Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 44 (2023) 83–90

Structural Integrity Procedia

www.elsevier.com/locate/procedia

XIX ANIDIS Conference, Seismic Engineering in Italy

Clustering of empirical damage data for the vulnerability classification of the Italian residential building stock

Annalisa Rosti[a]*, Maria Rota[b], Andrea Penna[a,b]

[a] *University of Pavia, via Ferrata 3, Pavia 27100, Italy*
[b] *Eucentre Foundation, via Ferrata 1, Pavia 27100, Italy*

Abstract

This paper proposes an innovative data-driven vulnerability model for the classification of the existing residential building stock, by clustering observational damage data gathered after the 2009 L'Aquila earthquake. The proposed model preserves the conceptual framework at the basis of the macroseismic approach, which allows for a thorough vulnerability classification of the built environment by resorting to vulnerability classes and by accounting for the uncertain association of building typologies to vulnerability classes. Novel aspects of this study are the adoption of peak ground acceleration for the ground motion characterisation, which allows for overcoming possible limitations related to the use of macroseismic intensity, and the use of unsupervised machine learning techniques for removing subjectivity in the definition of vulnerability classes. A probabilistic framework is then set up allowing for the attribution of a given building typology to multiple vulnerability classes, based on an ad-hoc strategy, involving the use of probability theory and using empirically-derived typological fragility functions as a target. The use of a detailed post-earthquake survey form also allows for an improved definition of building types representative of the Italian building stock.

Keywords: Seismic vulnerability; Fragility curves; Unsupervised machine learning techniques; Post-earthquake damage data; L'Aquila earthquake

---

* Corresponding author.
E-mail address: annalisa.rosti@unipv.it

10.1016/j.prostr.2023.01.012

1. Introduction

Besides fragility functions, which provide the expected distribution of damage in the different damage states as a function of the experienced ground shaking, an exhaustive vulnerability model should also supply indications on the vulnerability classification of the existing building stock. In this context, macroseismic approaches (e.g. Lagomarsino and Giovinazzi 2006; Bernardini et al. 2011) allow for a thorough classification of the seismic vulnerability of the built environment, by resorting to the six vulnerability classes of the EMS98 (Grünthal et al. 1998) and considering the uncertain association of building types to vulnerability classes. In spite of their mathematical refinement and methodological consistency, application of macroseismic approaches for seismic vulnerability and risk applications (e.g. da Porto et al. 2021; Dolce et al. 2021) requires approximate and uncertain laws for correlating macroseismic intensity values and peak ground motion parameters (e.g. Bernardini et al. 2011). Seismic input is indeed defined in terms of macroseismic intensity, which is a descriptive parameter affected by the characteristics, and therefore by the vulnerability, of the built environment.

In accordance with the conceptual framework of the macroseismic method (Lagomarsino and Giovinazzi 2006), this paper proposes an innovative model for the seismic vulnerability classification of the existing building stock, based on a data driven approach. Differently from macroseismic approaches, the peak ground acceleration is employed for characterizing the ground motion severity at the sites of damage observations. Seismic damages detected on Italian residential buildings hit by the 2009 L'Aquila earthquake are clustered via unsupervised machine learning techniques, allowing for the objective identification of vulnerability classes. An ad-hoc strategy, resorting to probability theory and using empirically-derived typological fragility curves as a target, is specifically built up to account for the uncertain attribution of building types to vulnerability classes.

2. Processing of the post-earthquake damage database

The proposed vulnerability model relies on statistical processing and clustering of Italian post-earthquake damage data available from the Da.D.O. web-gis platform (Dolce et al. 2019). Selection of the L'Aquila post-earthquake database is motivated by the significant number of inspected buildings and of municipalities completely-surveyed, identified by a completeness ratio (i.e. ratio of the number of surveyed buildings and the total number of residential buildings evaluated from national census data, ISTAT 2001) exceeding 90% (Rosti et al. 2021a, b). Furthermore, use of the L'Aquila damage database allows for suitably characterizing the negative evidence of damage in the territories less affected by the earthquake shaking, permitting to avoid bias in the subsequent fragility assessment. Following these operations, the post-earthquake dataset collects damage data of 37'406 residential buildings, then integrated by 197'528 undamaged buildings sited in the Abruzzi non-surveyed and partially-surveyed (with completeness ratio lower than 10%) municipalities (Rosti et al. 2022).

2.1. Adopted building taxonomy

Residential buildings are allocated to 42 building typologies, identified based on the main building attributes retrievable from the post-earthquake survey form. Typological classification of RC buildings accounts for both the level of seismic design (i.e. buildings with seismic design pre- and post-1981) and number of stories (i.e. 1, 2, 3, 4 and ≥ 5 stories). Masonry buildings are classified based on the number of stories (i.e. 1, 2, 3 and ≥ 4 stories), quality and layout of the masonry fabric (i.e. IRR: irregular layout or poor-quality; REG: regular layout and good-quality), in-plane stiffness of intermediate diaphragms (i.e. F: flexible; R: rigid) and presence (or lack) of connecting devices, such as tie-rods and/or tie-beams (i.e. CD: with connecting devices; NCD: without connecting devices). The considerable level of detail of the adopted typological classification system aims at identifying possible similarities or differences in the empirical seismic vulnerability of the exposed building stock, driven by the presence/lack of specific constructive details. Fig. 1 depicts the typological classification of the residential building stock, in terms of construction material (a), masonry type (b) and number of stories (c, d), with reference to the L'Aquila completely-surveyed municipalities.

2.2. Definition of seismic input

In line with the objective of the study, seismic input is represented by the peak ground acceleration (PGA) (e.g. Rosti et al. 2020a), extrapolated from the recently released INGV ShakeMap (Michelini et al. 2020), accounting for latest ground motion models, an updated V_{s30} map for local site effects and the latest USGS-ShakeMap version 4 (v.4) software (Worden et al. 2020).

2.3. Damage classification

Discrete damage states are defined in accordance with the EMS-98 (Grünthal et al. 1998), by using existing literature damage rules for suitably converting information on structural (Rota et al. 2008) and non-structural (Del Gaudio et al. 2017) damage available from the post-earthquake survey form. After evaluating damage on individual building components, a global level of damage is associated with each building, based on the maximum level of damage (e.g. Rota et al. 2008; Del Gaudio et al. 2017; Rosti et al. 2018). Damage classification of masonry and RC buildings is depicted in Fig. 1 (e) and (f), respectively, with reference to completely-surveyed municipalities.

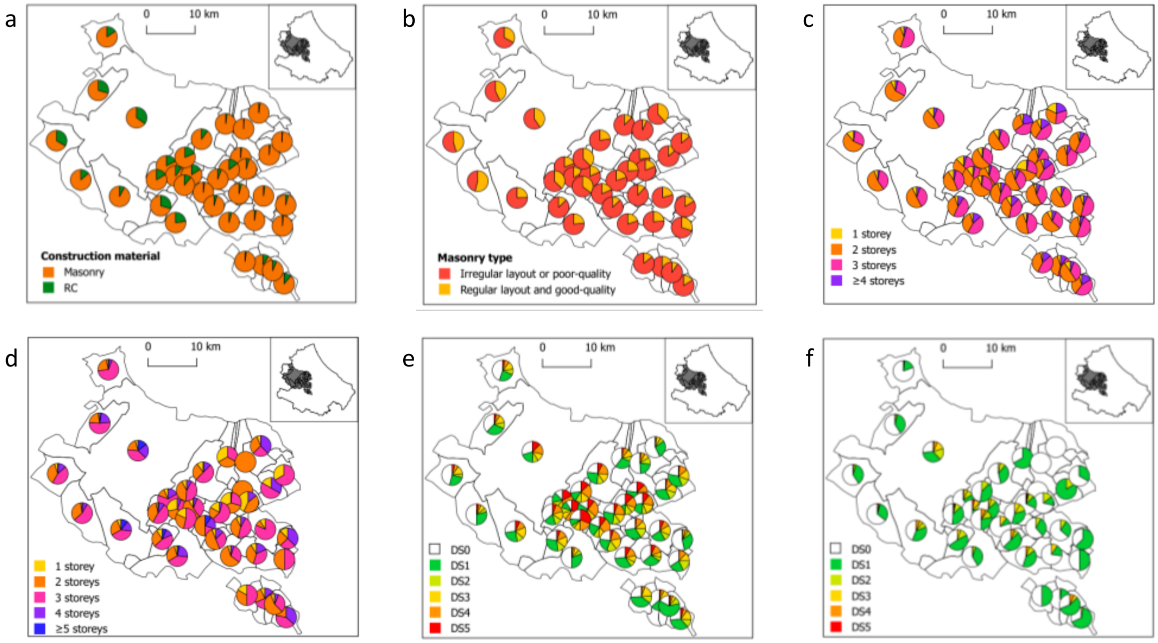

Fig. 1. Typological classification of the residential building stock for L'Aquila completely-surveyed municipalities (Rosti et al. 2022): (a) construction material; (b) masonry type; subdivision of (c) masonry and (d) RC buildings based on the number of stories. Damage distribution of (e) masonry and (f) RC buildings.

2.4. Seismic fragility assessment

Fragility curves are derived for predefined building typologies by fitting a suitable statistical model on observational data points. In line with existing studies (e.g. Ioannou et al. 2021; Rosti et al. 2021a, b), the multinomial distribution is employed for approximating the repartition of buildings in the different damage states, given the ground motion severity. The probability of exceedance of a given level of damage, as a function of the ground shaking, is described by the cumulative lognormal distribution (e.g. Rota et al. 2008; Del Gaudio et al. 2017; Ader et al. 2020; Rosti et al. 2020b). Fragility functions are simultaneously fitted on all damage levels and building typologies via the maximum likelihood estimate (MLE) approach, by enforcing a common dispersion value (e.g. Karababa and Pomonis 2011; Ader et al. 2020; Rosti et al. 2021a, b).

Probabilities of occurrence of the different damage states, resulting from the corresponding typological fragility functions, are then combined to get the mean level of damage, μ_D, as a function of the ground motion severity (e.g. Braga et al. 1982; Lagomarsino and Giovinazzi 2006). The reader is addressed to Rosti et al. (2022) for details on the adopted statistical model and fitting technique and for collection of the parameters of resulting typological fragility curves.

3. Identification of vulnerability classes by clustering of observational damage data

Vulnerability classes are identified by applying unsupervised machine learning techniques, overcoming possible subjectivity in the attribution of some choices. Empirically-derived mean damage data are allocated to multiple clusters (i.e. the vulnerability classes) with different membership degree by fuzzy c-means (FCM) clustering (Bezdek 1981). In line with the EMS-98, six vulnerability classes of decreasing vulnerability (from A to F) are considered. Depending on the construction material (i.e. masonry and RC), vulnerability classes are then split into two subgroups, to account for the different distance among damage levels observed in the typological fragility curves. Six vulnerability classes (i.e. A1, B1, C1, D1, E1, F1) are defined in case of masonry, whereas four out of six vulnerability classes (i.e. C2, D2, E2, F2) are considered in case of RC buildings, for which higher vulnerability classes (i.e. classes A2 and B2) lack. Following the implementation of FCM clustering, empirical mean damage data points are attributed to the most likely vulnerability class and to the other vulnerability classes with different membership degree. Sets of lognormal fragility curves are derived for each vulnerability class (Fig. 2). Details on the adopted statistical procedure and parameters of the cumulative lognormal fragility curves of vulnerability classes can be found in Rosti et al. (2022).

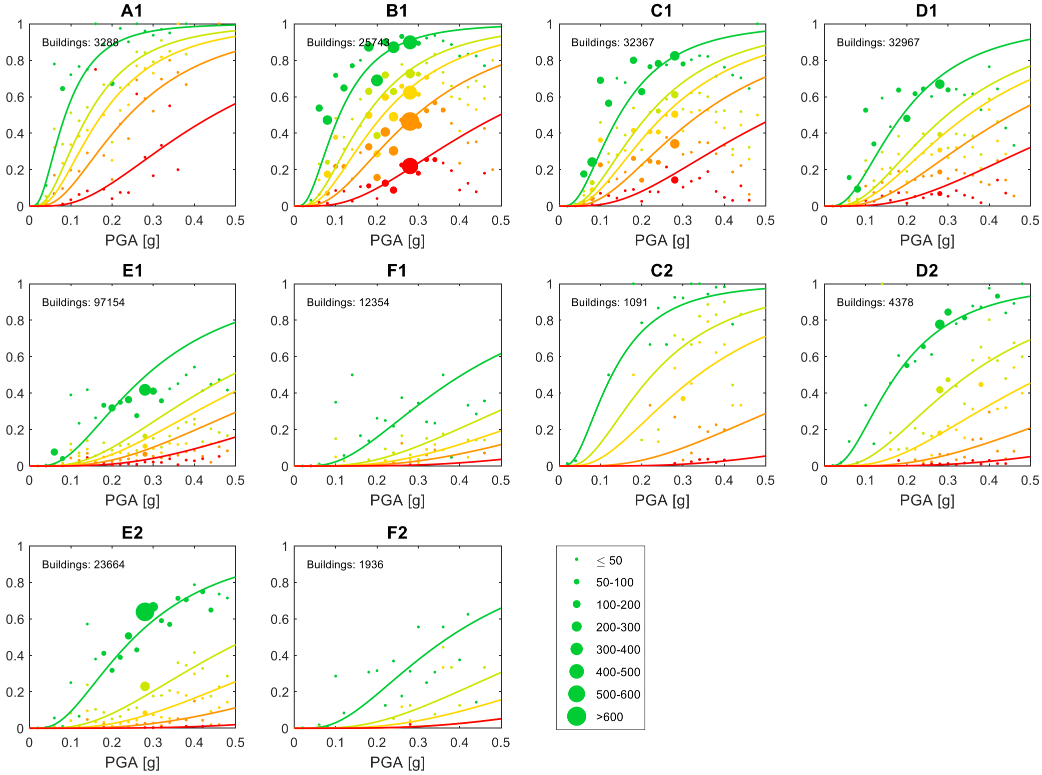

Fig. 2. Fragility functions of the vulnerability classes identified based on FCM clustering of observational mean damage values of masonry (subscript 1) and RC (subscript 2) building typologies (Rosti et al. 2022).

4. Vulnerability classification of the exposed building stock

Predefined building typologies are mapped to multiple vulnerability classes, allowing for a thorough vulnerability classification of the exposed building stock. The fragility curves of the vulnerability classes are linearly combined by means of the w_{jk} coefficients, representing the degrees of belonging of the j^{th} building typology to the k^{th} vulnerability class. The trend of the w_{jk} coefficients is approximated by the binomial model, with the advantage of describing the entire w_{jk} distribution by a single parameter (e.g. Rota and Rosti 2017; Rosti and Rota 2017). Two binomial distributions, one for "masonry" and the other one for "RC" vulnerability classes, are specifically defined and jointly used. Each binomial distribution is suitably scaled to account for the different weight that it takes in the global w_{jk} distribution. Based on this strategy, the fragility curve of damage level DS_i of the j^{th} building typology can be approximated as:

$$\Phi\left[\frac{\log(PGA/\theta_{DS_{ij}})}{\beta}\right] \approx c_{j,mas} \sum_{k_{mas}=0}^{5}\left[\frac{5!}{k_{mas}!(5-k_{mas})!}\left(\frac{y_{j,mas}}{5}\right)^{k_{mas}}\left(1-\frac{y_{j,mas}}{5}\right)^{5-k_{mas}}\right]\Phi\left[\frac{\log(PGA/\theta_{DS_{ik_{mas}}})}{\beta}\right] + (1 - c_{j,mas}) \sum_{k_{RC}=0}^{3}\left[\frac{3!}{k_{RC}!(3-k_{RC})!}\left(\frac{y_{j,RC}}{3}\right)^{k_{RC}}\left(1-\frac{y_{j,RC}}{3}\right)^{3-k_{RC}}\right]\Phi\left[\frac{\log(PGA/\theta_{DS_{ik_{RC}}})}{\beta}\right] \quad (1)$$

where c_{mas} is the scaling coefficient accounting for the weight that the "masonry" binomial distribution assumes in the global w_{jk} distribution, y_{mas} and y_{RC} are the binomial parameters of the "masonry" and "RC" binomial distributions, k_{mas} varies from 0 (F1) to 5 (A1), whereas k_{RC} ranges from 0 (F2) to 3 (C2).

The global deviation between the sets of approximating and target typological fragility functions is then minimized to obtain optimal values of the unknowns (i.e. y_{mas}, y_{RC} and c_{mas}).

Fig. 3 shows the outcome of the abovementioned procedure, with specific application to RC building types. Comparison of approximating (continuous lines) and target (dashed lines) fragility functions demonstrates the suitability of the adopted strategy. Results show that the evolution of building code enhances the seismic response of RC buildings. RC buildings seismically-designed based on obsolete (pre-1981) seismic provisions are indeed more vulnerable than the corresponding ones seismically-designed after 1981. Given the design level, impact of the building height on the seismic vulnerability of RC buildings is also significant. This finding prompts the need and relevance of accounting for number of stories in the seismic vulnerability classification of existing buildings. The reader is addressed to Rosti et al. (2022) for further details on the adopted strategy and for collection of the parameters (i.e. y_{mas}, y_{RC} and c_{mas}) necessary for modelling the uncertain attribution of each building type to vulnerability classes.

Analogously to the EMS-98, implementation of the proposed methodological framework to all predefined building types results in a vulnerability table (Fig. 4), enabling the vulnerability classification of the existing building stock based on essential building attributes. In the figure, black and white squared markers indicate the weighted mean "masonry" and "RC" vulnerability class, respectively, fully characterizing the uncertain distribution of building types to vulnerability classes. Grey and white bars respectively indicate the fraction of "masonry" and "RC" binomial distributions to be considered in the entire distribution of the membership degrees.

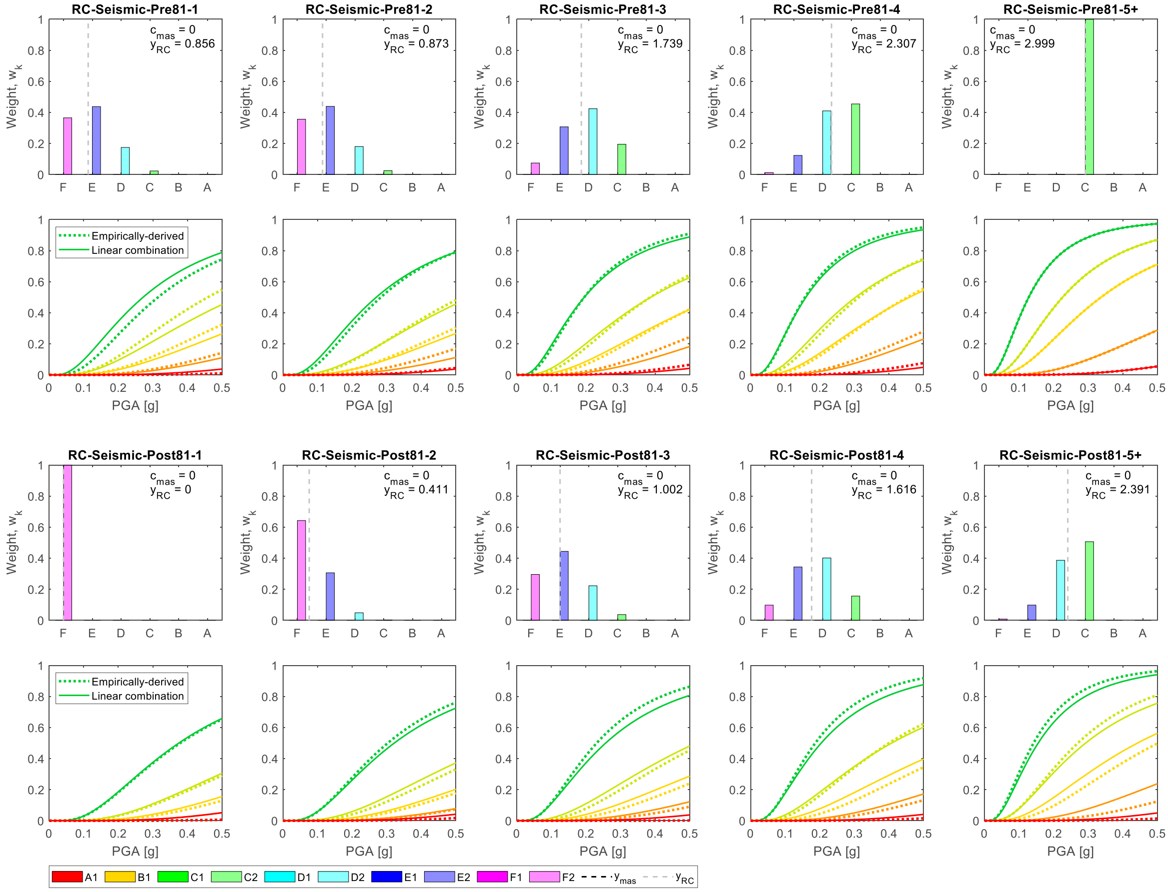

Fig. 3. Degrees of belonging of RC building types to vulnerability classes (Rosti et al. 2022).

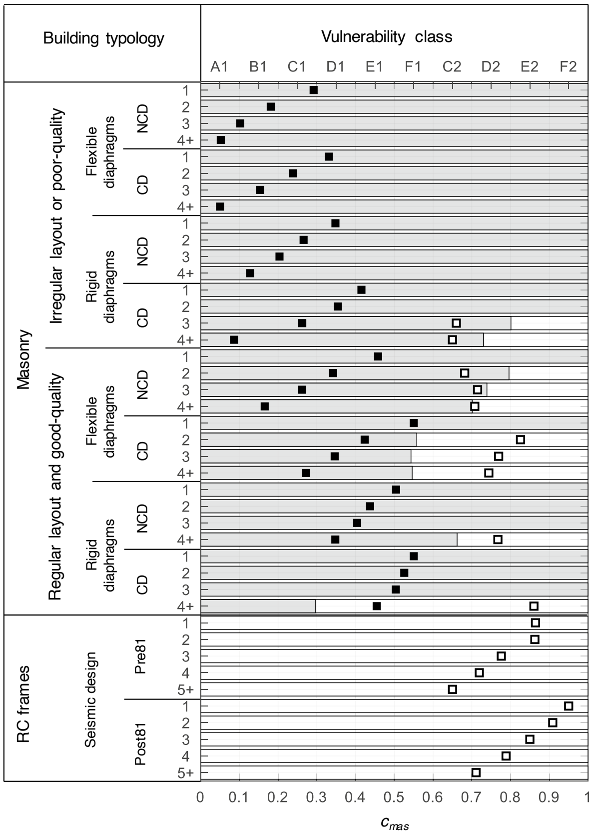

Fig. 4. Proposed vulnerability table for masonry and RC building types (Rosti et al. 2022).

5. Conclusions

This study proposes an exhaustive vulnerability model for the classification of the existing residential building stock, by clustering observed seismic damages detected in the aftermath of the 2009 L'Aquila seismic event. The proposed model allows for describing the seismic vulnerability of existing buildings by empirical fragility curves defined for vulnerability classes and for classifying the vulnerability of the exposed building stock by selected building attributes (e.g. vertical and horizontal structures, number of stories, structural details and design level). Original contributions of this work are the adoption of unsupervised machine learning techniques for the objective identification of vulnerability classes and the use of the peak ground acceleration for seismic input characterization. An ad-hoc procedure, grounding on probability theory and targeting empirical typological fragility functions, is developed for modelling the uncertain attribution of building types to vulnerability classes. For each building type, the weighted mean vulnerability class is provided, under the assumption of binomial distribution. In line with the EMS-98, a vulnerability table provides a synthetic representation of the seismic vulnerability of buildings, identified based on selected structural attributes. In this context, the availability of a robust post-earthquake database (Dolce et al. 2019), gathering both damage and typological information, allows for an improved definition of building typologies representative of the Italian built environment.

Acknowledgements

This work was developed with the partial financial support of the Italian Department of Civil Protection, within the 2022-2024 DPC-ReLUIS Research Projects WP4 "MARS – Seismic Risk Maps", which is gratefully acknowledged. The Authors also would like to sincerely thank Dr. Massimiliano Stucchi for the fruitful discussions on the topic and for his very useful suggestions on the work.

References

Ader, T., Grant, D.N., Free, M., Villani, M., Lopez, J., Spence, R., 2020. An unbiased estimation of empirical lognormal fragility functions with uncertainties on the ground motion intensity measure. J Earthq Eng, 24(7):1115–1133.

Bernardini, A., Lagomarsino, S., Mannella, A., Martinelli, A., Milano, L., Parodi, S., 2011. Forecasting seismic damage scenarios of residential buildings from rough inventories: a case-study in the Abruzzo Region (Italy). Proc IMech E Part OJ Risk Reliab, 224: 279–296.

Bezdek, J.C., 1981. Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York.

Braga, F., Dolce, M., Liberatore, D., 1982. A statistical study on damaged buildings and an ensuing review of the M.S.K.-76 scale. Proceedings of the 7th European Conf on Earthq Eng, Athens, Greece.

Census Data 2001. http://dawinci.istat.it/jsp/MD/dawinciMD.jsp.

da Porto, F., Donà, M., Rosti, A., Rota, M., Lagomarsino, S., Cattari, S., Borzi, B., Onida, M., De Gregorio, D., Perelli, F.L., Del Gaudio, C., Ricci, P., Speranza, E., 2021. Comparative analysis of the fragility curves for Italian residential masonry and RC buildings Bull Earthq Eng, 19: 3209-3252.

Del Gaudio, C., De Martino, G., Di Ludovico, M., Manfredi, G., Prota, A., Ricci, P., Verderame, G.M., 2017. Empirical fragility curves from damage data on RC buildings after the 2009 L'Aquila earthquake. Bull Earthq Eng, 15(4): 1425-1450.

Dolce, M., Speranza, E., Giordano, F., Borzi, B., Bocchi, F., Conte, C., Di Meo, A., Faravelli, M., Pascale, V., 2019. Observed damage database of past Italian earthquakes: the Da.D.O. WebGIS. Boll Geofis Teor Appl, 60(2): 141-164.

Dolce, M., Prota, A., Borzi, B., da Porto, F., Lagomarsino, S., Magenes, G., Moroni, C., Penna, A., Polese, M., Speranza, E., Verderame, G.M., Zuccaro, G., 2021. Seismic risk assessment of residential buildings in Italy. Bull Earthq Eng, 19: 2999-3032.

Grünthal, G. (ed.), Musson, R.M.W., Schwarz, J., Stucchi, M., 1998. European Macroseismic Scale. Cahiers du Centre Européen de Géodynamique et de Séismologie, Vol. 15 - European Macroseismic Scale 1998. European Centre for Geodynamics and Seismology, Luxembourg.

Ioannou, I., Bertelli, S., Verrucci, E., Arcidiacono, V., Rossetto, T., 2021. Empirical fragility assessment of residential buildings using data from the Emilia 2012 sequence of earthquakes. Bull Earthq Eng, 19: 1765– 1795.

Karababa, F.S., Pomonis, A., 2011. Damage data analysis and vulnerability estimation following the August 14, 2003 Lefkada Island, Greece, Earthquake. Bull Earthq Eng, 9: 1015–1046.

Lagomarsino, S., Giovinazzi, S., 2006. Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. Bull Earthq Eng, 4: 415–443.

Michelini, A., Faenza, L., Lanzano, G., Lauciani, V., Jozinovic, D., Puglia, R., Luzi, L., 2020. The New ShakeMap in Italy: Progress and Advances in the Last 10 Yr. Seismol Research Letters, 91(1): 317-333.

Rosti, A., Rota, M., 2017. Comparison of PSH results with historical macroseismic observations at different scales. Part 2: application to South-East France. Bull Earthq Eng 15(11): 4609-4633.

Rosti, A., Rota, M., Penna, A., 2018. Damage classification and derivation of damage probability matrices from L'Aquila (2009) post-earthquake survey data. Bull Earthq Eng, 16(9): 3687-3720.

Rosti, A., Rota, M., Penna, A., 2020a. Influence of seismic input characterisation on empirical damage probability matrices for the 2009 L'Aquila event. Soil Dyn Earthq Eng. https://doi.org/10.1016/j. soildyn.2019.105870.

Rosti, A., Del Gaudio, C., Di Ludovico, M., Magenes, G., Penna, A., Polese, M., Prota, A., Ricci, P., Rota, M., Verderame, G.M., 2020b. Empirical vulnerability curves for Italian residential buildings, Boll Geofis Teor Appl, 61(3): 357-374.

Rosti, A., Rota, M., Penna, A., 2021a. Empirical fragility curves for Italian URM buildings. Bull Earthq Eng, 19: 3057-3076.

Rosti, A., Del Gaudio, C., Rota, M., Ricci, P., Di Ludovico, M., Penna, A., Verderame, G.M., 2021b. Empirical fragility curves for Italian residential RC buildings. Bull Earthq Eng, 19: 3165-3183.

Rosti, A., Rota, M., Penna, A., 2022. An empirical seismic vulnerability model. Bull Earthq Eng, https://doi.org/10.1007/s10518-022-01374-3.

Rota, M., Penna, A., Strobbia, C.L., 2008. Processing Italian damage data to derive typological fragility curves. Soil Dyn Earthq Eng, 28(10): 933-947.

Rota, M., Rosti, A., 2017. Comparison of PSH results with historical macroseismic observations at different scales. Part 1: methodology. Bull Earthq Eng 15(11):4585–4607

Worden, C.B., Thompson, E.M., Hearne, M., Wald, D.J., 2020. ShakeMap Manual Online: technical manual, user's guide, and software guide, U. S. Geological Survey. http://usgs.github.io/shakemap/. DOI: https://doi.org/10.5066/F7D21VPQ.