampling to constitute the test set is performed: single buildings within each BD polygon are randomly picked until 20% (circa 400 buildings) of the entire dataset is reached. The training examples consist of the remaining 80% of the buildings, obtained by removing the test set from the entire dataset.

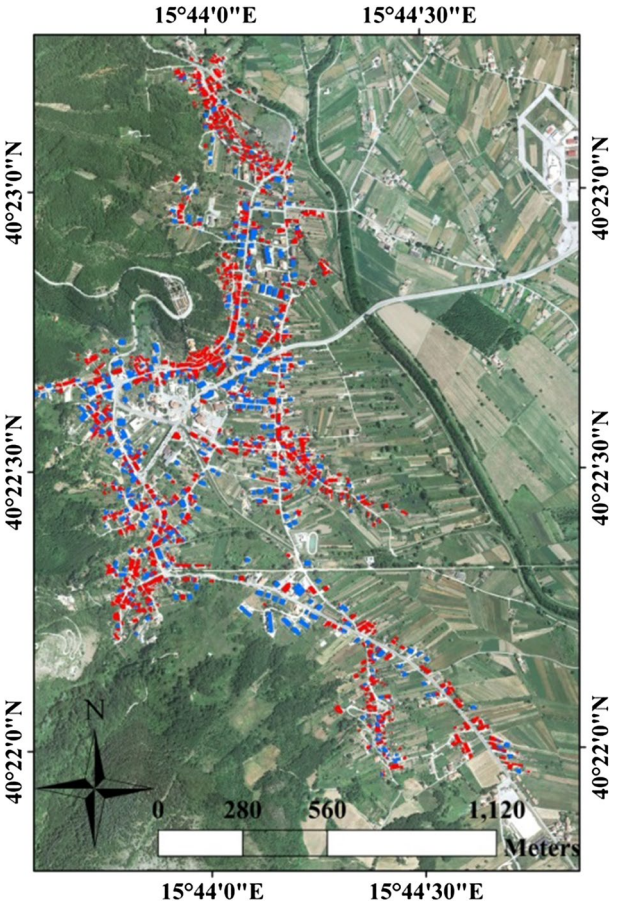


**Fig. 5** Reference data in Paterno village: labeled building polygons. Red: masonry; blue: RC

### 4.2 Learning the models

Representative for the entire Val d'Agri area, the building stock in the Paterno village is highly unbalanced (Table 1). With regard to their ST 1338 (73%) belong to the masonry class and 506 (27%) belong to the RC class. Therefore, in order to ensure an even distribution of training samples for the two classes when learning the model, the majority class has been decreased based on the procedure described in Sect. 3.2.2. In this way, 1212 buildings have been used to learn the model: 506 from the masonry class, 506 from the RC class. After the downsampling, 30 independent iterations of classification have been carried out, each for the RF as well as for the KNN approach. The implementation of the RF approach using the weighted classes approach allowed modifying the hyperparameter giving each class a weight. In the "balanced" mode weights are regulated inversely proportional to the class frequencies in the input data and the errors coming from the minority class are penalized more severely. In this study, the majority class (masonry) has been given weight 1, whereas the minority class (RC) was assigned weight 4. Such weights have been empirically obtained through prior experiments: classifications have been performed, based each time on random sample selection as explained in Sect. 4.1, while systematically varying the combination of weight values assigned to the two classes. The weights leading the model to achieve the best classification performance metrics have been considered to be used in the final model. After that, the RF model has been applied for 30 iterations using the complete set of features and 30 times adopting the selected features.

### 4.3 Feature selection

A RF-based feature selection procedure has been carried out together with the splitting of the data for 30 iterations. The random split has been performed for each iteration following the procedure described above. The complete set of features has been ranked according to their importance in the classification process. The 10 best features have been selected by considering their frequency in the ranking of the iterations.

### 4.4 Transferring the models

Since the results obtained for the Paterno area by using KNN were not satisfying, only the RF model has been applied to the rest of Val d'Agri area.

Therefore, the 10 best features calculated as described in Sect. 3.2.1 have been derived for the 15618 buildings in the remaining 17 villages, used for validating the model.

The building stock of the Paterno village has served as training dataset: when applying the downsampling approach described in Sect. 4.2 the selection of buildings used as sample for training is random. In this way a different number of training (and, as a consequence, of testing) examples is used; therefore, 30 iterations have been executed. Giving the classes a different weight enables exploiting the whole building stock in Paterno within the learning phase; hence, it has been implemented only once to classify the other 15,618 buildings.

## 5 Results

Results are reported as indicated in the following: Sect. 5.1 presents the results of the classification for the training area (Paterno). Section 5.2 describes the results of the feature importance ranking aimed at decreasing the dimensionality of the feature vector; Sect. 5.3 illustrates results of the classification of the entire Val d'Agri area.

### 5.1 Results in the training area

The accuracy scores obtained from the experiments illustrated in Sect. 4 are listed in Table 4. In general, the experiments reveal that the RF approach based on the complete set of features and the downsampled training dataset performs for the training area better than the KNN approach. With a mean OA of 0.90, a mean κ of 0.70 and a mean TSS of 0.66 the results prove the feasibility of the approach.

For the RF model deploying the downsampled training dataset and the selected features, we find only slightly lower accuracy scores than obtained using all the features (mean OA of 0.88, a mean κ of 0.69 and a mean TSS of 0.65). We consider this reduced approach meaningful since the results show that the decrease in the accuracy scores is minimal, whereas the improvement in terms of rapid applicability of the model is considerable. The RF approach based on the weighted classes approach and the complete feature set provides lower accuracy metrics (mean OA of 0.80, mean κ of 0.38 and mean TSS of 0.32). Nevertheless, such results are still adequate and suitable for applications (Congalton and Green 2008). The RF approach based on the weighted classes and using the ten best features obtains only slightly lower accuracy metrics (mean OA of 0.78; mean κ of 0.36 and mean TSS of 0.32).

In general, the KNN models perform consistently inferior than the RF models. Specifically, the OA in both cases is lower than 80% and κ shows only a fair agreement between prediction and reality (i.e., $\kappa < 0.4$; Sect. 3.3). For this reason KNN has not been used in transferring the model to the whole area.

The best models in the training area are RF based on the downsampled training data approach. The models based on the extended feature set reach higher accuracy values;

**Table 4** Classification accuracy statistics of RF and KNN based on the downsampled training dataset (left) and on the weighted classes approach (right) reported as mean and standard deviation (in brackets) from 30 iterations with a varying configuration of labeled samples on the training area (Paterno village)

| | Downsampled training data | | | | Weighted classes | |
|---|---|---|---|---|---|---|
| | RF | | KNN | | RF | |
| Accuracy statistics | All features | Selected features | All features | Selected features | All features | Selected features |
| RE | 0.96 (±0.01) | 0.95 (±0.01) | 0.88 (±0.05) | 0.90 (±0.02) | 0.66 (±0.03) | 0.66 (±0.04) |
| PR | 0.90 (±0.02) | 0.90 (±0.02) | 0.80 (±0.02) | 0.81 (±0.03) | 0.78 (±0.03) | 0.73 (±0.03) |
| OA | 0.90 (±0.02) | 0.88 (±0.02) | 0.75 (±0.03) | 0.78 (±0.02) | 0.80 (±0.02) | 0.78 (±0.03) |
| F1 | 0.93 (±0.02) | 0.92 (±0.01) | 0.84 (±0.03) | 0.85 (±0.02) | 0.68 (±0.04) | 0.68 (±0.04) |
| κ | 0.70 (±0.07) | 0.69 (±0.05) | 0.30 (±0.08) | 0.39 (±0.06) | 0.38 (±0.07) | 0.36 (±0.07) |
| TSS | 0.66 (±0.07) | 0.65(±0.064) | 0.28 (±0.07) | 0.36 (±0.06) | 0.32 (±0.07) | 0.32 (±0.07) |

nevertheless, all the accuracy metrics values obtained with the selected features are comparable with those obtained using all the features.

### 5.2 Results of the feature selection

In general, the feature selection analysis revealed that the 10 most important features refer to the geometrical extent and shape characteristics at the IB level. There is only one exception which is referring to the BD level. Specifically, the selected features are: area, perimeter, volume, length, width, fractal dimension, shape index 3D, height, number of pixels at IB level, and average built-up area at BD level. These most important features can be considered invariant: since they are geometry features, they are expected to be not influenced significantly by varying atmospheric situations or sensor properties and related spectral inhomogeneity—properties that are frequently related to spectral features. In addition, such features can be readily calculated, reducing the time cost related to the implementation of the classification model. Therefore, relying on these features is beneficial when transferring the model on unseen data. The importance ranking has been calculated for 30 classification iterations. The 10 best features, listed in Fig. 6, have been chosen depending on the frequency with which they were ranked among the 10 most important features.

### 5.3 Results of the model transfer

When the RF models are applied on the 15,618 buildings located in the remaining 17 villages resulting classification accuracy values (Table 5) are lower but still acceptable for application. Model transfer classification results of the two approaches for a selection of three different villages are visualized in Fig. 7 by means of the spatial distribution of correctly classified buildings (TPs and TNs) and misclassified buildings (FPs and FNs).

Comparing the accuracy values resulting from the transfer of RF model based on the downsampled training data and weighted classes RF model, all global accuracy measures consistently indicate that the latter performs better. Thereby the weighted class RF approach obtains an OA of 80%, a TSS value of 0.35 and a κ value of 0.31 expressing a fair agreement.

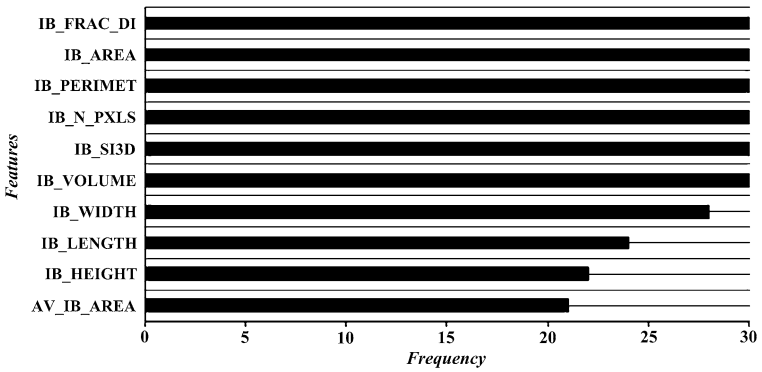


**Fig. 6** Outcome of the feature selection: the y-axis lists the ten best features; the x-axis shows the frequency with which each feature has been ranked among the best ten over 30 iterations

**Table 5** Results of RF based on the downsampled and weighted classes approaches on the 17 villages used for transferring the model. With regard to the downsampled training data RF approach mean and standard deviations from 30 iterations with a varying configuration of labeled samples are given. In case of the weighted classes RF approach all buildings were used to learn the final model to be transferred. Consequently only one a single value for each accuracy measure is given here

| Accuracy statistics | Downsampled training data RF | Weighted classes RF |
|---|---|---|
| RE | 0.66 (±0.01) | 0.64 |
| PR | 0.60 (±0.01) | 0.68 |
| OA | 0.60 (±0.03) | 0.80 |
| F1 | 0.56 (±0.02) | 0.65 |
| κ | 0.20 (±0.02) | 0.31 |
| TSS | 0.32 (±0.02) | 0.35 |

Figure 7 demonstrates that the accuracies for the single villages vary. While the weighted class RF classification result for Village 3 shows a substantial agreement with the reference data ($\kappa = 0.62$), the villages 2 and 1 show a moderate ($\kappa = 0.57$) and fair ($\kappa = 0.36$) agreement respectively. In general, however, spatial patterns of ST type distributions related to the settlement cores and outskirts are well captured.

## 6 Discussion

The potential of VHR remote sensing imagery for seismic building risk assessment has been the object of extensive research studies in the last decade. With VHR satellite images enabling covering large geographical areas in high geometric detail and with short revisiting time the capabilities have been shown. However, for seismic building vulnerability assessment some crucial information cannot directly be obtained from remote sensing images, but must be statistically inferred.

In order to produce classification maps useful for risk management applications, supervised classification methods are commonly deployed on remote sensing images. These algorithms are based on features describing the investigated classes for learning rules to discriminate between such classes. These rules, applied on a target domain, produce results having accuracies that are highly affected by the amount of data used to train the model. The retrieval of appropriate ground truth data for implementation and validation purposes requires large efforts in terms of time consumption and economic resources, as in the case of assessment of seismic building vulnerability. Therefore, in most areas in situ data are either lacking or are outdated and unreliable. The limited availability of in situ data is a challenge for classification problems, especially in the perspective of model transferability.

In this paper, the location of structural building typologies for the Val d'Agri area (South Italy) has been identified in order to assess seismic vulnerability. Our study is based on high detailed and extensive data sets and is thus not limited by data scarcity. We rely on an extensive, unique data set on buildings represented by geolocated polygons in the largest building-by-building geodatabase in Italy. Detailed structural knowledge has been collected during field surveys by teams of purposely trained engineers and stores very detailed physical characteristics of the constructions. Thus, for the first time ever, we have been able to systematically test such a remote sensing based approach for the complex morphological

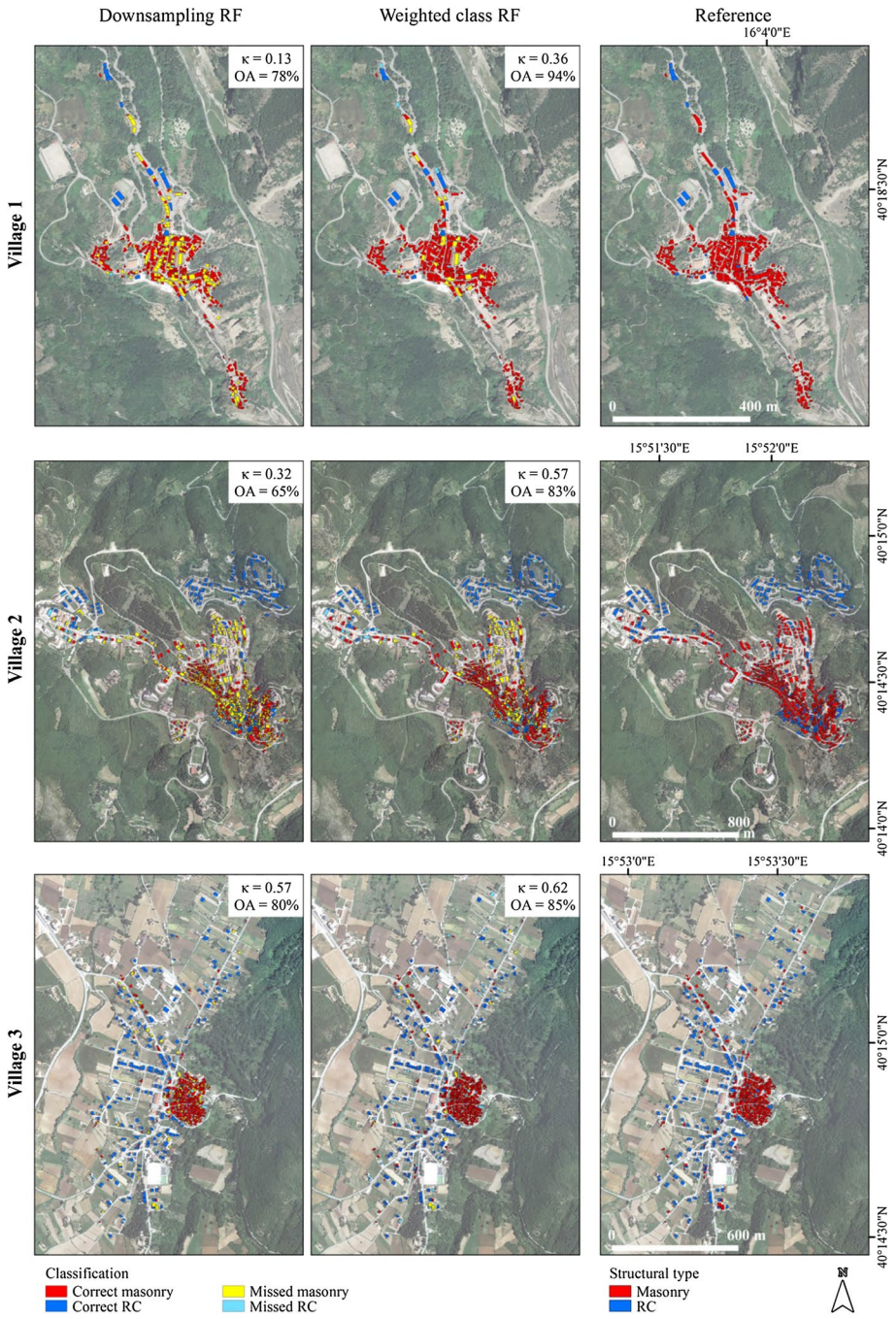


**Fig. 7** Classification maps obtained by transferring the models to three representative villages (left: downsampling approach; center: weighted class approach). Respective in situ reference data is shown on the right

situation in Europe. Its quality, detail level and reliability represent the added value of this research in providing solid contribution to the implementation of methods for image classification.

Due to the quantity of data used for validation, the developed models can be considered statistically robust as shown by the results presented in Sect. 5.

In detail we reveal, when transferring the models, the weighted class RF approach performing better than the RF approach based on downsampled training data. This is presumably due to the fact that the latter deploys less building inventory data in the training phase and consequently encodes less in situ information.

However, identifying in a more appropriate way such a percentage requires the effect in terms of estimated damage and expected consequences (fatalities, homeless people) to be verified. Percentages of misclassifications generating limited variations of the expected effects and/or conservative estimates of such effects may be considered acceptable. In this perspective, misclassifications of RC buildings should be considered more carefully: this building type is on average characterized by larger average volume—therefore, higher exposure—compared to masonry. Therefore, it can be expected to add more inaccuracies in the quantification of predicted impacts (Dolce et al. 2003; Masi et al. 2015). This implies that in practical applications the building volume should be taken into account as a weighting factor, since considering equally the individual buildings would lead to a wrong damage distribution map. The last component to be taken into consideration in determining the minimum level of accuracy of the classification necessary for the application of engineering DEMs is the intrinsic level of approximation of such models and other parameters used in the determination of seismic damage scenarios (Dolce et al. 2003).

The specific morphologic situation in the studied area presents a special distribution of the STs in only two classes; this may be considered representative of the Italian building stock with regard to villages. The discrimination of STs is essential in order to determine the expected seismic behavior of buildings; nevertheless, in the studied area such distinction is challenging due to the similar buildings' characteristics and the complex urban morphology. However, we find that the achieved results prove that the developed models still provide high accuracy; nevertheless, the decrease in the accuracy with respect to the transferred model calls for additional prior knowledge from the target domain or other areas in Italy, since the building STs are expected to be quite similar to those examined in this investigation. Furthermore, an important remark is relevant to the selected features: when only the 10 best features are deployed, the accuracy metrics show a slight decrease, whereas higher transferability is achieved. Such result expresses that a further step towards a rapid classification of building STs has been done, especially taking into account the selected geometrical features related to shape and extent.

## 7 Conclusion

Most previous studies on seismic vulnerability assessment of building structures tested their approaches in cities of the Global South and proved that remote sensing data allow for a large area assessment with acceptable accuracies. The building stock in large parts of Italy does not feature such a large morphological variety. This study, however, proves that also in this morphological situation where only two main STs are dominating the building

stock, remote sensing has the potential to be an appropriate method to contribute to fast and area-wide seismic vulnerability assessment of buildings.

The two STs, i.e. masonry and RC, were accurately identified using remote sensing data for the Val d'Agri area (Italy). Based on a very large in situ dataset we were able to learn the applied RF and KNN algorithms and validate them accordingly. With an extensive feature set (123 features at different spatial levels), it was proven that the two STs can be identified with viable accuracies in order to assess seismic building vulnerability, although other building features, among which age, height, are concurrently required to achieve this. Beyond, we found that 10 invariant geometric features allow obtaining good accuracy and are capable for transferring the model.

The dataset was affected by the class imbalance problem, with masonry being the majority class. Therefore, the models have been implemented adopting two different approaches: (i) downsampling the majority class, and (ii) changing the classifier internally. The model has been learned based on a small sample and then applied to the entire Val d'Agri area. Results show that the weighted classes approach produces better performance when RF is applied to the whole area (OA = 80%; $\kappa = 0.31$; TSS = 0.35). Therefore, the achieved results are encouraging and meaningful for practical applications. Although the considered area is representative of the building stock of Italian small villages, additional data can be encoded to render the model more invariant and further improve accuracy: further research may concern any other areas in Italy, or similar morphological situations in the Mediterranean countries.

## Appendix

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| *Geometry-extent 2D* | | |
| Area (*IB_AREA; AB_AREA*) | IB and AB polygon area | Steiniger et al. (2008), Hermosilla et al. (2014), Voltersen et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Length (*IB_LENGTH; AB_LENGTH*) | Length of the longer side of the minimum bounding rectangle comprising IB and AB polygons | Geiß et al. (2015) and Wurm et al. (2016) |
| Width (*IB_WIDTH; AB_WIDTH*) | Length of the shorter side of the minimum bounding rectangle comprising IB and AB polygons | Geiß et al. (2015) and Wurm et al. (2016) |
| Length/width (*IB_L_W; AB_L_W*) | Ratio between length and width of each IB and AB polygon | Steiniger et al. (2008), Voltersen et al. (2014), Geiß et al. (2015 and Wurm et al. (2016) |
| Perimeter (*IB_PERIMET; AB_PERIMET*) | IB and AB polygon perimeter | Hermosilla et al. (2014), Voltersen et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Inverted floor area ratio (*IB_IFAR*) | $\mathrm{IB} = \frac{polygon\ area}{polygon\ area \times number\ of\ floors}$ | Berger et al. (2013) |
| *BD* Area (*BD_AREA*) | Area of each BD polygon | Steiniger et al. (2008) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Average built-up area (*AV_IB_AREA*) | Average area of the IB polygons within each BD polygon | Yu et al. (2010) |
| Sum built-up area (*SUM_IB_AREA*) | Sum of the IB polygon areas within each BD polygon | Steiniger et al. (2008) |
| Average L/W IB (*AV_IB_L_W*) | Average length/width based on IB polygons within each BD polygon | Voltersen et al. (2014) |
| Average L/W AB (*AV_AB_L_W*) | Average length/width based on AB polygons within each BD polygon | Voltersen et al. (2014) |
| *Geometry-shape 2D* | | |
| Number of building vertices (building corners) (*IB_N_VERT, AB_N_VERT*) | Count of polygon points of the exterior ring | Steiniger et al. (2008) and Wurm et al. (2016) |
| Shape index (*IB_SI, AB_SI*) | Describes the smoothness of the outer shape of the polygon object. Calculation is based on the proportion between the real perimeter P of the polygon and an approximated square with the same area as the polygon $SI = \frac{P}{4\times\sqrt{A}}$ | Belgiu et al. (2014), Hermosilla et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Fractal dimension (*IB_FRAC_DI, AB_FRAC_DI*) | Provides a numerical description of the complexity and segmentation of a polygon by computing the proportion of area and perimeter $FD = \frac{\ln\left(\frac{U}{4}\right)^2}{lnA}$ | Hermosilla et al. (2014) and Wurm et al. (2016) |
| Density (*IB_DENS, AB_DENS*) | The distribution in space of the pixels of an image object, calculated by the number of pixels forming the image object divided by its approximated radius, based on the covariance matrix. (Trimble 2014) | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Border index (*IB_BI, AB_BI*) | $BI = \frac{perimeter}{2\times(length+width)}$ | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Compactness (*IB_COMP, AB_COMP*) | $COMP = \frac{length\times width}{number\ of\ pixels}$ | Hermosilla et al. (2014) and Belgiu et al. (2014) |
| Rectangular fit (*IB_RECT_FIT, AB_RECT_FIT*) | The calculation is based on a rectangle with the same area as the image object. The proportions of the rectangle are equal to the proportions of the length to width of the image object. The area of the image object outside the rectangle is compared with the area inside the rectangle. (Trimble 2014) | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Radius of largest enclosing ellipse (*IB_RLEE, AB_RLEE*) | The Radius of Largest Enclosed Ellipse feature describes how similar an image object is to an ellipse. The calculation uses an ellipse with the same area as the object and based on the covariance matrix. This ellipse is scaled down until it is totally enclosed by the image object. The ratio of the radius of this largest enclosed ellipse to the radius of the original ellipse is returned as feature value. (Trimble 2014) | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Radius of smallest enclosing ellipse (*IB_RSEE, AB_RSEE*) | The Radius of Smallest Enclosing Ellipse feature describes how much the shape of an image object is similar to an ellipse. The calculation is based on an ellipse with the same area as the image object and based on the covariance matrix. This ellipse is enlarged until it encloses the image object in total. The ratio of the radius of this smallest enclosing ellipse to the radius of the original ellipse is returned as feature value. (Trimble 2014) | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |
| Elliptic fit (*IB_EL_FIT, AB_EL_FIT*) | The Elliptic Fit feature describes how well an image object fits into an ellipse of similar size and proportions. While 0 indicates no fit, 1 indicates a perfect fit. The calculation is based on an ellipse with the same area as the selected image object. The proportions of the ellipse are equal to the length to the width of the image object. The area of the image object outside the ellipse is compared with the area inside the ellipse that is not filled by the image object | Belgiu et al. (2014), Geiß et al. (2015) |
| Roundness (*IB_ROUND, AB_ROUND*) | The Roundness feature describes how similar an image object is to an ellipse. It is calculated by the difference of the enclosing ellipse and the enclosed ellipse. The radius of the largest enclosed ellipse is subtracted from the radius of the smallest enclosing ellipse. (Trimble 2014) | Belgiu et al. (2014), Geiß et al. (2015) and Wurm et al. (2016) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Normalized perimeter index (*IB_NPI, AB_NPI*) | The NPI is the proportion of the perimeter of a circle with the same area as the building polygon ($PA_{circle}$) with the same perimeter of the building object $nPI = \frac{PA_{circle}}{P}$ | Wurm et al. (2016) |
| Normalized proximity index (*IB_NPRI, AB_NPRI*) | The proximity index is based on the calculation of Euclidian distances between single pixels of an object and the object center $nPrI = \frac{PA_{circle}}{PA_{object}}$ where the proximity of the circle is $PA_{circle} = \frac{2}{3} \times r_{A_{circle}}$ and $r_{A_{circle}}$ is the radius of the circle and for $P_{object}$ $P_{object} = \sum_{j=1}^{n} d_j \times \frac{1}{n}$ where $d$ is the Euclidian distance of the pixel to the object center | Wurm et al. (2016) |
| Normalized spin index (*IB_NSI, AB_NSI*) | The nSI (or moment of inertia) is similar to the proximity index, but the moment of inertia weights the extremities of the polygon higher: $nSI = \frac{J_{A_{circle}}}{J_{A_{object}}}$ the moment of inertia for the circle is: $J_{A_{circle}} = \frac{1}{2} r_{A_{circle}}^2$ and for the object: $J_{object} = \frac{1}{n} \times \sum_{j=1}^{n} d_j^2$ where $d_j^2$ equals the square Euclidian distance of the pixels to the object center | Wurm et al. (2016) |
| Areal asymmetry (*IB_ASYM, AB_ASYM*) | The asymmetry feature describes the relative length of an image object, compared to a regular polygon. An ellipse is approximated around a given image object, which can be expressed by the ratio of the lengths of its minor and the major axes. The feature value increases with this asymmetry | Wurm et al. (2016) |
| *Geomtry-extent 3D* | | |
| Building height (*IB_HEIGHT*) | For each IB polygon the height has been calculated as: *Average interstorey height* × *number of floors* × *area* | Hermosilla et al. (2014) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Maximum height AB *(AB_Max_HEIGHT)* | Building height of AB polygons calculated considering the maximum height value of the IB polygons forming each AB polygon | Hermosilla et al. (2014) |
| Area weighed maximum height AB *(AB_WEIGHT_A)* | Area-weighed height based on maximum heights of IB polygons within AB polygons | Hermosilla et al. (2014) |
| Building average height *(AB_MEAN_H)* | Building height of AB polygons calculated considering the average height value of the IB polygons forming each AB polygon | Hermosilla et al. (2014) |
| Area weighed mean height *(AB_MEAN_W_A)* | Area-weighed height based on the average height of IB polygons forming each AB polygon | Hermosilla et al. (2014) |
| Maximum height BD *(BD_MAX_HEIGHT)* | Maximum height value within *BD* polygon | Hermosilla et al. (2014) |
| Area-weighted maximum height *BD* (*BD_WEIGHT_A*) | Area-weighed height based on maximum heights of IB polygons within BD polygons | Hermosilla et al. (2014) |
| Mean of building height within *BD (AV_IB_HEIGHT)* | Average height of IB within BD polygons | Voltersen et al. (2014), Hermosilla et al. (2014) and Geiß et al. 2015 |
| Area-weighted on mean heights *BD* (*BD_AV_MA*) | Area-weighed height based on average heights of IB polygons within BD polygons | Yu et al. (2010) |
| Building volume (*IB_VOLUME*) | The building volume of the IB is calculated as: $IB_{VOLUME} = average\ floor\ area \times average\ interstorey\ height \times number\ of\ floors$ | Yu et al. (2010) |
| Maximum building volume based on maximum height (*AB_VOL_MAX_H*) | Maximum building volume of AB polygons within BD polygons based on maximum height of IB polygons | Hermosilla et al. (2014) |
| Mean building volume (*AB_AV_VOL*) | Average building volume of AB polygons within BD polygons based on maximum height of IB polygons | Hermosilla et al. (2014) |
| *Geometry-shape 3D* | | |
| IB Shape index 3D (*SI_3D*) | Describes the smoothness of the object in three dimension; it is calculated by the proportion of the real perimeter with the approximated perimeter of a cube with the same volume than the real object $SI_{3D} = \frac{P}{4 \times \sqrt[3]{V}}$ | Wurm et al. (2016) |
| Shape index 3D_maxHeight (*AB_SI3D_MAX_H*) | Shape index 3D of AB polygons based on area-weighted maximum heights of *IB* | Wurm et al. (2016) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Shape index 3D_meanHeight (*AB_SI3D_AV_H*) | Shape index 3D of AB polygons based on area-weighed mean heights of *IB* | Wurm et al. (2016) |
| *Spatial context/configuration* | | |
| Orientation (*IB_ORIENT, AB_ORIENT*) | Orientation of the major axis of the minimum bounding rectangle | Belgiu et al. (2014) and Geiß et al. (2015) |
| Distance to nearest building (*IB_DIST, AB_DIST*) | The shortest distance between buildings | Xie et al. (2015) |
| St. Dev. IB orientation (*IB_STDV_OR*) | Standard deviation of the IB polygons orientation within BD polygons | Cheriyadat (2014) |
| St. Dev. AB orientation (*STDV_AB_OR*) | Standard deviation of the AB polygons orientation within BD polygons | Gil et al. (2012) |
| Mean nearest IB (*IB_MEAN_NN*) | Mean distance to the nearest building based on IB polygons | Yu et al. (2010) |
| Mean of distances to nearest building based on *AB* (*AB_MEAN_NN*) | Mean distance to the nearest building based on AB polygons | Yu et al. (2010) |
| *IB* in *BD* (*IB_COUNT*) | Number of IB polygons within each BD polygon | Steiniger et al. (2008) |
| *AB* in *BD* (*AB_COUNT*) | Number of AB polygons within each BD polygon | Steiniger et al. (2008) |
| Mean IFAR (*AB_MEAN_IFAR, BD_ MEAN_IFAR*) | Mean of inverted floor area ratio based on *IB* polygons within AB and BD polygons | Voltersen et al. (2014) |
| Building aggregation measure (*BD_BA*) | $BA = \frac{\frac{A_b}{A_{AOI}}}{Median(D_b)} - \frac{Median(IFAR)}{N_B}$ where $A_b$, floor area covered by buildings; $A_{AOI}$, area of interest; $D_b$, distance to the nearest building; *IFAR*, Inverted Floor area Ratio; $N_B$, number of buildings | Berger et al. (2013) |
| Normalized building aggregation measure (*BD_NBA*) | $NBA = \frac{BA_i - BA_{min}}{BA_{max} - BA_{min}}$ where $BA_i$, BA in the i-th BD polygon; $BA_{max}$, maximum BA value; $BA_{min}$, minimum BA value | Berger et al. (2013) |
| *Spectral-1st order* | | |
| Mean blue (*IB_sMeanB, AB_sMeanB*) | Mean intensity in the blue channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| Mean green (*IB_sMeanG, AB_sMeanG*) | Mean intensity in the green channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| Mean red (*IB_sMeanR, AB_sMeanR*) | Mean intensity in the red channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| Mean NIR (*IB_sMeanNIR, AB_ sMeanNIR*) | Mean intensity in the NIR channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| St. Dev. blue (*IB_sSTDV_B, AB_ sSTDV_B*) | Standard deviation of the intensity in the blue channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| St. Dev. green (*IB_sSTDV_G, AB_ sSTDV_B*) | Standard deviation of the intensity in the green channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| St. Dev. red (*IB_sSTDV_R, AB_ sSTDV_R*) | Standard deviation of the intensity in the red channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| St. Dev. NIR (*IB_sSTDV_NIR, AB_sSTDV_NIR*) | Standard deviation of the intensity in the NIR channel | Bruzzone and Carlin (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| Blue/green (*IB_sB/G, AB_ sB/G*) | Mean blue/mean green | Bruzzone and Carli (2006), Geiß et al. (2015) and Leinenkugel et al. (2011) |
| Blue/red (*IB_sB/R, AB_ sB/R*) | Mean blue/mean red | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Blue/NIR (*IB_sB/NIR, AB_ sB/NIR*) | Mean blue/mean NIR | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Green/red (*IB_sG/R, AB_ sG/R*) | Mean green/mean red | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Green/NIR (*IB_sG/NIR, AB_ sG/NIR*) | Mean green/Mean NIR | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Red/NIR (*IB_sR/NIR, AB_ sR/NIR*) | Mean red/mean green | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Normalized green (*IB_sG/GRNIR, AB_ sG/GRNIR*) | Mean(green)/[mean(green) + mean(red) + mean(NIR)] | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Normalized red (*IB_sR/GRNIR, AB_ sR/GRNIR*) | Mean(red)/[mean(green) + mean(red) + mean(NIR)] | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Normalized NIR (*IB_sNIR/GRNIR, AB_ sNIR/GRNIR*) | Mean(NIR)/[mean(green) + mean(red) + mean(NIR)] | Bruzzone and Carlin (2006) and Geiß et al. (2015) |
| Brightness (*IB_sBRIGHT, AB_ sBRIGHT*) | Mean intensity of all channels | Bruzzone and Carlin (2006), Geiß et al. (2015) and Voltersen et al. (2014) |
| Normalized differenced vegetation index (*IB_sNDVI, AB_ sNDVI*) | $NDVI = \frac{NIR-red}{NIR+red}$ | Geiß et al. (2015) and Leinenkugel et al. (2011) |
| Soil-adjusted vegetation index (*IB_sSAVI, AB_ sSAVI*) | $SAVI = \frac{NIR-red}{NIR+red+L} \times (1+L)$ where L = 0.5 | Bruzzone and Carlin (2006) and Geiß et al. (2015) |

| Features | Description | References: studies on characterization of built environments |
|---|---|---|
| *Spectral-2nd order* | | |
| $GLCM_{inv.}$ (Angular 2nd moment) (*IB_GLCM_ANG, AB_GLCM_ANG*) | $\sum_{i,j=0}^{N-1} (Pi,j)^2$ where $i$, row number; $j$, column number; $P_{i,j}$, normalized value in the cell $i$, $j$; $N$ is the number of rows or columns | Geiß et al. (2015) and Zhang et al. (2006) |
| $GLCM_{inv.}$ (entropy) (*IB_GLCM_ENT, AB_GLCM_ENT*) | $\sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j})$ where $i$, row number; $j$, column number; $P_{i,j}$, normalized value in the cell $i$, $j$; $N$, number of rows or columns | Geiß et al. (2015) and Zhang et al. (2006) |
| $GLCM_{inv.}$ (homogeneity) (*IB_GLCM_HOM, AB_GLCM_HOM*) | $\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$ where $i$, row number; $j$, column number; $P_{i,j}$, normalized value in the cell $i$, $j$; $N$, number of rows or columns | Geiß et al. (2015) and Zhang et al. (2006) |

List of the state-of-the-art features deployed to describe the building stock in the present study

## References


Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43(6):1223–1232

Aravena Pelizari P, Spröhnle K, Geiß C, Schoepfer E, Plank S, Taubenböck H (2018) Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements. Remote Sens Environ 209:793–807

Balkaya C, Casciati F, Casciati S, Faravelli L, Vece M (2015) Real-time identification of disaster areas by an open-access vision-based tool. Adv Eng Softw 88:83–90

Belgiu M, Tomljenovic I, Lampoltshammer T, Blaschke T, Höfle B (2014) Ontology-based classification of building types detected from airborne laser scanning data. Remote Sens 6(2):1347–1366

Berger C, Voltersen M, Eckardt R, Eberle J, Heyer T, Salepci N, Hese S, Schmullius C, Tao J, Auer S, Bamler R (2013) Multi-modal and multi-temporal data fusion: outcome of the 2012 GRSS data fusion contest. IEEE J Sel Top Appl Earth Obs Remote Sens 6(3):1324–1340

Blaschke T (2010) Object based image analysis for remote sensing. ISPRS J Photogramm Remote Sens 65(1):2–16

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Bruzzone L, Carlin L (2006) A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans Geosci Remote Sens 44(9):2587–2600

Calvi GM, Pinho R, Magenes G, Bommer JJ, Restrepo-Vélez LF, Crowley H (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. ISET J Earthq Technol 43(3):75–104

Casciati S, Chen ZC, Faravelli L, Vece M (2016) Synergy of monitoring and security. Smart Struct Syst 17(5):743–751

Chawla NV, Japkowicz N, Kotcz A (2004) Special issue on learning from imbalanced data sets. ACM SIGKDD Explor Newsl 6(1):1–6

Chen C, Liaw A, Breiman L (2004) Using random forest to learn imbalanced data. Univ Calif Berkeley 110:1–12

Cheriyadat AM (2014) Unsupervised feature learning for aerial scene classification. IEEE Trans Geosci Remote Sens 52(1):439–451

Chiauzzi L, Masi A, Mucciarelli M, Vona M, Pacor F, Cultrera G, Gallovič F, Emolo A (2012) Building damage scenarios based on exploitation of Housner intensity derived from finite faults ground motion simulations. Bull Earthq Eng 10(2):517–545

Congalton RG, Green K (2008) Assessing the accuracy of remotely sensed data: principles and practices. CRC Press, Boca Raton

Dolce M, Masi A, Marino M, Vona M (2003) Earthquake damage scenarios of the building stock of Potenza (Southern Italy) including site effects. Bull Earthq Eng 1(1):115–140

Fardis MN (2009) Seismic design, assessment and retrofitting of concrete buildings: based on EN-Eurocode, vol 8. Springer, Berlin

Feller W (1968) The strong law of large numbers. In: Feller W (ed) An introduction to probability theory and its applications, vol 1(3). Wiley, New York, pp 243–245

Ganganwar V (2012) An overview of classification algorithms for imbalanced datasets. Int J Emerg Technol Adv Eng 2(4):42–47

Geiß C, Taubenböck H, Tyagunov S, Tisch A, Post J, Lakes T (2014) Assessment of seismic building vulnerability from space. Earthq Spectra 30(4):1553–1583

Geiß C, Aravena Pelizari P, Marconcini M, Sengara W, Edwards M, Lakes T, Taubenböck H (2015) Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. ISPRS J Photogramm Remote Sens 104:175–188

Geiß C, Jilge M, Lakes T, Taubenböck H (2016) Estimation of seismic vulnerability levels of urban structures with multisensor remote sensing. IEEE J Sel Top Appl Earth Obs Remote Sens 9(5):1913–1936

Geiß C, Aravena Pelizari P, Schrade H, Brenning A, Taubenböck H (2017) On the effect of spatially non-disjoint training and test samples on estimated model generalization capabilities in supervised classification with spatial features. IEEE Geosci Remote Sens Lett 14(11):2008–2012

Genuer R, Poggi JM, Tuleau C (2008) Random forests: some methodological insights. arXiv preprint arXiv:0811.3619

Géron A (2017) Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media Inc, Newton

Gil J, Beirão JN, Montenegro N, Duarte JP (2012) On the discovery of urban typologies: data mining the many dimensions of urban form. Urban Morphol 16(1):27

Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3(March):1157–1182

Haralick RM, Shanmugam K (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 6:610–621

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference and prediction. Springer, New York, 763 pp

He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 9:1263–1284

Hermosilla T, Palomar-Vázquez J, Balaguer-Beser Á, Balsa-Barreiro J, Ruiz LA (2014) Using street based metrics to characterize urban typologies. Comput Environ Urban Syst 44:68–79

INGV – Instituto Nazionale di Geofisica e Vulcanologia (2018) Database of individual seismogenic source. DISS version 3. http://diss.rm.ingv.it/diss/. Accessed 23 Feb 2017

Jeni LA, Cohn JF, De La Torre F (2013) Facing imbalanced data-recommendations for the use of performance metrics. In 2013 humaine association conference on affective computing and intelligent interaction. IEEE, pp 245–251

Khalilia M, Chakraborty S, Popescu M (2011) Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 11(1):51

Klotz M, Kemper T, Geiß C, Esch T, Taubenböck H (2016) How good is the map? A multi-scale cross-comparison framework for global settlement layers: evidence from Central Europe. Remote Sens Environ 178:191–212

KrishnaVeni CV, Sobha Rani T (2011) On the classification of imbalanced datasets. IJCST 2(SP1):145–148

Leichtle T, Geiß C, Wurm M, Lakes T, Taubenböck H (2017) Unsupervised change detection in VHR remote sensing imagery: an object-based clustering approach in a dynamic urban environment. Int J Appl Earth Obs Geoinf 54:15–27

Leinenkugel P, Esch T, Kuenzer C (2011) Settlement detection and impervious surface estimation in the Mekong Delta using optical and SAR remote sensing data. Remote Sens Environ 115(12):3007–3019

Masi A (2003) Seismic vulnerability assessment of gravity load designed R/C frames. Bull Earthq Eng 1(3):371–395

Masi A, Vona M (2012) Vulnerability assessment of gravity-load designed RC buildings: evaluation of seismic capacity through non-linear dynamic analyses. Eng Struct 45:257–269

Masi A, Santarsiero G, Nigro D (2013) Cyclic tests on external RC beam-column joints: role of seismic design level and axial load value on the ultimate capacity. J Earthq Eng 17(1):110–136
Masi A, Chiauzzi L, Samela C, Tosco L, Vona M (2014) Survey of dwelling buildings for seismic loss assessment at urban scale: the case study of 18 villages in Val D'agri, Italy. Environ Eng Manag J (EEMJ) 13(2):471–486
Masi A, Digrisolo A, Manfredi V (2015) Fragility curves of gravity-load designed RC buildings with regularity in plan. Earthq Struct 9(1):1–27
McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med Biochem Med 22(3):276–282
Mueller M, Segl K, Heiden U, Kaufmann H (2006) Potential of high-resolution satellite data in the context of vulnerability of buildings. Nat Hazards 38(1–2):247–258
Nguyen GH, Bouzerdoum A, Phung SL (2008) A supervised learning approach for imbalanced data sets. In 2008 19th international conference on pattern recognition. IEEE, pp 1–4
Pacifici F, Chini M, Emery WJ (2009) A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. Remote Sens Environ 113(6):1276–1292
Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12(October):2825–2830
Penazzi D, Valluzzi MR, Cardani G, Binda L, Baronio G, Modena C (2000) Behaviour of historic masonry buildings in seismic areas: lessons learned from the Umbria-Marche earthquake. In 12th international brick-block masonry conference, pp 217–235
Rottensteiner F, Briese C (2002) A new method for building extraction in urban areas from high-resolution LIDAR data. In International archives of photogrammetry remote sensing and spatial information sciences, vol 34, no 3/A, pp 295–301. Natural Resources Canada
RNDT – Repertorio Nazionale dei Dati Territoriali (2018) Geoportal. http://geodati.gov.it/geoportale/eng/. Accessed 23 Feb 2017
Sarabandi P, Kiremidjian A, Eguchi RT, Adams BJ (2008) Building inventory compilation for disaster management: application of remote sensing and statistical modeling. Technical Report Series MCEER-08-0025, Buffalo, MCEER
Steiniger S, Lange T, Burghardt D, Weibel R (2008) An approach for the classification of urban building structures based on discriminant analysis techniques. Trans GIS 12(1):31–59
Taubenböck H, Post J, Roth A, Zosseder K, Strunz G, Dech S (2008) A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing. Nat Hazards Earth Syst Sci 8(3):409–420
Taubenböck H, Roth A, Dech S, Mehl H, Münich JC, Stempniewski L, Zschau J (2009) Assessing building vulnerability using synergistically remote sensing and civil engineering. In: Krek A, Rumor M, Zlatanova S, Fendel E (eds) Urban and regional data management. Taylor & Francis Group, London, pp 287–300
Tuia D, Persello C, Bruzzone L (2016) Domain adaptation for the classification of remote sensing data: an overview of recent advances. IEEE Geosci Remote Sens Mag 4(2):41–57
Voltersen M, Berger C, Hese S, Schmullius C (2014) Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level. Remote Sens Environ 154:192–201
Winham SJ, Freimuth RR, Biernacka JM (2013) A weighted random forests approach to improve predictive performance. Stat Anal Data Min ASA Data Sci J 6(6):496–505
Wurm M, Taubenböck H, Schardt M, Esch T, Dech S (2011) Object-based image information fusion using multisensor earth observation data over urban areas. Int J Image Data Fusion 2(2):121–147
Wurm M, Schmitt A, Taubenböck H (2016) Building types' classification using shape-based features and linear discriminant functions. IEEE J Sel Top Appl Earth Obs Remote Sens 9(5):1901–1912
Xie Y, Weng A, Weng Q (2015) Population estimation of urban residential communities using remotely sensed morphologic data. IEEE Geosci Remote Sens Lett 12(5):1111–1115
Yu B, Liu H, Wu J, Hu Y, Zhang L (2010) Automated derivation of urban building density information using airborne LiDAR data and object-based method. Landsc Urban Plan 98(3–4):210–219
Zhang L, Huang X, Huang B, Li P (2006) A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. IEEE Trans Geosci Remote Sens 44(10):2950–2961