

“DEVELOPING SEISMIC INTENSITY MAPS FROM TWITTER DATA; THE CASE STUDY OF LESVOS, GREECE 2017 EARTHQUAKE: ASSESSMENTS, IMPROVEMENTS AND ENRICHMENTS ON THE METHODOLOGY”

Stathis G Arapostathis¹, Efthimios Lekkas², Kostas Kalabokidis³, Gavriil Xanthopoulos⁵, Giorgos Drakatos⁴, Nafsika Spirou²,
Ioannis Kalogeras⁴

¹Department of Geography, Harokopio University, Athens, Greece

²Department of Geology, University of Athens, Athens, Greece

³Department of Geography, University of the Aegean, Mytilene, Greece

⁴Institute of Geodynamics of the National Observatory of Athens, Greece

⁵Institute of Mediterranean Forest Ecosystems and Forest Products Technology, Athens, Greece

KEYWORDS: seismic intensity, twitter, volunteered geographic information, VGI

ABSTRACT

This article presents an effort to validate and further improve a previously published innovative approach for drawing macroseismic intensity maps from data extracted from sources of volunteered geographic information (VGI). Our approach involves classification of macroseismic observations (extracted from social media sources) to values of the EMS 98 intensity scale, leading to the drawing of isoseismal maps. The earthquake of June 12th, 2017 (Mw 6.3) that occurred off the south coast of Lesvos Island, Greece, was used as a case study; its main shock was located at depth of about 13 km. This specific event, which claimed the life of a woman and caused at least 15 injuries due to collapsing buildings and falling debris (mainly in the town of Vrissa), was chosen for the specific geomorphological characteristics of the meizoseismal area, time of occurrence and distribution of damage. Twitter was chosen as a VGI source mostly for reasons of consistency with the original published work, generating comparable findings that can be assessed more readily to facilitate further development of the methodology. Results of the dataset analysis include the drawing of the isoseismal maps from Tweets published within different time periods (6h, 12h, 24h, 48h); and the identification of various text patterns regarding the evaluation of the macroseismic observations that result into intensity values. The present work offers additional empirical evidence regarding the validity of the methodology presented in the scientific literature, and further enriches it by providing additional text patterns and specific improvements related to the classification of the information in certain values of seismic intensity. Assessment of the results is enriched by the progress that has been noted in the field and has been presented in the international scientific literature since 2016.

INTRODUCTION

The importance of seismic intensity maps is very well known among the scientific community that is specialized in the management of earthquake events. Through them, scientists can extract useful information regarding the intensity of an earthquake event, can identify interesting trends of the spatial patterns and can also assess and compare the intensities of different earthquakes within the same or similar geographic regions. Concerning macroseismic observations, there are various ways of collecting them, and a plethora of ways for mapping them.

Until 2015 the most common ways for collecting macroseismic observations were by the distribution of questionnaires usually through post, e-mail, telephone, radio, TV, or by instant distribution (Cecic I. and Musoon R., 2004). Various community approach initiatives had also appeared prior to 2015, such as the “Did you feel it” Project of the United States Geological Survey (USGS) or the online macroseismic questionnaires of the European Mediterranean Seismological Center (EMSC). Similar approaches can be found in other important national Seismological Institutes including the National Observatory of Athens, Greece (NOA). In these initiatives, the citizens can report what they have exactly felt during the earthquake.

Moreover, and as technology evolves, other approaches have been developed especially in the part of collecting macroseismic observations. The most interesting ones are those which make use of drones (Antoniou et al 2017, Yamazaki et al 2015). The results of the related studies are impressive and include 3D representation of the area that is captured through a drone system, and the most accurate collection of spatial data.

However, even the drone approach requires time to setup the necessary equipment and natural presence in the area in which the earthquake event occurs.

In 2016, a different approach for creating seismic intensity maps was introduced, based on the extraction of macroseismic observations from social media (Arapostathis et al. 2016). Moreover, Kropivnitskaya et al. (2016) presented a hybrid research for the creation of seismic intensity maps, using both physical and social sensors. It is a marvelous work which however considers the use of tweets as a full-filling component, limiting the research as far as this specific part is concerned only to the geolocated tweets. In general, these approaches constitute a new, challenging research field, which aspires to create seismic intensity maps in real time, without even the need to have natural presence in the area in which an earthquake event occurs. Taking also into consideration the fact that the production of information through the social media has an increasing rhythm, it is assessed that these kind of approaches is possible to replace the conventional ones, in near future.

In this research article, an effort to improve various parts of the methodology that was published in 2016 is presented. Specifically, we are applying the method, as it is described in the next section, in a seismic event that has some characteristics that no other case study had. That earthquake event is the one of Lesvos island, Greece 2017 (ML = 6.1).

MAIN BODY

CASE STUDY

The earthquake event of Lesvos (Greece, 2017 ML=6.1) occurred on the 12th of June 2017 at 12:28 GMT, in a geographic area of about 15 km south of the SE coast of Lesvos Island. The damage in parts of the island was catastrophic and had also caused the death of a woman (in the local town of Vrissa) and at least 15 people were injured from the catastrophic damage of the surrounding premises (Antoniou et al. 2017, Lekkas et al. 2017)

At least 12 villages had serious damages, while according to an official report more than 1,000 premises were assessed as unsafe for use. Various seismic results of the earthquake were reported even in the Turkish coast. Regarding the physical environment, various ground cracks and slope movements were reported (Antoniou et al. 2017) along with a tsunami at the port of Plomari town.

From a social media perspective, the earthquake event has also some interesting characteristics. It took place during the touristic period in Greece and Turkey, so it is assumed that the information is to be enriched by the tweets of many tourists that were possibly touring in the general area during the earthquake event. Finally, the earthquake occurred during the day, with all people being awoken, resulting to a big number of tweets that is expected to be posted almost instantly.

DATASET

The dataset used was acquired from the data provider sifter and contained 80,020 tweets. This is the total number of the tweets that were published from the 12th of June 00:00:01 GMT until the 14th of June 23:59:59 GMT and contain at least one of the following keywords: earthquake (in Greek, Latin, transliterated Greek and English), Mytilene (in Greek and English), Vrissa (in Greek and English) and the words disasters, residents, panic, houses, emergency and situation in Greek.

The cost of the dataset was about 135 US dollars. Moreover, the dataset contains few geolocated tweets, a slight difference since the last relevant acquisition that was performed during 2016. In any case, the geolocated tweets are excluded from this research as according to literature, the location from which a tweet was published does not

necessarily reflect the location of a macroseismic observation; i.e. someone could observe something while moving and publish it afterwards (Huiji and Barbier 2011).

DESCRIPTION OF THE METHODOLOGY

The basic principles of the methodology are consisted in few basic parts. The first part is related to the preparation of the dataset that is needed to be analyzed. Various filters were applied to isolate the tweets that are written in Greek and in English. The next parts include the selection of the tweets that are relevant to the earthquake event of Lesvos, and further to the selection of those that include macroseismic observations and geographic reference. Afterwards, the tweets that fulfill the above criteria are classified into values of the EMS 98 macroseismic scale. The classification process is based on various criteria, mostly based on the official description of the scale (Grunthal et al. 1998). Various text-patterns that had been identified in previously published research (Arapostathis et al. 2015, 2016) are applied automatically, improving thus the process in terms of speed. Moreover, during the analysis, new text- patterns are identified and added to the table (Table I).

After the classification process, the tweets are geo-referenced, by adding the geographic coordinates of the centroid of the geographic area in which they are referred in the texts. Next, they are imported in a Geographic Information System (GIS) software environment as layers, and various geo-processing techniques are applied. These techniques include the randomization of the spatial distribution of the macroseismic observations and the interpolation of the data by using the kriging method that is the most applicable method when the data are related to physical events (Schenková et al. 2005). The kriging interpolation is applied many times in different sub-sets of the data, leading thus, to the development of different seismic intensity maps. These sub-sets are consisted of tweets published within different time periods (up to 48h) and in two spatial precision groups. The first one, contains tweets that their geographical reference is up to a precision score 2 (medium sized Greek city such as Plomari), and the second one, contains tweets that their geographical reference is up to a level of a modern municipality (such as Lesvos island).

STATISTICS

The total number of tweets checked was 56,170 (both in English and Greek). From those, 1,038 tweets in the Greek language and 2,420 in the English language contained information related to macroseismic observations and a geographical reference (about 6.15%). 1,764 of those (about 50%) were published within the first 6 hours. The maximum value that a macroseismic observation received was 11 (XI)

and the minimum was 2 (II). In total, 764 observations receive a precision score up to 2 and 2117 receive score I, II or 2-3.

IMPROVEMENTS OF VARIOUS WEAKNESSES IN THE METHODOLOGY AND SUGGESTIONS

At first, the geo-referencing of the information was based on criteria that are more compliant to the general directions and guidelines of the European Macroseismic Scale EMS 98. The most precise geographic area associated to a macroseismic observation was a hamlet and the largest a municipality. This information was found contradicting in few cases as in Greece, few years after the publication of the official description of the EMS 98 macroseismic scale, there was an administrative re-organization of the country and as a result, a lot of areas that were considered prefectures are now considered municipalities. These contradicting cases were overcome by adding a precision score, classifying thus the geographic accuracy of each observation to specific values (1, 2, 2-3, 3, 3+, 5-6+). A medium sized Greek city corresponds to value 2. For the final development two subsets of seismic observations were used. The first one, contains only the observations that receive precision score 1 and 2 (Map 1) and the second one is consisted of observations that receive score 1, 2 and 2-3 (Map 2).

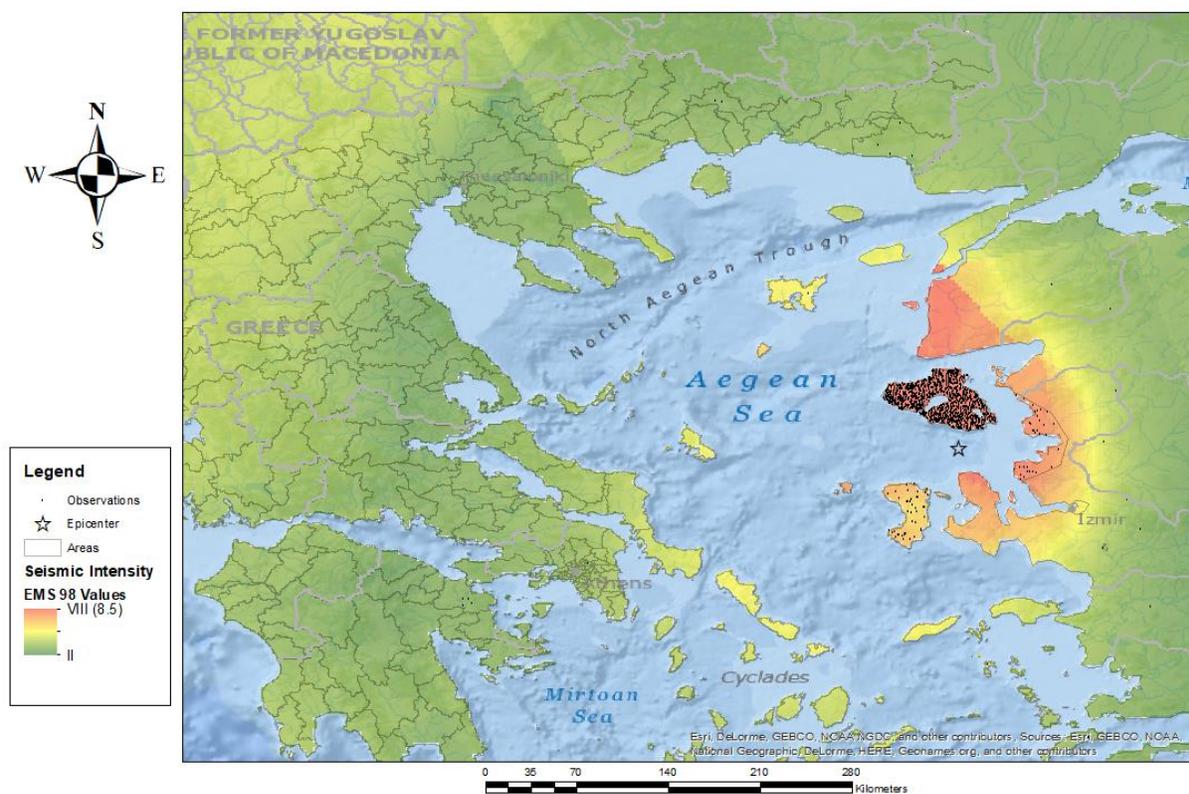
Another issue, was that when the official description of the scale was published in 1998 none of technological tools like social media even existed or there was not anyone that could imagine the contribution of those tools in the field of observation collection. As a result, in some cases, it was not very easy to classify a short text observation to the most suitable value; it is strongly recommended that all these technological developments will be considered in a next version of the official description of the EMS 98 macroseismic scale.

Another improvement had to do with the enrichment of the text patterns. By that term, we mean the classification to certain values of the EMS 98 intensity scale of the tweets that contain certain words within their text. While these text-patterns radically improve the method in terms of speed, there were some contradicting incidents. For instance, in the same text there were more than one words associated to different values. In these cases, only the words classified to the higher value of the intensity scale were considered. All the new along with the previous text patterns are presented in table I.

MAPS AND TABLES

Map 1: Seismic intensity map: Tweets published within 48 hours from the Earthquake event occurrence. Tweets in geographic areas with precision score up to 2-3.

Seismic Intensity Map Based on observations extracted from twitter within 48h (Lesvos 12/6/2018 Mw = 6.3)



Map 2: Seismic intensity map: Tweets published within 48 hours from the Earthquake event occurrence. Tweets in geographic areas that receive precision score up to 2.

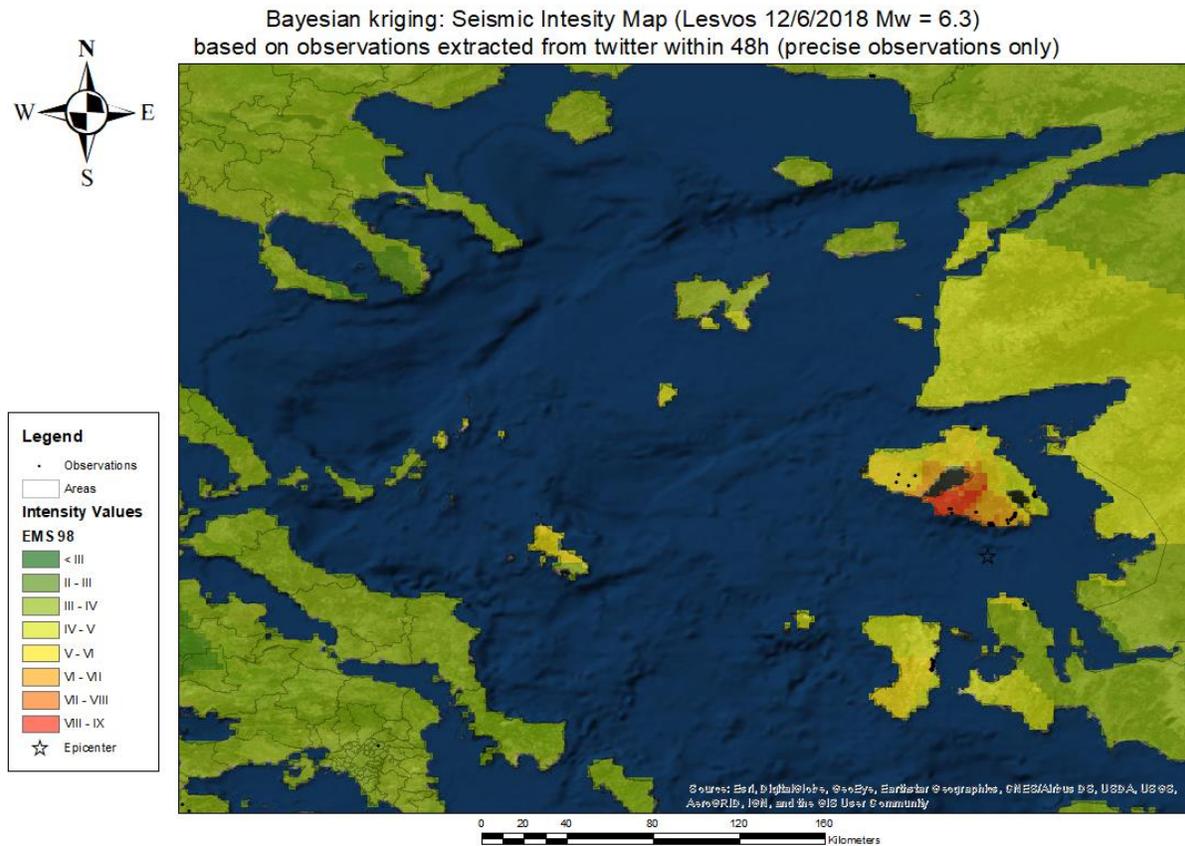


Table I: Text patterns and corresponding intensity values (left columns previously published, right columns new patterns).

Words in tweet-text (translated, previous and new patterns)	Intensity value	Words in tweet-text (translated, previous and new patterns)	Intensity value
Disruption	VI	Major damage	VIII
Powerful, strong, big	V	Heavily damaged	IX
Minor damages	VI	Ruins, guts, debris, rubble	VIII – IX
Damages / major damages	VII	Ghost village	X
General alert, emergence situation,	VII	Most houses in area collapsed	X – XI
Felt	III	Earthquake victims	VIII
Felt enough, felt strongly	IV	Evasion	VIII
Cracks on roads	VII-VIII	Many houses have been collapsed	IX
Evacuation	V	Houses have been collapsed	VIII
Windows crack/break	VI	Bombed landscape	IX
Landslides	VI		
Fear, terror, panic	VII		
Stadium evacuation	VII		
Emergency situation	VII		
Catastrophe, catastrophic	VII		
Broken glasses			
Agony	VI		

CONCLUSIONS

According to the map I the observed values range from about II to IX; the maximum values are located in the whole area of Lesvos island (which is also a municipality); Chios Island and the western coast of Turkey receive the second higher values of about VI. The more we get away from Lesvos Island the lower the intensity appears to be. According to map 2, the observed values have the same range but the maximum values (of about VIII+) are located in the south-central area of Lesvos island. The rest of the island receives a value range between V and VII while in Chios island the maximum values are of about V. In this map as well, the more away we are from the earthquake epicenter, the weaker intensity is observed. The main difference between the two maps is about the geo-referencing part. In particular the first map, that is based on more than 2000 observations contains a lot of information that has a more general geographic precision (for instance: Lesvos island, of municipality of Lesvos, instead of Vrissa town which is located in the south

of Lesvos island). As a result, after the geoprocessing technique of randomization it is possible to have observations graded to VIII, IX or even X in the north part near the Skala Sikamniias in which there are precise observations that receive values of about V. This specific part is something that needs to be improved in a future in a modeled way, probably by weighting each value according to its geographic precision. Thus, a most accurate map that will consider the total amount of macroseismic observations will be created.

Finally, by comparing the Maps with other published work, it seems that Map 2, the one that contains the precise geographic observations only, is closer to the results that have been published through from the hybrid approach of the National Observatory of Athens (NOA) and from the macroseismic intensity map published by the Institute of Engineering Seismology and Earthquake Engineering (ITSAK). Moreover the results seem to be more compliant with the published maps of Antoniou et al. (2017)

considering though that their work covers a really precise mapping of intensity that focuses on the town of Vrissa and receives a score of about X. Map 1 which has a limited geographic accuracy, in reality maps the seismic intensity at a municipality level which in this specific paradigm is a quite big area.

Map 3: Seismic intensity maps of Lesvos earthquake published from the ITSAK (left) and the NOA (right).

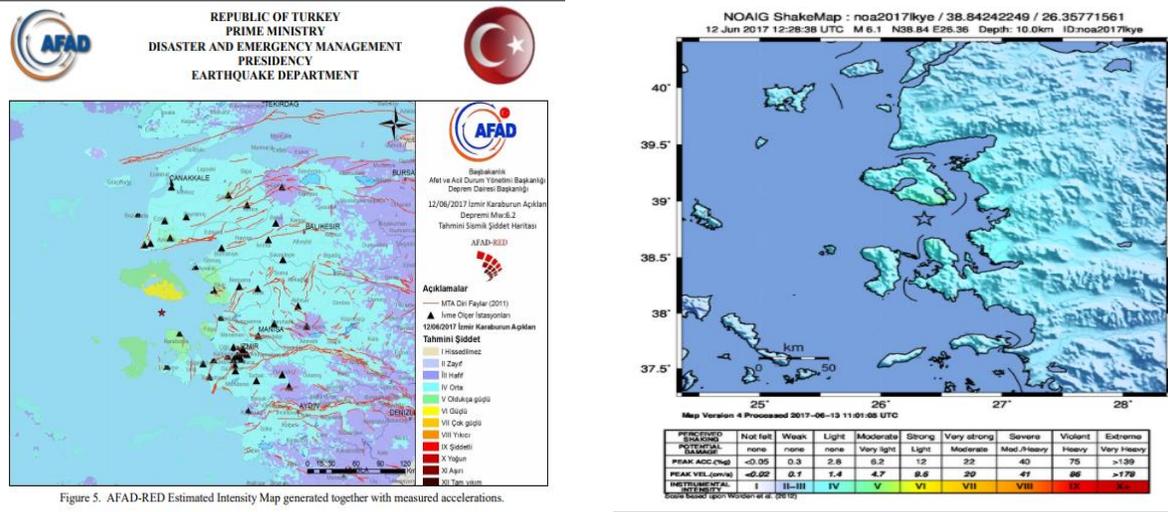


Figure 5. AFAD-RED Estimated Intensity Map generated together with measured accelerations.

Acknowledgements

Various open data have been used during the geo-processing procedure. These include a shape file of the administrative areas of Greece that is available through the geodata.gov.gr portal. Moreover, a shape file of the European Countries and of Turkey available at the Eurostat official page, was used. The software used for the geo-processing of the data was ArcMap 10.5.1 along with the open source software Quantum GIS. The tweet dataset was acquired from the data provider Sifter; expenses were covered by the main author.

REFERENCES

- Antoniou V., Andreadakis E., Mavroulis S., Spyrou N., Skourtsos E., Lekkas E., Carydis P. 2017. Post-earthquake rapid assessment using unmanned aircraft systems (UAS) and GIS online applications: the case of Vrissa settlement after the 2017 June 12, Mw 6.3 Lesvos earthquake (North Aegean Sea, Greece) 8th International INQUA Meeting on Paleoseismology, Active Tectonics and Archeoseismology (PATA) 13 – 16 November, 2017, New Zealand
- Arapostathis S., Parcharidis I., Kalogeras I., Drakatos G. 2015. [Mapping seismic intensity using twitter data: A case study: The February 26th, 2014 ML=5.9 Kefallinia \(Greece\) earthquake.](#) Abstract in Proceedings of the EGU International Congress 17-22 of April, Vienna, Austria, 2015
- Arapostathis S. G., Parcharidis I., Stefanakis E., Drakatos G. and Kalogeras I. 2016. A Method for Developing Seismic Intensity Maps from Twitter Data. *Journal of Civil Engineering and Architecture* 10 (2016) 839-852. Doi: 10.17265/1934-7359/2016.07.013
- Cecic, I., and Musoon, R. 2004. "Macroseismic Surveys in Theory and Practice." *Natural Hazards* 31: 39-61
- Grünthal, G. 1998. European Macroseismic Scale 1998. Centre Européen de Géodynamique et de Séismologie. Accessed January, 14, 2015. http://media.gfz-potsdam.de/gfz/sec26/resources/documents/PDF/EMS-98_Original_englisch.pdf.
- Huiji, G., and Barbier, G. 2011. Harnessing the Crowdsourcing Power of Social Media for Disaster Relief. *IEEE Intelligent Systems* 26 (3): 1541-672.
- Kropivnitskaya Y., Tiampo K. F., Qin J., Bauer M. 2016. Real-Time earthquake intensity information using streaming data analysis of social and physical sensors. *Pure and Applied Geophysics* Vol 174(6):2331-2349
- Lekkas E., Voulgaris N., Karydis P., Tselentis G-A., Skourtsos E., Antoniou V., Andreadakis E., Mavroulis S., Spirou N., Speis F., Papadimitriou P., Kouskouna V., Kassaras G., Kaviris G., Pavlou K., Sakkas V., Chouliaras G. 2017. Lesvos Earthquake: Preliminary Report Mw 6.3. http://paleoseismicity.org/wp-content/uploads/2017/07/2017-06-12-Lesbos-EQ-PaleoseismicityReport_small.pdf
- Schenkóvá, Z., Kalogeras, I., Schenk, V., Pichl, R., Kourouzidis, M., and Stavrakakis, G. 2005. Atlas of Isoleismal Maps of Selected Greek Earthquakes (1956-2003). Athens: Institute of Geodynamics of the National Observatory of Athens and the Institute of Rock Structure and Mechanics of the Czech Academy of Sciences.
- Yamazaki F., Matsuda T., Denda, S., Liu, Wen 2015. Construction of 3D models of buildings damaged by earthquakes using UAV aerial images. The 10th Pacific Conference on Earthquake Engineering, 6-8 of November, Sindney, Australia