



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## LIQUEFACT

Assessment and mitigation of Liquefaction potential across Europe: a holistic approach to protect structures/infrastructure for improved resilience to earthquake-induced Liquefaction disasters.

H2020-DRA-2015

GA no. 700748



### DELIVERABLE D2.6

**Report to describe the adopted procedure for the development of the European liquefaction hazard map**

Author(s):	Carlo G. Lai, Claudia Meisina, Francesca Bozzoni, Daniele Conca, Roberta Bonì
Responsible Partner:	Università degli Studi di Pavia/Eucentre
Version:	1.0
Date:	31/07/2019
Distribution Level (CO, PU)	CO



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## DOCUMENT REVISION HISTORY

Date	Version	Editor	Comments	Status
10/07/2019	1	UNIPV/Eucentre: Carlo G. Lai, Claudia Meisina, Francesca Bozzoni, Daniele Conca, Roberta Boni	First Draft	Draft

## LIST OF PARTNERS

Participant	Name	Country
UNIPV/Eucentre	Università degli Studi di Pavia/Eucentre	Italy

## GLOSSARY

Acronym	Description
GIS	Geographical Information System
DEM	Digital Elevation Model
CTI	Compound Topographic Index
PGA	Peak Ground Acceleration
PGAm	Magnitude-weighted PGA
GHSL	Global Human Settlement Layer
RD	River Distance
CD	Coast Distance
WBD	Waterbody Distance
TPI	Topographic Position Index
TRI	Terrain Roughness Index
EV	Explanatory Variable
SMOTE	Synthetic Minority Over-sampling Technique



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

ADASYN	Adaptive Synthetic
AUC	Area Under the Curve
ROC	Receiver Operating Characteristics
TPR	True Positive Rate
FPR	False Positive Rate
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
Pd	Population Density
AHP	Analytical Hierarchy Process
CI	Consistency Index



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## CONTENTS

1.	INTRODUCTION AND PURPOSE OF THIS DOCUMENT .....	7
2.	GIS DATABASE FOR MACROZONING THE LIQUEFACTION RISK IN EUROPE .....	8
2.1	Liquefaction events historical catalogue .....	8
2.2	Geological, hydrogeological and geomorphological data at the European scale .....	9
2.3	Seismological data collected for Europe .....	9
2.4	Proxy data of exposure available in Europe .....	10
3.	MAPPING THE LIQUEFACTION SUSCEPTIBILITY IN EUROPE .....	11
3.1	Methodology overview.....	11
3.2	Dataset.....	12
3.3	Macrounits for liquefaction susceptibility.....	13
3.4	Analytic Hierarchy Process (AHP) .....	15
4.	GEOSPATIAL METHODOLOGY TO ASSESS LIQUEFACTION RISK AT THE EUROPEAN SCALE .....	22
4.1	Methodology overview.....	22
4.2	Dataset.....	23
4.3	Explanatory variables selection .....	25
4.4	Development of a prediction model .....	28
4.5	Adopted exposure indicators for Europe .....	30
4.6	Macrozoning the liquefaction risk in Europe .....	32
4.7	Filtering of the maps.....	34
5.	MAPS DISPLAYING THE LIQUEFACTION HAZARD AND RISK AT CONTINENTAL SCALE.....	35
5.1	Liquefaction hazard maps for Europe .....	35
5.2	Validation by superimposing historical liquefaction occurrences.....	37
5.3	European maps of liquefaction risk.....	41
6.	CONCLUDING REMARKS .....	43
	REFERENCES .....	44



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## LIST OF FIGURES AND TABLE

### FIGURES

Figure 3-1: Workflow of the liquefaction susceptibility assessment. ....	11
Figure 3-2: Map of the environment deposition (Panagos et al. 2012). ....	13
Figure 3-3: European macrounits .....	13
Figure 3-4: Liquefaction susceptibility map of Europe .....	20
Figure 3-5: Detail of the Liquefaction susceptibility map for the MHL4 .....	20
Figure 3-6: Percentage of the observed liquefaction phenomena versus the liquefaction susceptibility classes of the MLH4.....	21
Figure 4-1: Workflow of the methodology applied .....	22
Figure 4-2: 0 cells regions definitions, for (a) events that triggered liquefaction, (b) events that did not trigger the phenomenon .....	25
Figure 4-3: flowchart of the procedure adopted to plot the liquefaction probability values for each explanatory variable. An example is shown, where the values of two bins (highlighted in orange and green) are calculated. ....	27
Figure 4-4: (a) Population density map and (b) CORINE land cover map.....	31
Figure 4-5: Final exposure Europe model, subdivided into 5 classes.....	32
Figure 5-1: Maps showing the binary liquefaction prediction relative to the three models developed, (a) ADASYN, (b) SMOTE, (c) Undersampling with a PGA relative to the 475 return years period .....	36
Figure 5-2: Maps developed with the ADASYN model. For each return period is presented a binary model prediction map and a probability of liquefaction map. Specifically, (a) and (b) refers to 475 years return period, (c) and (d) refers to 975 years return period, (e) and (f) refers to 2475 years return period. ....	37
Figure 5-3: Number of manifestations of earthquake-induced liquefaction included in the European earthquake catalogue grouped for different ranges of return periods. ....	38
Figure 5-4: Distribution of manifestations of earthquake-induced liquefaction included in the catalogue for which cases were collected. ....	38
Figure 5-5: Validation maps for the 475 years return period, which show (a) the binary model prediction map with the liquefaction events superimposed and (b) a zoom to the Italian/Balkan region .....	39
Figure 5-6: Validation maps for the 475 years return period, which show (a) the liquefaction probability map with the liquefaction events superimposed and (b) a zoom to the Italian/Balkan region .....	40
Figure 5-7: Maps showing European liquefaction risk for (a) 475 years, (b) 975 years, (c) 2475 years return period. ....	42

### TABLES

Table 3-1: Characteristics of the European macrounits .....	14
Table 3-2: Analysis of the historical earthquake-induced soil liquefaction occurrences in Europe .....	17
Table 3-3: Selected variables for the liquefaction susceptibility assessment according the literature overview (Youd and Perkins, 1978; Papathanassiou et al. 2010; Zhu et al. 2017).....	17
Table 3-4: Normalized class weights .....	18
Table 3-5: Weight of the factor for the different macrounits .....	19



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

Table 4-1: Example of the layout of the dataset developed .....	23
Table 4-2: results of the Luco & Cornell methodology, in terms of practicality, efficiency and proficiency. The values for each criteria are marked with a color in a scale from green (good result) to red (poor result). ...	28
Table 4-3: Classification of the variables for each criteria. The variables in each column are ordered from the top (best variables) to the bottom (worst variable). .....	28
Table 4-4: Calibration results of the three logistic regression models along with the Zhu et al. (2015) model. ....	30
Table 4-5: Relative importance for comparison between alternatives (Saaty, 1980).....	33



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

LIQUEFACT  
Deliverable 2.6

*Report to describe the adopted procedure for the development of the European liquefaction hazard map*  
v. 1.0

# 1. INTRODUCTION AND PURPOSE OF THIS DOCUMENT

Work Package 2 (WP2) of LIQUEFACT project deals with the zonation of a territory for liquefaction hazard at both continental and municipal or submunicipal scale. More specifically, the goal of Task 2.5 in WP2 is the definition of a European liquefaction hazard map (*macrozonation*). In a map of liquefaction hazard, the territory is subdivided into an appropriate number of homogeneous zones where the likelihood of earthquake-induced soil liquefaction is displaced according to a specified chromatic scale.

At a first glance, zonation of a large territory for liquefaction risk seems an almost impossible task since liquefaction is a phenomenon of soil instability occurring at a very local scale, that is it may or it may not occur at a specific location and depth from the ground surface depending on whether certain conditions of soil susceptibility and severity of ground shaking are met at that particular depth. Thus, the macrozonation of liquefaction hazard at the continental scale is a truly hard facing challenge. Yet, a qualitative representation of the variability of liquefaction potential within a single country is within reach considering the resolution and accuracy of geological and geotechnical information that is currently available in the most developed nations. The availability of a macrozonation map of liquefaction risk of a country can be useful to policy makers and administrators of that country in identifying territories that are potentially at risk of earthquake-induced ground failures. This in turn could motivate the interest in drafting plans for further investigations and in-depth studies in those territories.

More specifically, the University of Pavia and EUCENTRE lead the effort of constructing geo-referenced European earthquake-induced soil liquefaction risk maps for various return periods. They are built using available datasets at a continental scale on the expected seismic hazard and on the geological, geomorphological, hydrogeological, shallow lithology and digital terrain information.

A crucial step of the work was the selection of the best variables connected to the liquefaction manifestation. Subsequently, a prediction model was developed employing a logistic regression, a data-driven algorithm. A *knowledge-driven methods* like the analytical hierarchy process was applied to develop the final risk map.

A validation of this work was carried out by superimposing on the calculated macrozonation maps of liquefaction hazard, a GIS-based catalogue of liquefaction manifestations occurred in Europe and well-documented in historical earthquakes. This catalogue has been one of the deliverables of LIQUEFACT project (i.e. Deliverable 2.4). The final liquefaction risk maps of Europe were computed by convolving soil susceptibility to liquefaction, expected severity of ground motion and exposure, the latter being alternatively described by a combination of the European population density or the land use of the European territory.

Aim of this document is to present the procedure adopted in for the development of the European liquefaction hazard and risk maps (macrozonation).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## 2. GIS DATABASE FOR MACROZONING THE LIQUEFACTION RISK IN EUROPE

In this section the geological, hydrogeological and seismological data collected within WP2 by University of Pavia and EUCENTRE, along with the liquefaction events historical catalogue developed, will be shortly presented.

All the data collected were harmonized in a GIS environment. The harmonization in a GIS environment represented a crucial step of the methodology: it was important, indeed, that all the raster data had the same resolution (900mx900m, i.e. the dimension of each pixel) and the same spatial reference system (ETRS 1989 LCC). It was also very important that the rasters were perfectly overlapped (i.e. the edges of the cells of each raster are snapped to those of the other rasters), to compute spatial calculation.

A more detailed report regarding the GIS databased developed and the liquefaction events catalogue can be found, respectively, in the Deliverable 2.5 and in the Deliverable 2.4 of the LIQUEFACT project.

### 2.1 Liquefaction events historical catalogue

Within WP2, aim of Task 2.3 was the construction of a GIS-based catalogue of historical liquefaction occurrences in Europe. In this framework, a database containing historical information regarding the liquefaction-related phenomena occurred in Europe, including sand ejects and boils, soil settlements and lateral spreading, ground and structural failures, was developed.

To build the catalogue of liquefaction manifestations, a critical bibliographic review was carried out to identify the most suitable sources to be used, such as existing databases for specific areas (e.g. for Italy), studies, reports and tales concerning earthquakes, chronicles and diaries, archival documentation and seismic bulletins.

In this research, one of the most important starting points was represented by the earthquake catalogue set up for the European territory within recent research projects (i.e. SHARE "*Seismic Hazard Harmonization in Europe*"). Descriptions of liquefaction manifestations triggered by earthquakes, including, if possible, photos and figures, were gathered from the collected references and used to construct a European database under a GIS environment. Thus, the GIS-based catalogue includes two pieces of information: main seismological features of the seismic events (date, geographic coordinates, magnitude, focal mechanism if known, etc.) and liquefaction site parameters (epicentral distance, type of failure, etc.).

All the liquefaction manifestation events are represented in the GIS environment as point features with associated all the information gathered.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## 2.2 Geological, hydrogeological and geomorphological data at the European scale

A quaternary geological map of Europe (<https://produktcenter.bgr.de>) has been obtained, allowing to discern soil deposits susceptible and not susceptible to liquefaction: indeed, soil deposits susceptible to liquefaction are not randomly distributed but occur within a range of specific sedimentary environments. Liquefaction resistance increases with age, the mode of deposition also has influence on liquefaction susceptibility. Thus, an evaluation of geological units and depositional process can be both used as a screening for identification of liquefaction prone areas. Surficial lithological maps have been also obtained.

Hydrogeological maps (<https://produktcenter.bgr.de>) have been collected. The soil saturation represents a significant influence on the liquefaction susceptibility. In fact, only saturated sediments or sediments capable of becoming saturated with ground water table are susceptible to liquefaction.

The Digital Elevation Model (DEM), obtained from Shuttle Radar Topography Mission (SRTM) dataset (Jarvis et al., 2007), and his derived products have been helpful to gather morphological and hydrological information. Indeed, SRTM DEM was geoprocessed to obtain the following data:

- Local slope;
- Compound Topographic Index (CTI) as defined by Wilson (2000);
- Stream network;
- Euclidean distance from streams network.

The average shear-wave velocity down to 30 m ( $V_{s30}$ ) has been useful for the definition of soil stiffness, this because soft sandy soils are more susceptible to liquefaction. The global topographic-slope based  $V_{s30}$  map was downloaded from <https://earthquake.usgs.gov/data/vs30/>.

## 2.3 Seismological data collected for Europe

From a seismological viewpoint, the following data were gathered from the deliverables of the European project SHARE (<http://portal.share-eu.org>):

- Probabilistic seismic hazard maps for Euro-Mediterranean region such as the map for peak ground acceleration (PGA).
- European earthquake catalogue, which includes harmonized moment magnitude ( $M_w$ ) estimates and provides uncertainty estimates. The most recent version of the Italian earthquake catalogue (<http://emidius.mi.ingv.it/CPTI15-DBMI15/>) was integrated within the GIS platform.
- Seismogenic zones for Europe.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

- Seismogenic faults, i.e. the European database EDSF (<http://diss.rm.ingv.it/share-edsf/>), a compilation of fault sources deemed to be capable of generating earthquakes of magnitude equal to or larger than 5.5 in the Euro-Mediterranean area.

## 2.4 Proxy data of exposure available in Europe

Concerning exposure, population density is a well-established proxy in case of residential and public buildings. This is combined with additional open-access databases such as CORINE which provides the geo-referenced distribution of non-residential areas in Europe (Sousa et al., 2017). Indeed, the European initiative, named Global Human Settlement Layer (GHSL; <https://ghsl.jrc.ec.europa.eu/index.php>) provides a free tool for assessing the presence of human settlements on the planet. From GHSL, the spatial raster dataset, which depicts the distribution and density of population expressed as the number of inhabitants per unit cell, was used as input for macrozonation of liquefaction risk in Europe.

The other proxy for exposure collected was the CORINE land cover map 2012 (<http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view>). The CORINE map subdivides the European territory into 44 different classes of land cover, thus representing an useful tool to define the exposure of the territory.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## 3. MAPPING THE LIQUEFACTION SUSCEPTIBILITY IN EUROPE

### 3.1 Methodology overview

The liquefaction susceptibility was assessed through a knowledge driven approach, in particular, we have used the Analytical Hierarchy Process (AHP, introduced by T. Saaty, 1980), where the judgment of an expert is applied in order to rank the explanatory variables (the highest the rank, the highest the contribution to the liquefaction hazard). The choice of this method is motivated by the fact that the expert decides which factor is important for the liquefaction occurrence and it does not require an accurate inventory (we remember that the liquefaction inventory still remains incomplete in some areas).

Small scale liquefaction zoning over complex terrains like Europe prove difficult since different liquefaction conditioning factors can exert specific control on liquefaction susceptibility (e.g. coastal areas, fluvial areas, etc.). In order to apply the AHP we decided to divide the Europe in macroareas homogeneous from the geological, physiographical and geomorphological point of view (European Macrounits) (Figure 3-1).

The validation of the liquefaction susceptibility map of Europe was obtained through the comparison with the GIS-based catalogue of liquefaction manifestations occurred in Europe (see Section 2.1) and the comparison with the already existing national and regional liquefaction hazard map (Portugal and Greece, Jorge and Coelho, 1994; Papathanassiou et al. 2010).

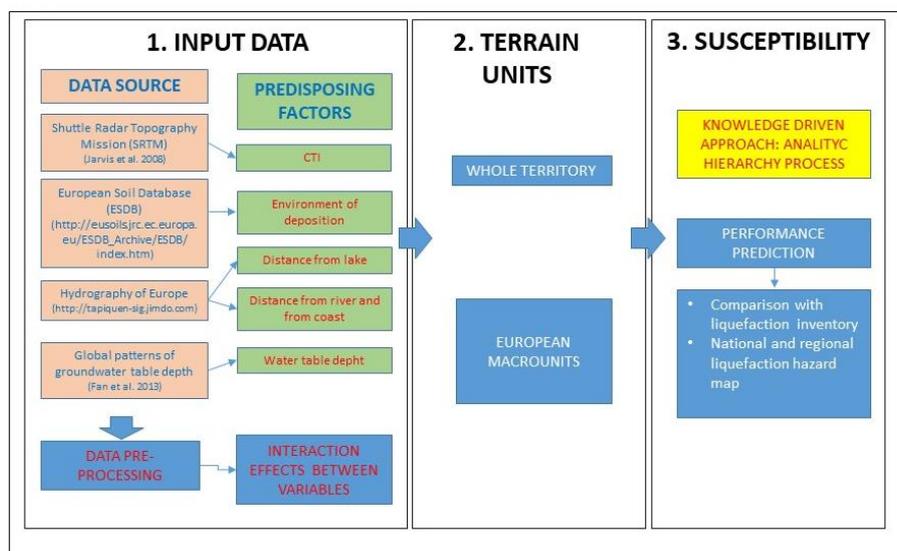


Figure 3-1: Workflow of the liquefaction susceptibility assessment.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

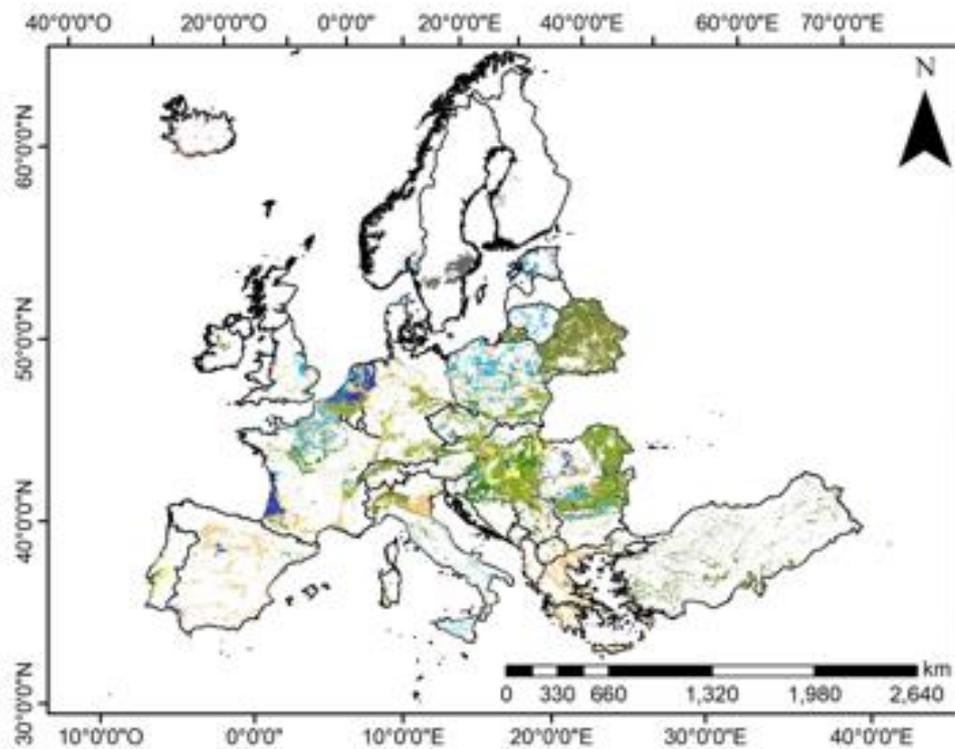
## 3.2 Dataset

A dataset was developed to proceed with the development of the liquefaction susceptibility map of Europe. It was built starting from the data collected and described in Section 2. Furthermore, two additional datasets were used for the analysis such as the (1) European Soil Database (Panagos et al. 2012) including the information about the type of environment deposition of the soils and the (2) water table depth derived from the Global patterns of groundwater table depth (Fan et al. 2013).

Overall, the dataset used for the liquefaction susceptibility map contains the following characteristics:

- Environment of deposition
- Water table depth
- Distance from water bodies (lake and rivers).

In particular, the European Soil Database (Panagos et al. 2012) was exploited to extract the susceptible soils (Figure 3-2) in the different environments of deposition such as coastal and continental areas and including artificial deposits.





This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748



Figure 3-2: Map of the environment deposition (Panagos et al. 2012).

### 3.3 Macrounits for liquefaction susceptibility

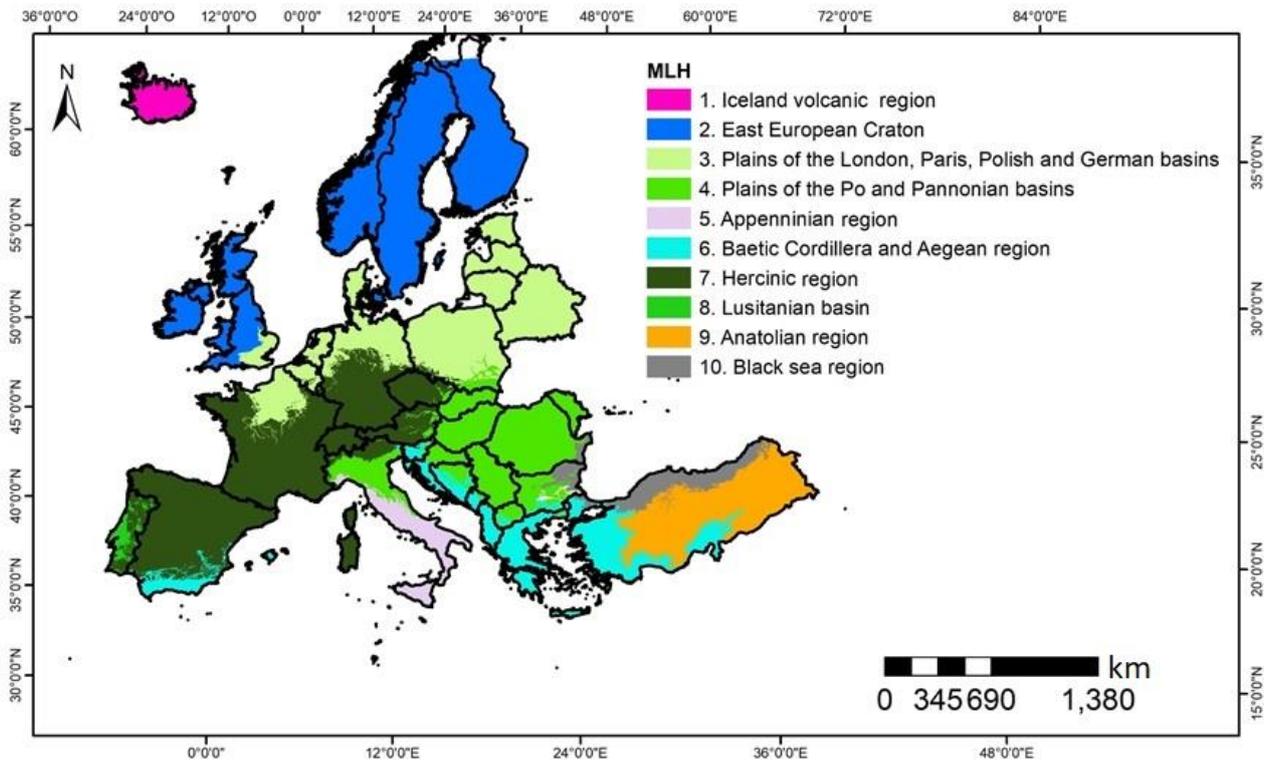


Figure 3-3: European macrounits



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

**Table 3-1: Characteristics of the European macrounits**

MACROUNITS	STRUCTURES	LANDFORM	SEDIMENT ENVIRONMENT
MLH1 Iceland volcanic region	Caledonian orogeny	Mountains and plateau	<ul style="list-style-type: none"> <li>volcanic sand deposits</li> </ul>
MLH2 East European Craton	Caledonian orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>Morenic deposits</li> <li>Lake deposits</li> </ul>
MLH3 Plains of the London, Paris, Polish and German basins	Hercynian orogeny	Plain	<ul style="list-style-type: none"> <li>Fluvial deposits, coastal and eolian deposits</li> </ul>
MLH4 Plains of the Po and Pannonian basins	Alpine orogeny	Plain	<ul style="list-style-type: none"> <li>Fluvial deposits</li> </ul>
MLH5 Appenninian region	Alpine orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>Marine and estuarine clays and silts</li> <li>Fluvial clays, silts and loams</li> </ul>
MLH6 Baetic Cordillera and Aegean region	Alpine orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>Fluvial clays, silts and loams</li> </ul>
MLH7 Hercinic region	Hercynian and Alpine orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>Fluvial clays, silts and loams</li> <li>eolian sands</li> <li>unconsolidated deposits (alluvium, weathering residuum and slope deposits)</li> </ul>
MLH8 Lusitanian basin	Hercynian orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>marine and estuarine sands</li> <li>fluvial sands and gravels</li> </ul>
MLH9 Anatolian region	Alpine orogeny	Mountains, plateau and plains	<ul style="list-style-type: none"> <li>Quaternary, Alluvium fan, slope debris, cone of dejection etc.</li> </ul>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

MLH10 Black sea region	Alpine orogeny	Mountains and plateau	• Beach and dune
------------------------	----------------	-----------------------	------------------

The macrounits for the assessment of the liquefaction susceptibility Europe represents homogeneous areas from the geological point of view. The European zonation for the macrounits was obtained through the combination of different map such as (1) the structural map of Europe (Plant et al. 2003), (2) the Geological map of Europe (<http://www.europe-geology.eu/onshore-geology/geological-map/>) derived from the Onegology Project, (3) the landform map (Meybeck et al. 2019) and (4) the sediment environment map (Plant et al. 2003). Finally, we obtain 10 macrounits (Figure 3-3, Table 3-1).

### 3.4 Analytic Hierarchy Process (AHP)

The AHP is designed to solve complex problems involving multiple criteria (Saaty, 1980). The process requires the decision maker to provide judgments about the relative importance of each criterion and then specify a preference for each decision alternative on each criterion. The output of the AHP is a prioritized ranking indicating the overall preference for each of the decision alternatives.

More in detail, AHP is an expert-based, stepwise classification technique designed to hierarchically organise criteria (here the factors) to solve complex decisions through pairwise comparisons of their relative importance on a scale from 1 to 9.

The steps of AHP are the following:

1. Selection of the variables.
2. Relative importance of each variable.
3. Preference scale and ratings for each variable.
4. Synthetizing procedure.
5. Consistency checking.

The first step in the AHP approach is the selection of the variables (here factors). There are three broad factors that contribute to the likelihood of liquefaction (Youd and Perkins 1978, 1987, Ishihara 1996): density, saturation (or water table depth), and dynamic load on the soil from an earthquake (both intensity and duration). Typically, the first two factors are measured on a site-specific basis using geotechnical logs and penetration data. Regional liquefaction hazard mapping projects have predominantly relied on criteria that relate Quaternary surficial deposits to liquefaction susceptibility, taking into account factors such as depositional environment, dominant grain size, and relative age (Youd and Perkins, 1978). This methodology commonly leads to the identification of large regions of susceptible material. Youd and Perkins (1987) discussed how the resulting maps show geologic units that likely contain liquefiable sediments but do not identify the precise location of the liquefiable sediments within the geologic unit. Therefore, it is possible that within a susceptible unit only a small discrete area or areas will actually liquefy during a given



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

earthquake. In order to select the most important variables we have analysed the literature about the historical earthquake-induced soil liquefaction occurrences in Europe (Table 3-2 and Table 3-3). The following variables seem to have had a great influence on the liquefaction occurrence:

- environment of deposition (the type of sediments) (ED, Figure 3-1);
- distance from water bodies (lake, coast and rivers) (LD, RD). Characterizing the degree of saturation or depth of the water table is one of the most important factors in predicting liquefaction of soils. Knudsen and Bott (2011) identified several candidate proxies for soil saturation, though most can simultaneously be considered proxies for density. The most promising proxies for saturation were the distance to the closest water body. Within coastal regions, the distance from the coast generally correlates with the age of the sediment because older and denser sediments are generally located farther from the coast due to the transport and depositional process of marine sediments. Young, loose sediment is also found along large rivers. We also include the compound topographic index (CTI; Beven and Kirkby 1979) as a proxy for saturation. CTI is defined as the natural logarithm of the ratio of contributing area to the tangent of slope (Moore et al. 1991). In order to compute the contributing area, the flow direction at each pixel is estimated from the digital elevation model (DEM). The contributing area at a specific location is the number of upstream pixels. Thus, CTI is increases on flat areas with large contributing areas.
- Water table depth (GW) (Fan et al. 2013).

In order to find the relative importance of each variable we reclassified the variables in a reduced number of classes. The relative weights for the individual classes of each factors were directly assigned based on our understanding of liquefaction susceptibility over Europe (Table 3-4).

Next, the relative importance of the used criteria in liquefaction susceptibility was decided. For this instance, a pairwise comparisons of the factors were performed within the AHP for each macrounit (Table 3-5).

The integration of the weight parameter classes into a liquefaction susceptibility index was determined using their weighted linear sum.

Finally, the weighted linear summation of the criteria classes was classified into five levels though equal interval slicing.

We acknowledge that the relative importance of a single criteria may not work everywhere in the study area. Therefore, susceptibility evaluation was performed individually for each macrounits to obtain the liquefaction susceptibility map across Europe (Figure 3-4).

A validation of the outcomes of the liquefaction susceptibility maps of Europe obtained with AHP is carried out by superimposing the computed maps to the GIS-based catalogue of liquefaction occurrences in Europe, already presented in Section 2.1 (Figure 3-5).

A good correlation was obtained between the observed liquefaction phenomena mapped in the inventory and the highest susceptibility classes (Figure 3-6).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

Table 3-2: Analysis of the historical earthquake-induced soil liquefaction occurrences in Europe

COUNTRY	LOCATION	EARTHQUAKE (Moment Magnitude, Mw)	DEPOSIT	AGE	WATER TABLE DEPTH (m)	DISTANCE OF WATER BODY	REFERENCE
Agean region	Gulf of Corinth, the islands of the Ionian Sea, coastal zone of the Sea of Marmara (Turkey)	-	alluvial and fluvial deposits, dunes, coastal, deltaic, marsh and artificial fill deposits	Holocene sediments	-	0-100	Papathanassiou and Pavlides (2011)
Albania	Northwestern Albania (Shkodra district)	March 18, 1962 Fier earthquake (Mw 6.2)	Gravels, sands, silty sands, silty clays, and clays	Quaternary deposits	0.50-1	-	Daja et al. (2013)
		April 15, 1979 Adriatic sea (Mw 6.9)	Sandy beaches and lagoons	Quaternary sediments	-	on both sides of Buna river	Kocio (2004)
Bulgaria	Kresna	Kresna earthquake 1904 (Mw 7.1)	Small valley at the confluence of rivers and valleys	-	-	2 km from Bobochevo and Struma river	Ambraseys (2001)
Croatia	Kopa Valley Zagreb area	October 8, 1909 Kopa Valley earthquake (Mw 5.8)	-	-	-	-	Herak D. and Herak M. (2010)
		Great Zagreb earthquake, 1880 (Mw 6.2)	Alluvium Sava River deposits	-	-	-	Veinovic et al. (2010)
Hungary	Komárom	Komárom earthquakes of 1763 (Mw 5.7), 1783 (Mw 5.35)	Sandy layers at a depth between 4 and 17m below the surface	-	-	left bank of the Danube	Györi et al. (2015)
		Kecskemét earthquake, 1911 (Mw 5.6)	Coarse bluish quartz-sand formation	-	5 m	-	Györi et al. (2015)
Iceland	Olafus region	2008, Olafus earthquake (Mw 6.3)	Alluvial volcanic sand deposits	Quaternary deposits	-	along the Olafus River	Green et al. 2012
		Avezzano earthquake, 1915 (Mw 6.99)	paleoliquefaction of intermountain Quaternary lacustrine deposits	Quaternary deposits	1.7 m	-	Galli (2000)
Italy	Po Plain	Emilia Romagna earthquakes (May 20 Mw 5.9 and May 29 Mw 5.8)	sandy and silty-sand deposits, which are spatially related with river channels, levees and crevasse splays	poorly consolidated fluvial deposits dated from the Upper Pleistocene to the Holocene	-	near abandoned riverbed	Lanfredi et al. 2018
Greece	Cephalonia	Cephalonia earthquakes, 2014 (Mw 6 and 6.1)	Coastal deposits and artificial fill	-	-	-	Papathanassiou et al. 2016
	Lefkada	South Lefkada earthquake, Ionian Sea, 2015 (Mw 6.4)	Coastal and fluvial deposits	Holocene deposits	-	-	Ganas et al. (2015)
Portugal	-	-	'streams and rivers', 'low and wetlands' and 'floodplains' and 'moorlands' deposits	-	-	along the coast	Vaz and Zezere (2016)
Romania	Bucharest	Vrancea earthquake, 1977 (Mw 7.4)	medium-dense to dense sand gravel, fine to medium sand and aeolian loess deposits on the flat	Holocene deposits	-	old riverbed of the Dambovitza river	Hannich et al. (2007)
Spain	Betic Cordillera	-	paleoliquefaction of alluvial fan facies, fluvial and lagoon facies and littoral facies	Pleistocene-Holocene deposits	-	-	Alfaro and López-Casado (2001)
Turkey	Izmit Bay	Kocaeli (Izmit) earthquake, 1999 (Mw 7.4)	Young sediments of marine and continental facies	Holocene deposits	-	along the coast of Izmit Bay	Cetin et al. 2004
	Adana	Adana-Ceyhan earthquake, 1998 (Mw 6.2)	young alluvial soils of mainly silty clay, overlying silty fine sands with occasional gravel.	-	-	1-2 Km from the riverbanks	Ulusay et al. 2000
	East Anatolian region of Turkey	Van earthquake, 2011 (Mw 7.2)	sands (fine-coarse) mixed with silt and rare clay.	Quaternary deposits	0-5	along the sides of the streams and Lake van shores.	Karakas A., Corik O. (2013)

Table 3-3: Selected variables for the liquefaction susceptibility assessment according the literature overview (Youd and Perkins, 1978; Papathanassiou et al. 2010; Zhu et al. 2017)

Variable	Data source	Reference	Spatial resolution	Role of the variable for liquefaction phenomena
Compound Topographic Index (CTI)	Shuttle Radar Topography Mission (SRTM)	Jarvis et al. 2008	90 m	Proxy for the soil saturation
Euclidean distance from river (RD)	Hydrography of Europe	<a href="http://tapiquen-sig.jimdo.com">http://tapiquen-sig.jimdo.com</a>	-	Proxy for the soil saturation
Water table depth (GW)	Global patterns of groundwater table depth	Fan et al. 2013 <a href="https://www.eea.europa.eu/data-and-maps/data/wise-large-rivers-and-large-lakes">https://www.eea.europa.eu/data-and-maps/data/wise-large-rivers-and-large-lakes</a>	-	Proxy for the soil saturation
Distance lake (LD)	Large Lake of Europe	(Heineke et al. 1998; Panagos et al. 2012; <a href="http://eusoirs.jrc.ec.europa.eu/ESDB_Archive/ESDB/index.htm">http://eusoirs.jrc.ec.europa.eu/ESDB_Archive/ESDB/index.htm</a> )	1:1,000,000	Proxy for the soil saturation
Environment of deposition (ED)	European Soil Database (ESDB)			Proxy for the density



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

Table 3-4: Normalized class weights

Factor		Class	Class weight
ED	Coastal zone	pre-quadernary sand, tertiary sand, pre-quadernary clay and silt, tertiary clay, tertiary silt	0.07
		holocene coastal sand with shells, holocene clay, holocene silt	0.20
		marine and estuarine sands, quadernary sand, marine and estuarine clays and silts, quadernary clay and silt	0.27
		delta sand, beach and dune*	0.33
	Continental deposits	river terrace sand or gravel, river terrace sand, river terrace gravel, terrace clay and silt, river loam, terrace loam, overbank deposit, eolian deposits, eolian sands, cover sand	0.07
		unconsolidated deposits (alluvium, weathering residuum and slope deposits), talus scree, unconsolidated glacial deposits/glacial drift	0.13
		glaciofluvial deposits, alluvium fan, slope debris, cone of dejection etc.*, gravel sand**	0.20
		floodplain sand or gravel, floodplain sand, floodplain gravel, floodplain clay and silt, floodplain clay and silt, floodplain loam, lake deposits, lake sand and delta sand, lake marl, bog lime, lake silt, colluvial deposit, loess, loamy loess, sandy loess, dune sand	0.27
		fluvial sands and gravels, fluvial clays, silts and loams, river clay and silt, undifferentiated Quadernary*	0.33
	Artificial	anthropogenic deposits	0.27
		redeposited natural materials, sand and gravel fill	0.33
	CTI	12.01 - 15.69	0.07
		15.69 - 17.53	0.13
		17.53 - 19.18	0.20
19.18 - 20.78		0.27	
20.78 - 24.36		0.33	
GW	>20 m	0.07	
	15-20 m	0.13	
	15-20 m	0.20	
	5-10 m	0.27	
	0-5 m	0.33	
RD	>10 km	0.07	
	5-10 km	0.13	
	3-5 km	0.20	
	1.5-3 km	0.27	
	0-1.5 km	0.33	
LD	>10 km	0.07	
	5-10 km	0.13	
	3-5 km	0.20	
	1.5-3 km	0.27	



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

	0-1.5 km	0.33
CD	>10 km	0.07
	5-10 km	0.13
	3-5 km	0.20
	1.5-3 km	0.27
	0-1.5 km	0.33

Table 3-5: Weight of the factor for the different macrounits

	ED	GW	RD	LD	CTI
<b>MLH1</b>	33	No data	-	-	67
<b>MLH2</b>	49	11	28	-	12
<b>MLH3</b>	46	28	10	-	16
<b>MLH4</b>	54	17	7	7	15
<b>MLH5</b>	40	16	16	-	28
<b>MLH6</b>	23	49	14	-	14
<b>MLH7</b>	33	33	14	-	20
<b>MLH8</b>	48	25	9	-	18
<b>MLH9</b>	46	9	-	14	31
<b>MLH10</b>	54	16	-	-	30



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

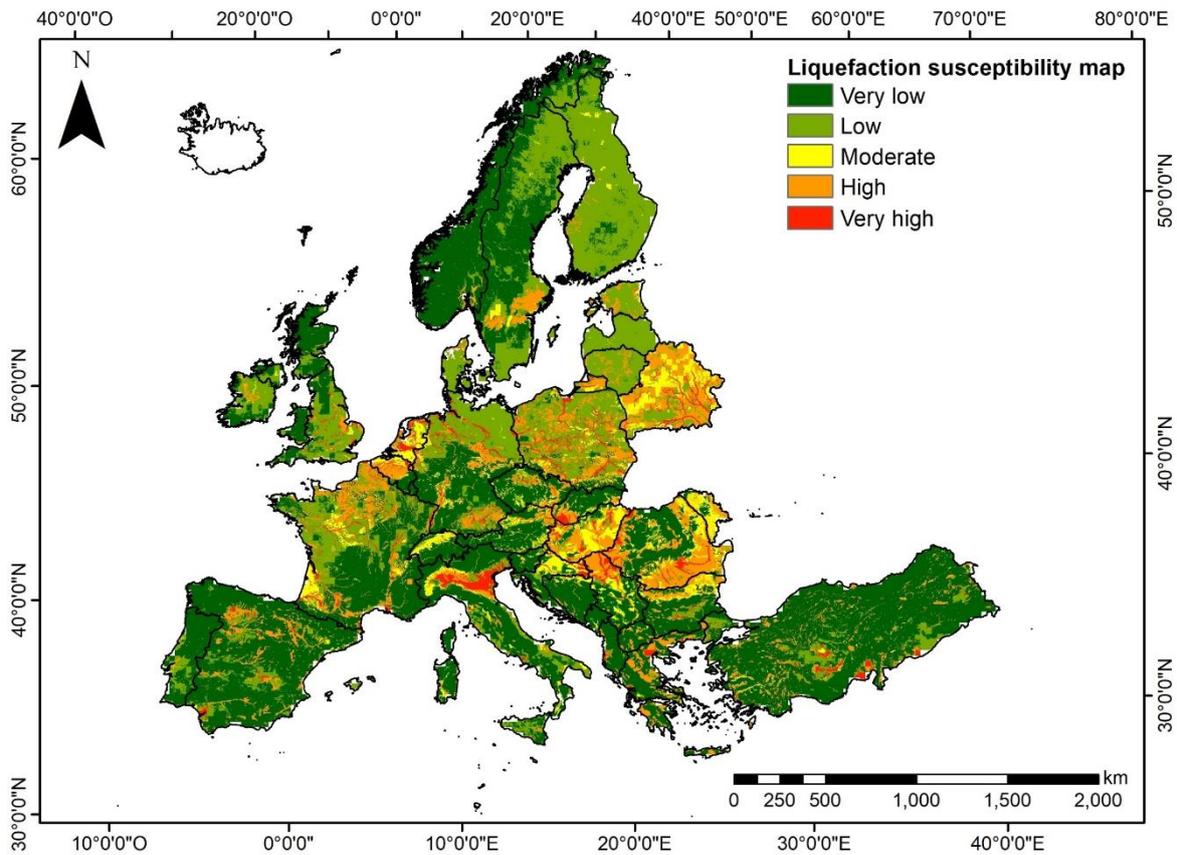


Figure 3-4: Liquefaction susceptibility map of Europe

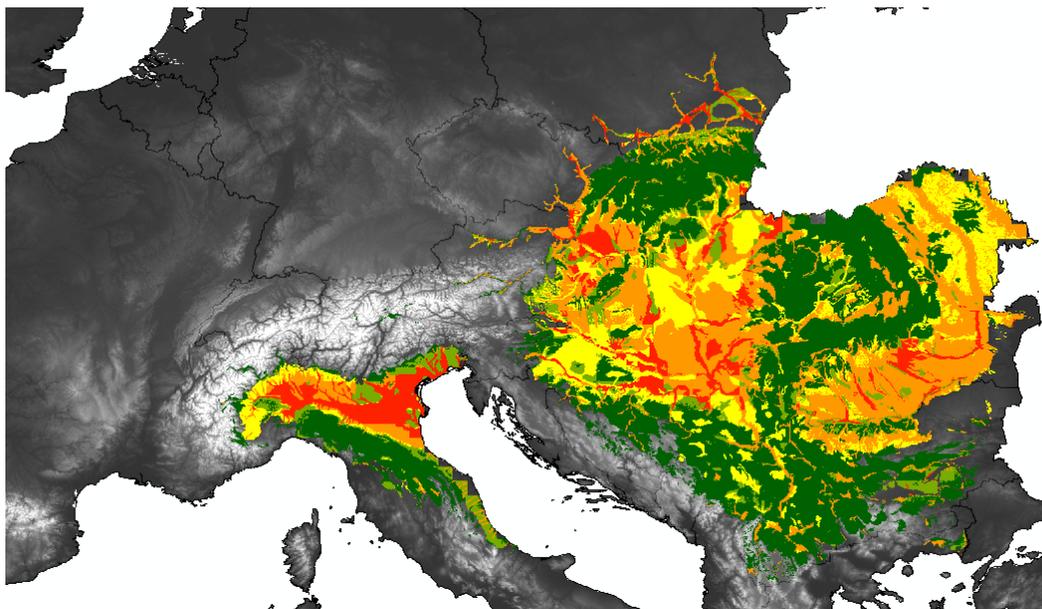


Figure 3-5: Detail of the Liquefaction susceptibility map for the MHL4



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

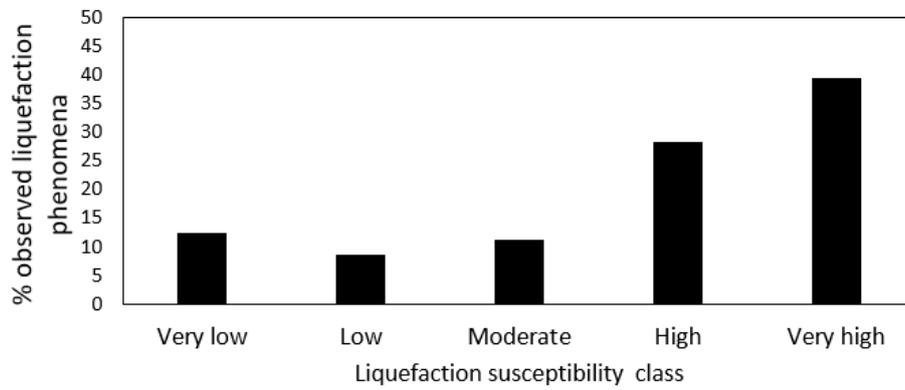


Figure 3-6: Percentage of the observed liquefaction phenomena versus the liquefaction susceptibility classes of the MLH4



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## 4. GEOSPATIAL METHODOLOGY TO ASSESS LIQUEFACTION RISK AT THE EUROPEAN SCALE

### 4.1 Methodology overview

The assessment of liquefaction hazard and risk across Europe was composed of different stages. A brief overview of the methodology's workflow applied is shown in Figure 4-1.

The first step was represented by the collection of the data and their subsequent harmonization in a GIS environment, explained in Section 2.

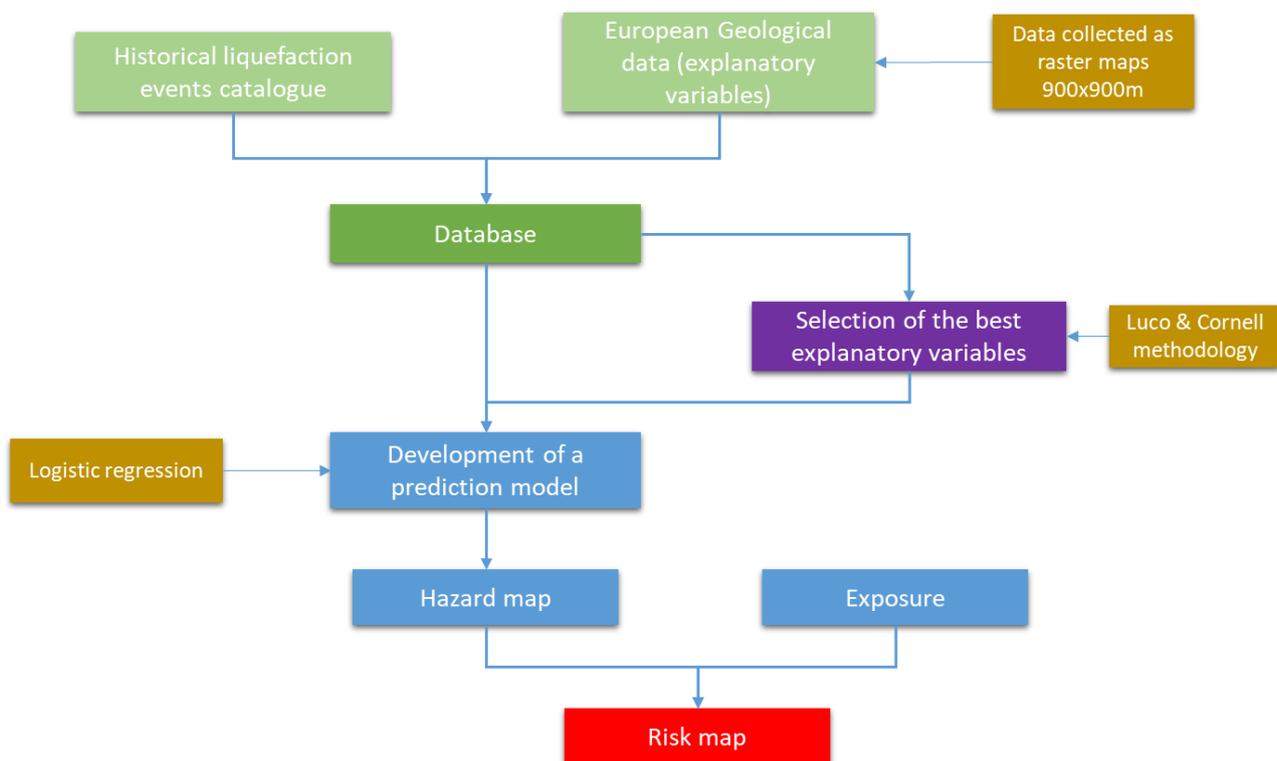


Figure 4-1: Workflow of the methodology applied

The final input data (or explanatory variables) employed in the methodology assessment were:

- PGA (peak ground acceleration, referred to three different return periods: 475, 975 and 2475 years, extracted from SHARE)
- PGAm ( $PGA \cdot MWF$ , where  $MWF = M^{2.56} / 10^{2.24}$  is the Magnitude-Weighting Factor)
- CTI (Compound Topographic Index, derived from DEM)
- River distance (in km, derived from DEM)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

- Coast distance (in km)
- Waterbody distance (in km, i.e. distance from the nearest river/coast/lake)
- TPI (Topographic Position Index, derived from DEM. It compares the elevation of each cell in the DEM to the mean elevation of a specified neighborhood around that cell)
- TRI (Terrain Roughness Index, derived from DEM. It provides a quantitative measure of topographic heterogeneity)

## 4.2 Dataset

A dataset was firstly developed. It represented the starting point of the subsequent analysis and, for this reason, it was carefully built starting from the data collected and the historical catalogue of liquefaction events.

The final dataset was structured in order to contain, for each of the corresponding cell of each raster:

- One value for each explanatory variable
- A binary label (1/0, in the following also referred as 1 cells and 0 cells, respectively) indicating whether if liquefaction was detected in that cell or not (information gathered from the catalogue)

Table 4-1 shows an example of the dataset layout.

**Table 4-1: Example of the layout of the dataset developed**

ID	Vs30 (m/s)	CTI	PGA (g)	...	Y/N
0001	150	25	0,2	...	1
0002	700	15	0,18	...	0
...	...	...	...	...	...

The second point needs some further explanation. The reason why a label is applied to each record of the dataset is that liquefaction manifestation is treated as a binary dependent variable, while the independent variables are the data presented before. If inside a cell is present a liquefaction feature (represented in the catalogue as georeferenced points), that cell is labeled as 1 (positive), otherwise is labeled as 0 (negative).

The peak acceleration values to be inserted in the dataset deserve a separate discussion. It's important to point out that those values are referred to shakemaps relative to the earthquake event that triggered (or did not trigger) liquefaction. The analysis that will be carried out need, indeed, the actual values that triggered liquefaction, or, in the opposite case, that were no strong enough to trigger (on equal terms) the phenomenon. In the light of this, an events selection had to be made. The criteria adopted to select the



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

events were the availability of shakemaps, the number of liquefaction features, the magnitude of the event and the environment features.

A total of 4 events were selected, 3 that caused liquefaction to a various level and 1 that did not cause the phenomenon, even though the environment features would have allowed it.

The selected events were:

- 2012 Emilia earthquake (<https://earthquake.usgs.gov/earthquakes/eventpage/usp000jkn8/executive>);
- 2014 Cephalonia earthquake (<https://earthquake.usgs.gov/earthquakes/eventpage/usb000m8ch/shakemap/intensity>);
- 2009 L'Aquila earthquake (<https://earthquake.usgs.gov/earthquakes/eventpage/usp000gvtu/shakemap/metadata>);
- 2008 Parma earthquake (<http://shakemap.rm.ingv.it/shake/1870169/products.html>)

The first three are events that caused liquefaction (also referred in the following as “liquefaction events”), while for the last one none liquefaction events were detected. In particular, for the 2012 Emilia earthquake the liquefaction occurrences are well detailed, while the 2014 Cephalonia earthquake is characterized by just a few events. The 2009 L'Aquila earthquake presents also very few liquefaction features. The shakemaps for the 4 events selected were harmonized.

The subsequent issue that had to be faced was how to define the range of the regions from which extract data. Indeed, while all the 1 cells were selected, the selection of the 0 cells needed some further discussion. In fact, potentially there could be a huge number of 0 cells, depending on the spatial extension of the shakemaps data, leading to an imbalanced dataset (i.e. a dataset in which classes 0, major class and 1, minor class - are not represented equally). Thus, 0 cells were extracted only from purposely devised regions, whose extensions depended on the type of event, namely event that triggered or did not trigger liquefaction. For the former case, as suggested in Zhu et al. (2017), two different buffers were calculated around liquefaction features: one of 1km and one of 15km. The 1km represents the non-sampling regions: all the 0 cells that falls into this buffer are not available for selection. The non-sampling regions represent thus holes in the sampling region, which is defined as the area characterized by a maximum distance of 15km from the nearest liquefaction feature. In the case of events that did not trigger liquefaction, the sampling region is simply defined as the area of maximum distance of 40km around the epicenter of the event. Figure 4-2 shows the regions definitions for the two different types of event.

At this point, the dataset was composed by the 0 cells extracted from the sampling region and the 1 cells. Each cell represents a vector, containing a value for each explanatory variable along with its label.

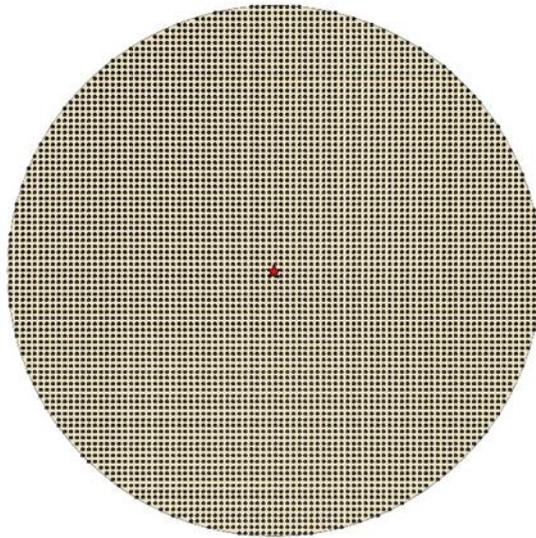
The final dataset was composed of about 160 cells with label 1 and about 13000 cells with label 0, thus characterized by a ratio of about 1:100 between minor and major class. Further methods were applied in the following analysis in order to reduce this ratio.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748



(a)



(b)

Figure 4-2: 0 cells regions definitions, for (a) events that triggered liquefaction, (b) events that did not trigger the phenomenon

### 4.3 Explanatory variables selection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

Starting from the dataset developed, an evaluation of the variables best correlated with liquefaction occurrence were carried out. The Luco & Cornell (Cornell and Luco, 2001; Luco and Cornell, 2007; Padgett et al. 2008) methodology was applied. The methodology evaluates the different variables according to three different criteria: *practicality*, *efficiency* and *proficiency*.

Efficiency and practicality are estimated in this work on the results of linear regressions carried out for each explanatory variable (EV), in the *EV vs Probability of liquefaction* plot. How the probability of liquefaction values was calculated for each explanatory variable will be discussed in the following.

Efficiency expresses the amount of variation in the probability of liquefaction, and is represented in this work by the standard deviation  $\beta$  of the regression analysis. A lower  $\beta$  yields a more efficient EV. Practicality refers to a possible direct correlation between an EV and the probability and is measured by the linear regression gradient parameter  $b$ . A more practical EV is characterized by a higher gradient. Proficiency measures the composite effect of practicality and efficiency. Also called modified dispersion, it is calculated as the ratio between the dispersion (efficiency parameter) and gradient (practicality parameter). A lower value of this ratio yields to a more proficient EV.

Thus, with the objective of applying the Luco & Cornell methodology, starting from the database developed linear regression were carried out for each explanatory variable (EV), in the EV vs Probability of liquefaction plot. The strategy adopted involved different stages, represented schematically in Figure 4-3.

At this stage, as already stated, the dataset presented a ratio of about 1:100 between minor and major class (namely 1 and 0 cells or records). In order to overcome this issue, in this stage was decided to carry on the analysis on a 1:1 sub-dataset. The technique adopted to reduce the number of 0 records can be referred as *undersampling*: 0 records are randomly sampled from the dataset, in a number equal to that of the 1 records. To improve model stability and to avoid the potentially discard of useful or important samples, the sampling procedure is repeated 1000 times.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

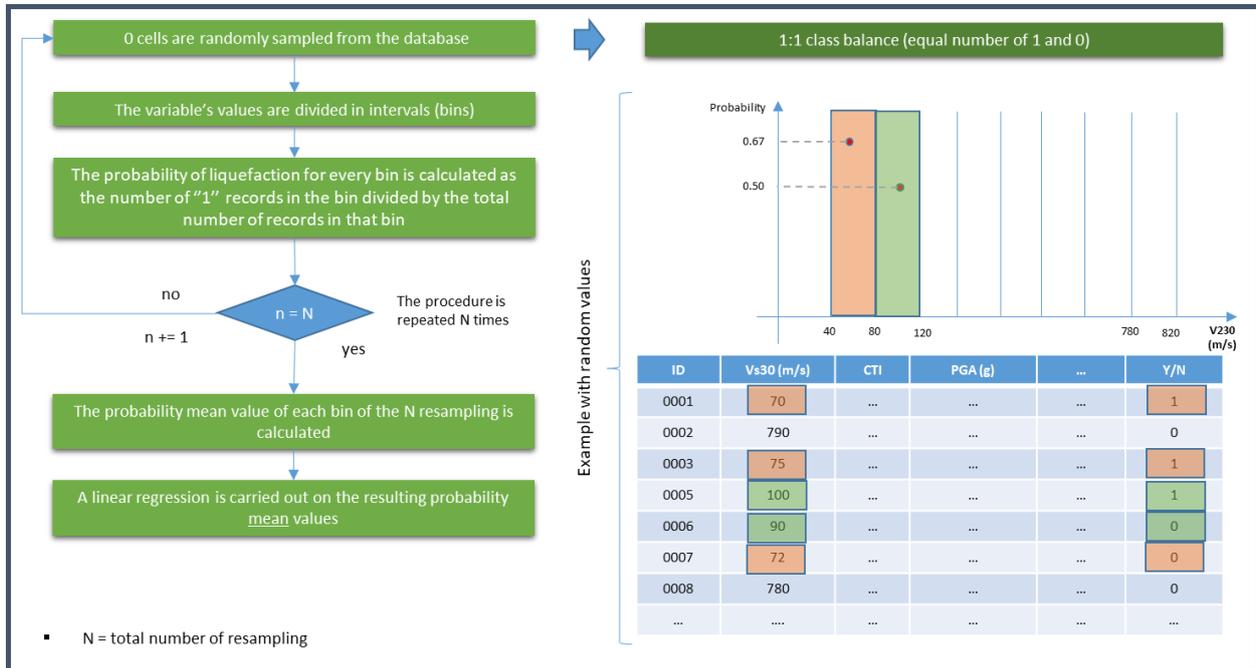


Figure 4-3: flowchart of the procedure adopted to plot the liquefaction probability values for each explanatory variable. An example is shown, where the values of two bins (highlighted in orange and green) are calculated.

For each resampling and for each explanatory variable, the variable's values are discretized in intervals (bins). The probability of liquefaction for every bin is calculated as the number of 1 records in the bin divided by the total number of records in that bin. This procedure is repeated for each resampling. At the end, a mean probability value is calculated for every bin. These values are finally plotted, and a linear regression is carried on.

The variables were evaluated also in alternative plots, such as a semi-logarithmic plot. Vs, PGA, PGAm, WBD, CD, RD showed better results by considering their natural logarithm, while TRI was considered with its square root ( $TRI^{0.5}$ ).

The results of the Luco&Cornell methodology are presented in Table 4-2 and in Table 4-3. The best variables, considering proficiency (which considers the composite effect of practicality and efficiency), resulted to be  $\ln(Vs)$ , CTI and  $\ln(PGAm)$ . This result is in accordance with the results of Zhu et al. (2015).

$\ln(PGA)$  also showed a good performance, comparable to that of  $\ln(PGAm)$ . In more general terms, efficiency was the criteria for which all the variables presented a good score. For what concern practicality, instead, has been observed a higher variability. Especially, TPI and  $TRI^{0.5}$  showed a poor performance.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

**Table 4-2: results of the Luco & Cornell methodology, in terms of practicality, efficiency and proficiency. The values for each criteria are marked with a color in a scale from green (good result) to red (poor result).**

Variable	Practicality*	Efficiency#	Proficiency#
CTI	0,13	0,13	1,02
Ln(Vs)	0,50	0,13	0,26
TPI	0,01	0,30	35,61
TRI <sup>0.5</sup>	0,00	0,16	144,85
Ln(WBD)	0,08	0,16	2,06
Ln(RD)	0,09	0,18	2,07
Ln(CD)	0,11	0,25	2,21
Ln(PGA)	0,19	0,24	1,24
Ln(PGAm)	0,18	0,22	1,21

\* The lowest the value the better the variable  
# The highest the value the better the variable

**Table 4-3: Classification of the variables for each criteria. The variables in each column are ordered from the top (best variables) to the bottom (worst variable).**

Practicality	Efficiency	Proficiency
Ln(Vs)	CTI	Ln(Vs)
Ln(PGA)	Ln(Vs)	CTI
Ln(PGAm)	TRI <sup>0.5</sup>	Ln(PGAm)
CTI	Ln(WBD)	Ln(PGA)
Ln(CD)	Ln(RD)	Ln(WBD)
Ln(RD)	Ln(PGAm)	Ln(RD)
Ln(WBD)	Ln(PGA)	Ln(CD)
TPI	Ln(CD)	TPI
TRI <sup>0.5</sup>	TPI	TRI <sup>0.5</sup>

## 4.4 Development of a prediction model

Once determined the best explanatory variables best correlated with liquefaction occurrence, special effort has been reserved to the development of a prediction model.

The logistic regression was employed to model the liquefaction probability. Logistic regression is a statistical approach for analyzing a dataset in which several independent variables determine a binary outcome. In this particular case, the outcome is represented by the liquefaction label (1 liquefaction - 0 no liquefaction) and the independent variables are the selected explanatory variables, namely CTI, Ln(PGAm) and Ln(Vs).

In logistic regression, liquefaction probability is expressed by the following expression:



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

$$P(X) = \frac{1}{1+e^{-X}} \quad (1)$$

Where  $X = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_k x_k$ , with  $x_k$  explanatory variables and  $\gamma_k$  coefficients of the regression calibrated upon the dataset.

The dataset, as already explained in the previous sections, is highly imbalanced. In this stage, two more strategies to overcome the problem were identified, in addition to the undersampling method.

- Undersampling: a number of 0 records such that the ratio is respected is randomly sampled from the database and a logistic regression is calibrated upon the resulting subset. The procedure is repeated  $n$  times and the mean values are extracted.
- SMOTE (Synthetic Minority Over-sampling Technique; Chawla et al. 2002): new minority records between existing (real) minority records are synthesized, in a number such that the ratio imposed is respected. The logistic regression is calibrated upon the resulting set.
- ADASYN (ADaptive SYNthetic; He et al., 2008): improved version of SMOTE, more synthetic data is generated for minority class examples that are harder to learn compared to those minority examples that are easier to learn.

ADASYN and SMOTE fall into the category of the oversampling methods.

The ratio between minority and majority class imposed was 1:2. The dataset was split into a training set and a test set (in each group is maintained the ratio between classes). The training set contains the records upon which the model is developed, while the test set contains the unseen data upon which the developed model is tested. The performance of each model is expressed in terms of AUC (Area Under the Curve) ROC (receiver Operating Characteristics) curve. In a Receiver Operating Characteristic (ROC) curve the true positive rate (Sensitivity, y-axis) is plotted in function of the false positive rate ( $1 - \text{Specificity}$ , x-axis). Each point on the ROC curve represents a sensitivity/specificity result corresponding to a particular decision threshold. The decision threshold is the probability value beyond which a data is classified as 1.

True positive rate is calculated as:

$$TPR = \frac{TP}{TP+FN} \quad (2)$$

and false positive rate is calculated as:

$$FPR = \frac{FP}{TN+FP} \quad (3)$$

where TP is the number of True Positive data (namely data labeled as 1 and classified by the model as 1), FP is the number of False Positive (namely data labeled as 0 and classified by the model as 0), TN and FN are the number of True Negative and False Negative (namely positive or negative data misclassified).

Subsequently, the AUC value is calculated as the area under the ROC curve. The AUC parameter ranges between 1, a perfect classifier, and 0.5, a random classifier. The highest the parameter is, thus, the better the model.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

The three model were compared, along with the model developed by Zhu et al. (2015). The results are shown in Table 4-4, where are reported the values of each coefficients of the logistic regression, given the general form:  $X = A + B \cdot \ln(PGAm) + C \cdot CTI + D \cdot \ln(Vs)$ , with X being the expression to be inserted in equation (1) to obtain the liquefaction probability value. AUC and optimal threshold are presented as well.

It can be seen how the best model resulted to be that developed with the ADASYN technique, although the others showed good performance as well. The Zhu et al (2015) model resulted the one with the poorest performance, but that was evident because it was calibrated with different data, while the others have been calibrated and tested on the same dataset. In the table is also present a column indicating for each model the perfect threshold at which the best performance is reached.

Table 4-4: Calibration results of the three logistic regression models along with the Zhu et al. (2015) model.

Method	A	B	C	D	AUC	Optimal threshold
ADASYN	-11.489	3.864	2.328	-0.091	0.95	0.57
SMOTE	30.281	2.348	0.22	-4.575	0.93	0.33
Undersampling	28.371	2.248	0.223	-4.279	0.91	0.41
Zhu 2015	24.1	2.067	0.355	-4.784	0.86	0.2

## 4.5 Adopted exposure indicators for Europe

The population density has been adopted as a proxy for exposure. As mentioned in Section 2, the population density data for Europe was obtained from the European GHSL database. The census data refer to the year 2015, and two different resolutions are available, 250m and 1km. The data are provided in a raster format, in which each cell contains the estimated number of inhabitants in that cell. The resolution adopted for this study is 1km to be consistent with the resolution of other input data. The raster map with a resolution of 1km represents the population density in terms of inhabitants/km<sup>2</sup> unit, which is the most common format to express the population density. Figure 4-4a shows an excerpt from the map of population density for Central Europe. The data were divided in 5 classes of exposure, as done for hazard. In particular:



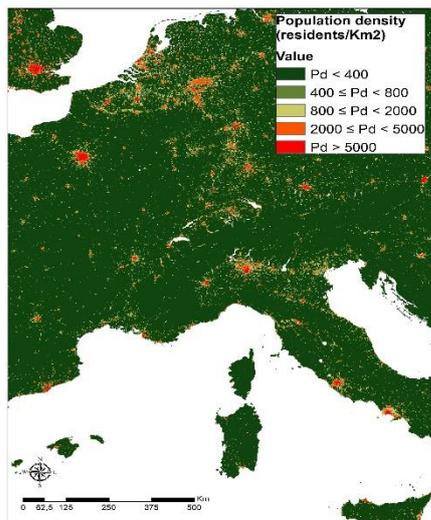
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

- Very low:  $Pd < 400$
- Low:  $400 \leq Pd < 800$
- Medium:  $800 \leq Pd < 2000$
- High:  $2000 \leq Pd < 5000$
- Very high:  $Pd \geq 5000$

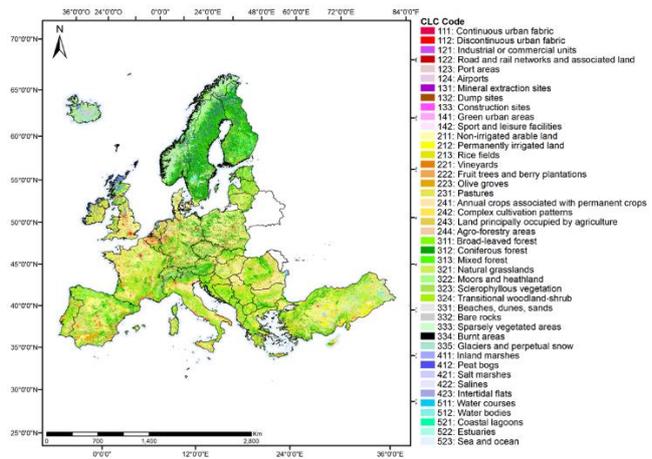
with Pd population density in pop./km<sup>2</sup>.

An additional second proxy for exposure was found in the CORINE land cover Europe map, which provides the geo-referenced inventory on land cover areas in Europe. The CORINE land cover Europe map is shown in Figure 4-4b. The data referred to the land use was particular helpful, in order to identify those areas with a high exposure that population density could not identify.

The population density data and the CORINE land cover were thus mixed. In particular, areas relative to airports, ports, roads and railways were assigned to the highest exposure class (very high). The final exposure model is shown in Figure 4-5.



(a)



(b)

Figure 4-4: (a) Population density map and (b) CORINE land cover map



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

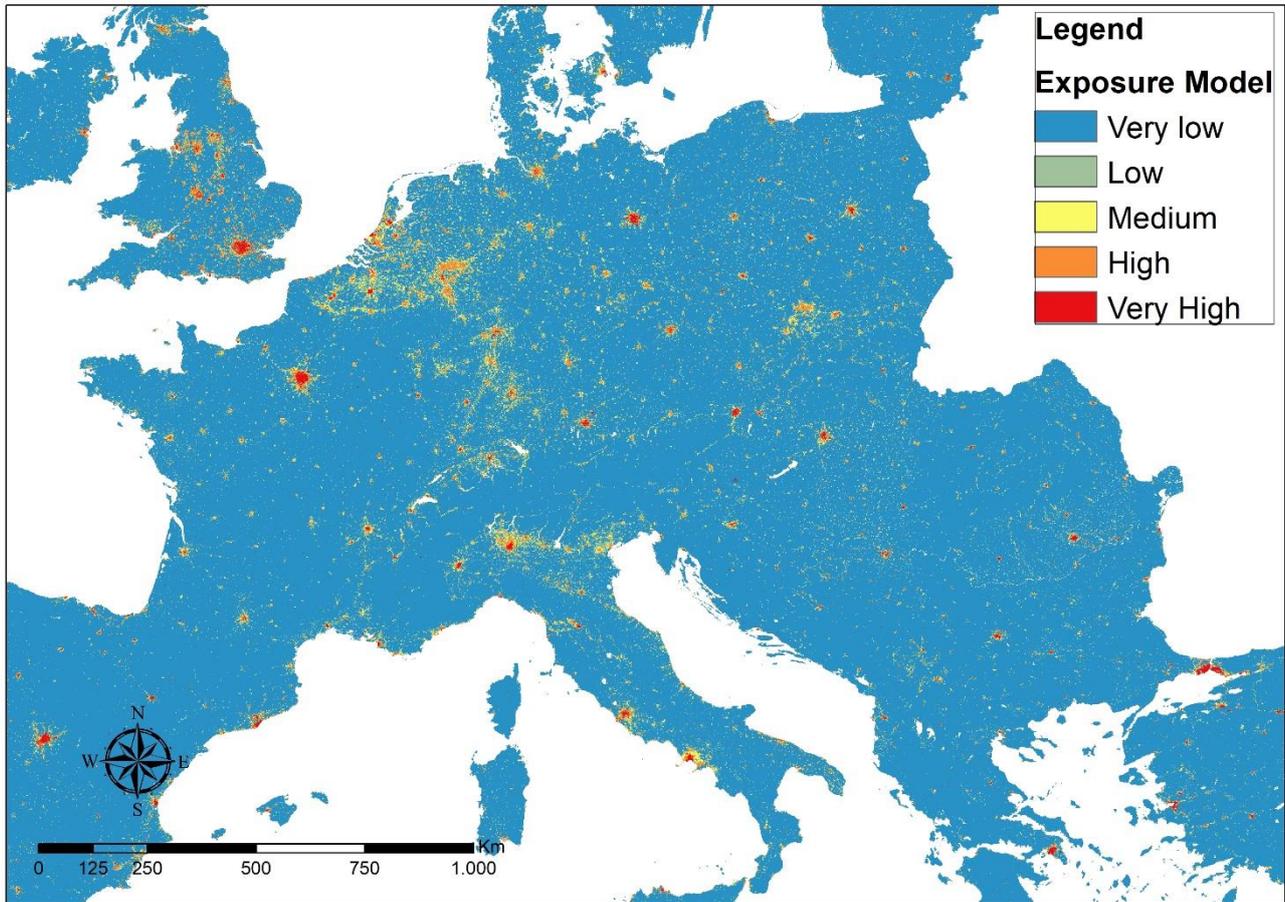


Figure 4-5: Final exposure Europe model, subdivided into 5 classes

## 4.6 Macrozoning the liquefaction risk in Europe

The assessment of large-scale risk connected to the soil liquefaction phenomenon is rarely treated in the literature. Recently, Yilmaz et al. (2018) performed a large-scale liquefaction risk assessment with reference to Portugal by extending simplified geotechnical methodologies to estimation of damage and economic losses within a probabilistic framework. Unfortunately, this approach can't be applied in the context of European risk map, because the lack of data needed by the methodology.

Since the risk is defined as the combination of hazard, vulnerability and exposure, the basic idea to compute a European liquefaction risk map is to combine geospatial data, available at continental scale, representing these three parameters.

The most suitable approach appeared to be the Analytical Hierarchy Process (AHP), a multi-criteria decision analysis technique, introduced by Saaty (1980) and then successfully applied to map the seismic hazard by Karimzadeh et al. (2014), Panahi et al. (2014) and Moustafa (2015). AHP can be defined a *knowledge-driven technique*, in which a set of explanatory variables are ranked, and their relative importance, in the light of a



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

certain objective, is assessed by assigning weights via calculation of a pairwise comparison matrix. The final map is calculated based on a weighted sum and ratings assignments via overlay operations. A shortcoming of the method is represented by the subjectivity of the assigned ranking which is therefore expert-based.

In the specific case, the explanatory variables are those representing hazard, vulnerability and exposure, and the objective, in the light of which the variables are compared, is the liquefaction risk. The more a variable influences risk, the higher will be its weight in the calculation of the final map.

The main steps of AHP method procedure is explained below.

In the first step, the alternatives are arranged in a GIS environment and their values are classified into different classes. The class are ranked, from the highest class (i.e. the value that has the greatest importance in the light of the objective) to the lowest.

In the second step of the methodology the data are organized in a matrix and the pairwise comparison of those alternatives on a qualitative scale is performed. Experts can rate the comparison as equal, marginally strong, strong, very strong, and extremely strong, as shown in Table 4-5. For example, the row corresponding to the alternatives A in the column corresponding to alternatives B presents the value 9 indicate that A is "Extremely strong" compared to B in the light of the objective. In general terms, the alternatives in the  $i_{th}$  row is stronger than that in the  $j_{th}$  column if the value of the matrix  $(i, j)$  is more than 1; otherwise the alternatives in the  $j_{th}$  column is stronger than that in the  $i_{th}$  row. Consequently, the  $(j, i)$  element of the matrix is the reciprocal of the  $(i, j)$  element.

**Table 4-5: Relative importance for comparison between alternatives (Saaty, 1980).**

<b>Weight/rank</b>	<b>Relative importance</b>
<b>1</b>	equal
<b>3</b>	moderately dominant
<b>5</b>	strongly dominant
<b>7</b>	very strongly dominant
<b>9</b>	extremely dominant
<b>2,4,6,8</b>	intermediate values
<b>Reciprocals</b>	for inverse judgements

The third step consist in the calculation of the principal eigenvalue and the corresponding normalised right eigenvector of the comparison matrix built at step 2. The elements of the normalised eigenvector are termed weights with respect to the objective and the comparison of the alternatives.

In the fourth step the consistency of the matrix built is evaluated. Indeed, the alternatives comparisons made in this method are subjective and the AHP tolerates inconsistency through the amount of redundancy in the



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

approach. If this consistency index results lower than required level, the comparisons may be re-examined. The consistency index, CI, is calculated as:

$$CI = (\lambda_{max} - n)/(n - 1) \quad (2)$$

where  $\lambda_{max}$  is the maximum eigenvalue of the judgement matrix and n is the dimension of the matrix. CI is then compared with that of a random matrix, RI. The ratio derived, CI/RI, is termed the consistency ratio CR. Saaty suggests that the upper threshold value of CR should be 0.1.

In the final step, the value of each alternatives is multiplied by its own weight. Subsequently the values obtained are summed up and the final rank is calculated. This last step is developed in a GIS environment. The alternatives are represented by overlapped raster files; every pixel of each raster contains a value calculated in the first step. The final raster consists in a map representing in each pixel the sum of the value contained in the pixels of the alternatives.

The results will be inherently coarse, for at this scale is out of the scope to reach high-detailed maps. The aim of the maps is, instead, to distinguish areas that may that are likely to experience soil liquefaction in case of strong ground shaking from areas where liquefaction is unlikely. These maps should be used with caution as they only provide a rough identification of the territories in Europe that may be affected by earthquake-induced liquefaction.

## 4.7 Filtering of the maps

The maps for both hazard and risk were filtered in order to exclude a priori from the analysis those area that either are characterized by non-susceptible soils or by a PGA value not high enough to trigger liquefaction.

The soil filter, as already explained in Section 3.2, was obtained using the lithological information derived from (i) the International Hydrogeological Map of Europe 1:1,500,000 (IHME1500), (ii) the Geological Map of Turkey 1:2.500.000 compiled by the General Directorate of Mineral Research and Exploration (MTA) and (iii) the Dominant parent material information available in the Soil Geographical database of Eurasia at scale 1:1.000.000 (Panagos et al. 2012, see Figure 3-2). In particular, the lithological classes considered as non-susceptible soils to liquefaction are the bedrock and clayey soils.

The filter referred to the PGA was obtained from the SHARE PGA maps, imposing a threshold value for PGA was equal to 0.1g. This assumption was based on recommendations from the literature (e.g. Italian Building Code, NTC2018). Therefore, for PGA values smaller than 0.1g, liquefaction occurrence is very unlikely. The filters were implemented in the GIS environment where the risk maps for Europe are computed. Therefore, it was assumed that liquefaction occurrence is very unlikely at any site of the European territory where the expected PGA is smaller than 0.10g (Green and Bommer, 2018). The unfiltered maps could display a medium or, in some cases, a high liquefaction hazard and/or risk level.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

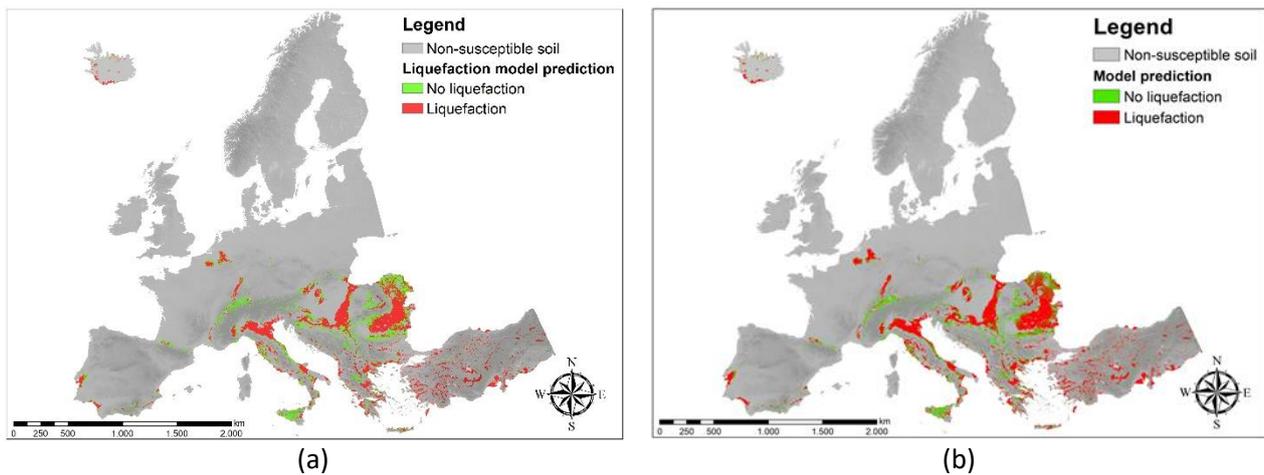
## 5. MAPS DISPLAYING THE LIQUEFACTION HAZARD AND RISK AT CONTINENTAL SCALE

### 5.1 Liquefaction hazard maps for Europe

The three models developed were employed to produce liquefaction hazard maps in Europe, with PGAm referred to the return period of 475, 975 and 2475 years obtained from SHARE. The peak ground acceleration values extracted from SHARE with reference to standard ground conditions (outcropping bedrock and level site) was multiplied to the soil coefficient of Eurocode 8 Part 1 (hereinafter, EC8) to take into account site effects. Ground categories of EC8 were assigned on the basis of  $V_s30$ .

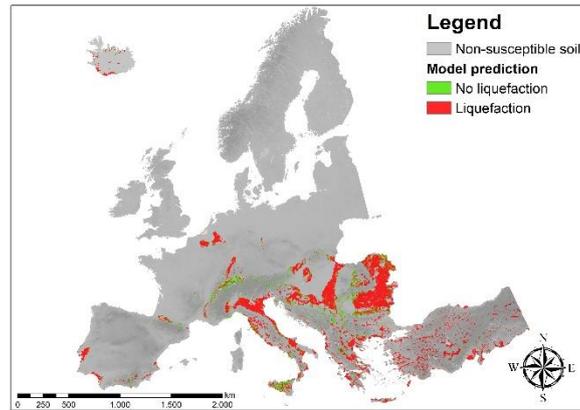
The maps are showed in Figure 5-1, with the filters exposed in Section 4.7. Those maps represent a binary liquefaction prediction, namely they distinguish areas where the model predicts liquefaction occurrence (i.e. 1) from areas where the model predicts non-occurrence of the phenomenon (i.e. 0). Being the output of the logistic regression model a probability, in order to obtain binary maps, the optimal thresholds of Table 4-4 were employed.

The ADASYN model was selected as the final model, being the model that showed the best performance. The model was implemented to produce a continuous map, which is a map displaying the probability value obtained from the logistic regression, divided into 5 different hazard classes (very low, low, medium, high, very high). These maps are shown in Figure 5-2.



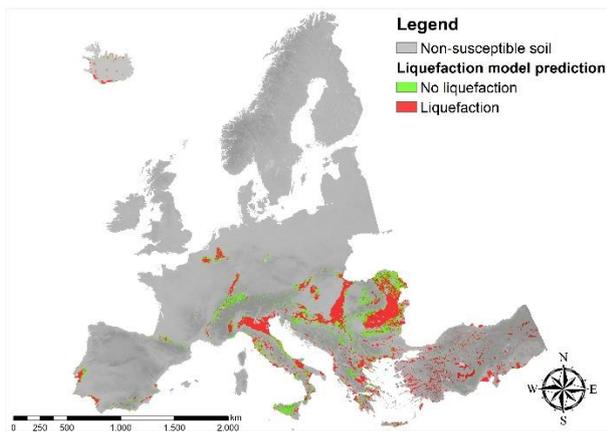


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

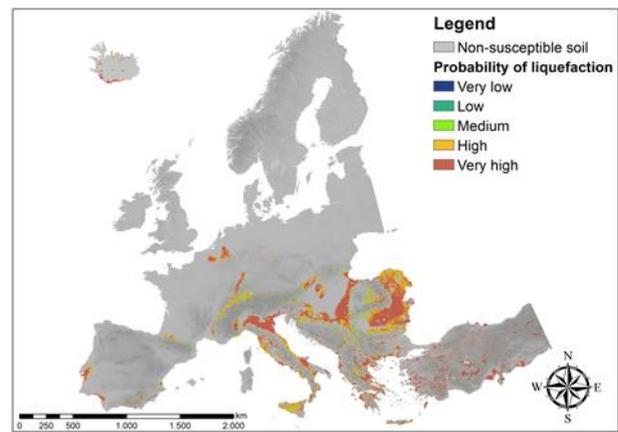


(c)

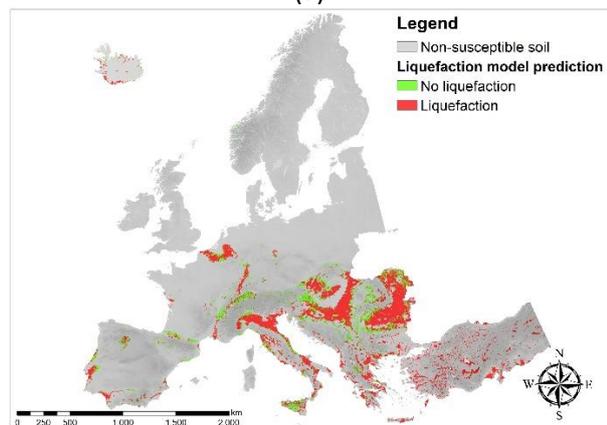
Figure 5-1: Maps showing the binary liquefaction prediction relative to the three models developed, (a) ADASYN, (b) SMOTE, (c) Undersampling with a PGA relative to the 475 return years period



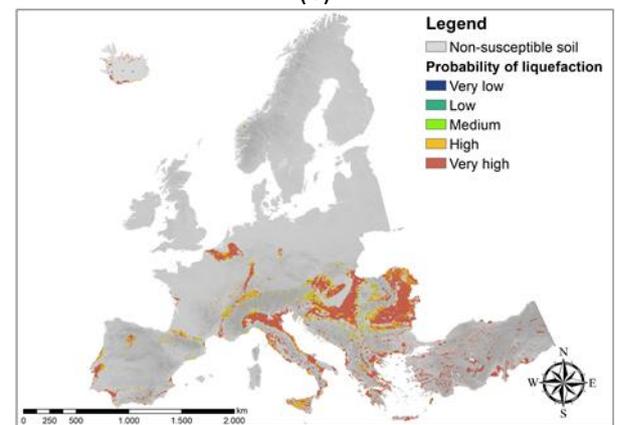
(a)



(b)



(c)



(d)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

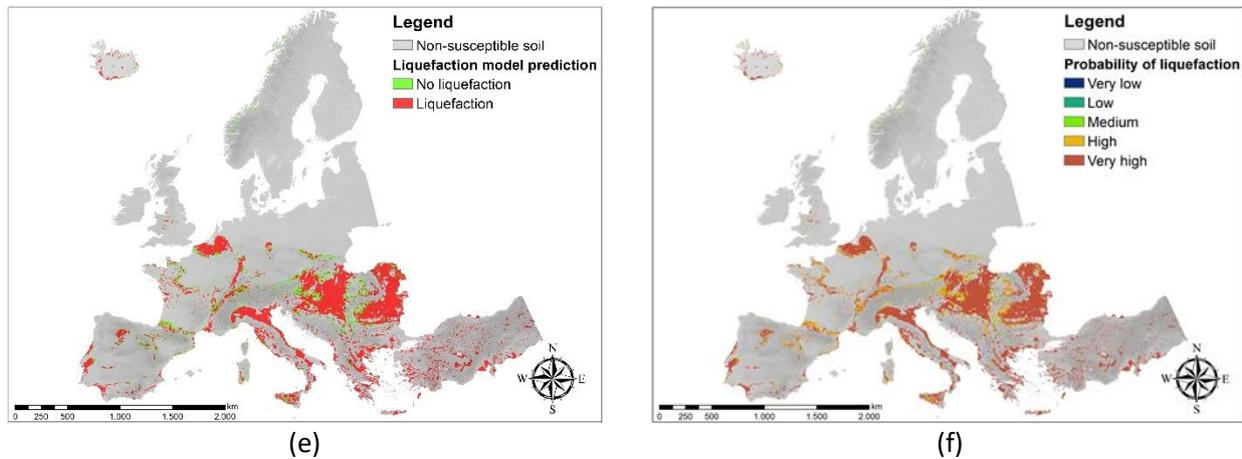


Figure 5-2: Maps developed with the ADASYN model. For each return period is presented a binary model prediction map and a probability of liquefaction map. Specifically, (a) and (b) refers to 475 years return period, (c) and (d) refers to 975 years return period, (e) and (f) refers to 2475 years return period.

## 5.2 Validation by superimposing historical liquefaction occurrences

A validation of the outcomes of the liquefaction hazard maps of Europe obtained with ADASYN is carried out by superimposing the computed maps to purposely selected cases from the GIS-based catalogue of liquefaction occurrences in Europe, already presented in Section 2.1. The European catalogue includes about 1000 manifestations of liquefaction phenomenon. A return period was associated to the events of the catalogue by using a procedure based on the identification of the sequences through the Gardner and Knopoff (1974) algorithm. The return period of each mainshock was calculated based on seismogenic zoning used in the SHARE project. The return period of each mainshock was then associated to the entire sequence. Figure 5-3 shows the number of manifestations of liquefaction grouped for different ranges of return periods. It seems reasonable to expect that the liquefaction cases increase when the return period increases, because the magnitude of the earthquake increases. However, the numbers of liquefaction manifestations associated to higher return periods decrease in the graph of Figure 5-3. This may be explained by considering that the manifestations of liquefaction phenomenon in many European Countries were collected only in the last centuries, except for the case of Italy whose catalogue spans a period starting in 1117 (Figure 2). The role played by the completeness periods of the earthquake catalogue associated to different magnitude bins is currently under investigation.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

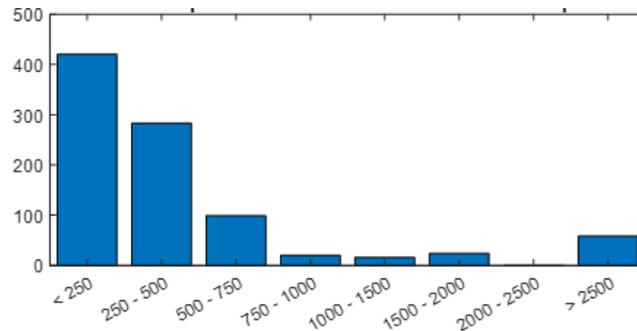


Figure 5-3: Number of manifestations of earthquake-induced liquefaction included in the European earthquake catalogue grouped for different ranges of return periods.

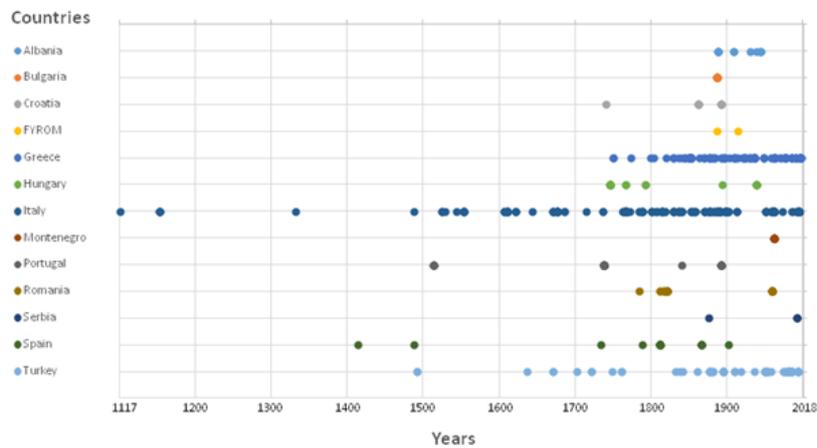
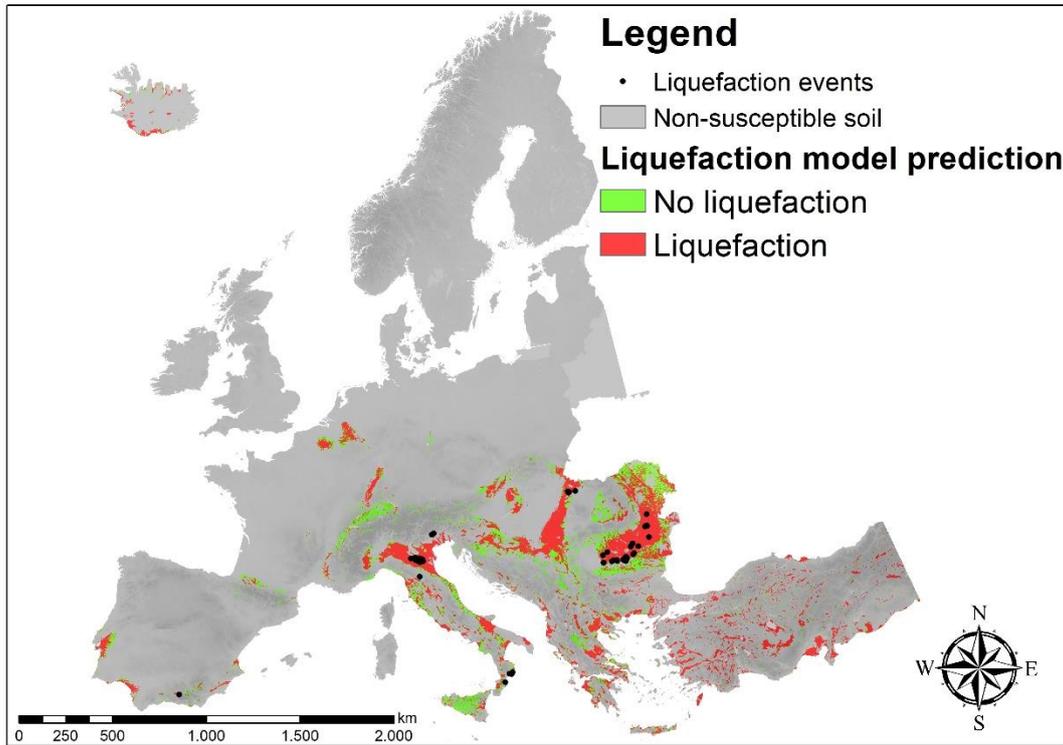


Figure 5-4: Distribution of manifestations of earthquake-induced liquefaction included in the catalogue for which cases were collected.

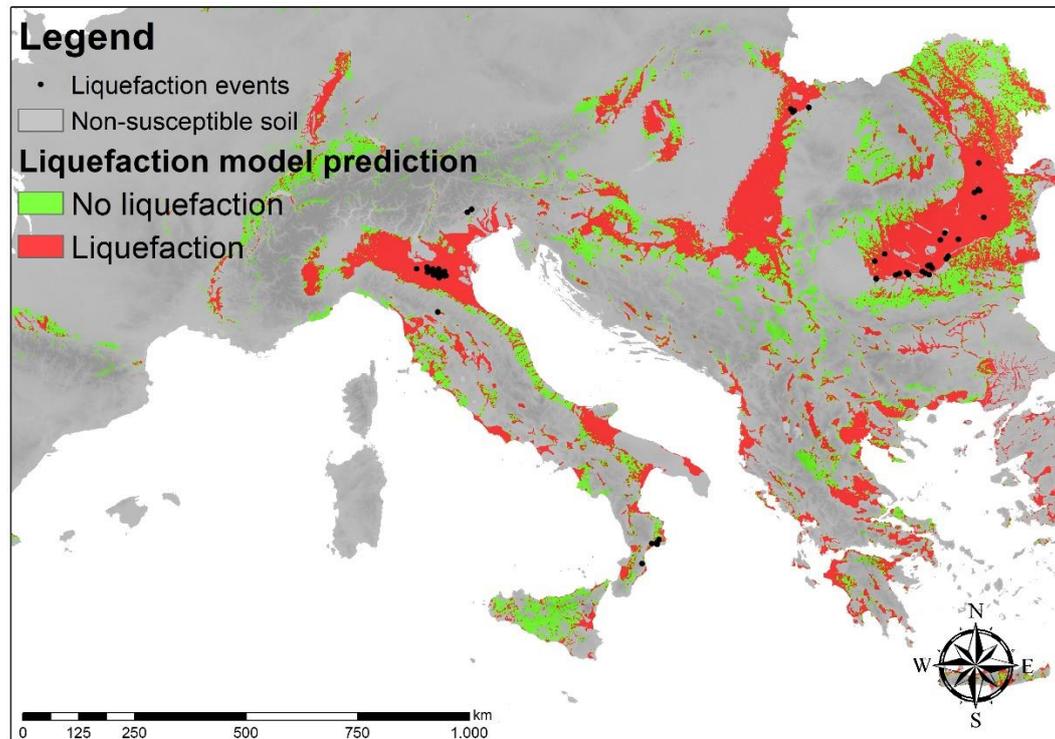
The validation was carried out, thus, only for the 475 years maps. In the maps of Figure 5-5 and Figure 5-6 the locations of the manifestations of soil liquefaction associated to a return period of about (i.e. +/- 10%) 475 years are superimposed. The liquefaction historical cases are mainly located within territory at high probability of liquefaction. This is particularly evident in the Balkan region and in Emilia-Romagna region in Italy (Figure 5-5b and Figure 5-6b).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748



(a)

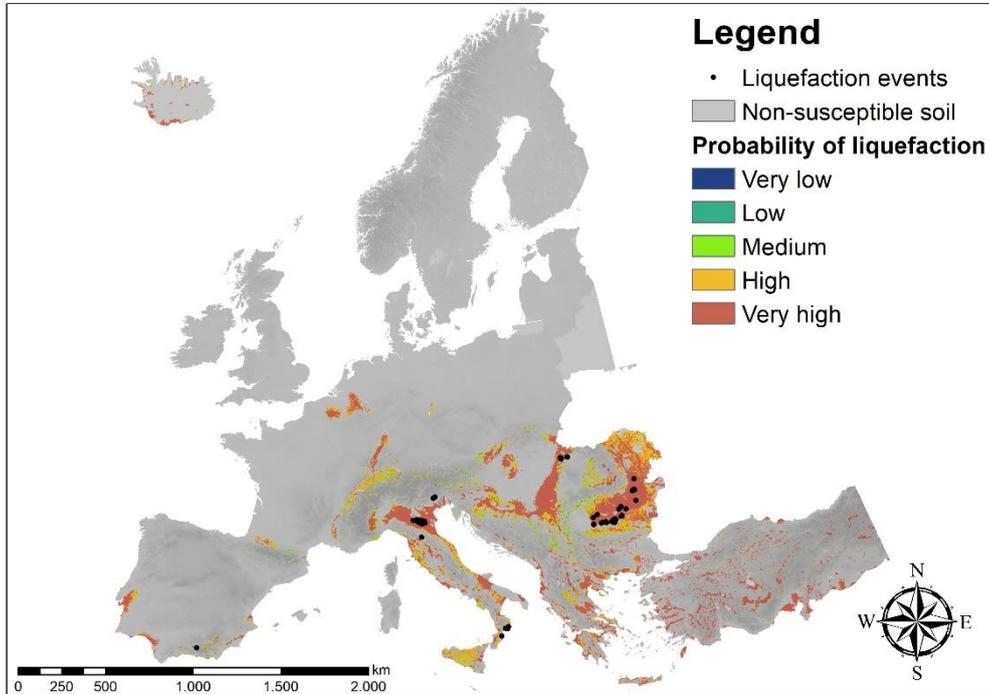


(b)

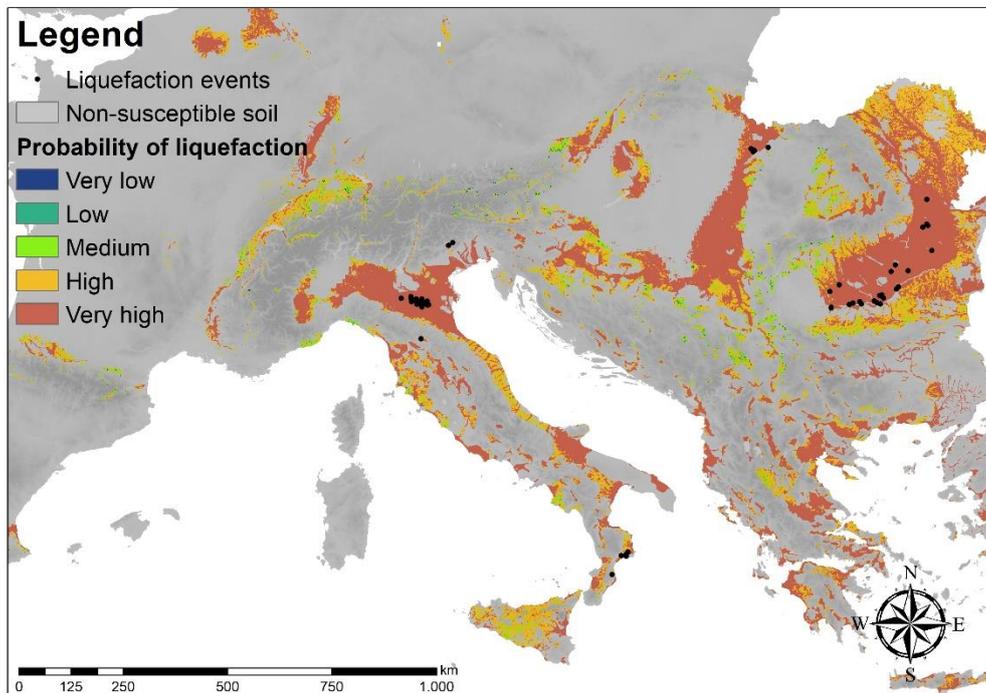
Figure 5-5: Validation maps for the 475 years return period, which show (a) the binary model prediction map with the liquefaction events superimposed and (b) a zoom to the Italian/Balkan region



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748



(a)



(b)

Figure 5-6: Validation maps for the 475 years return period, which show (a) the liquefaction probability map with the liquefaction events superimposed and (b) a zoom to the Italian/Balkan region



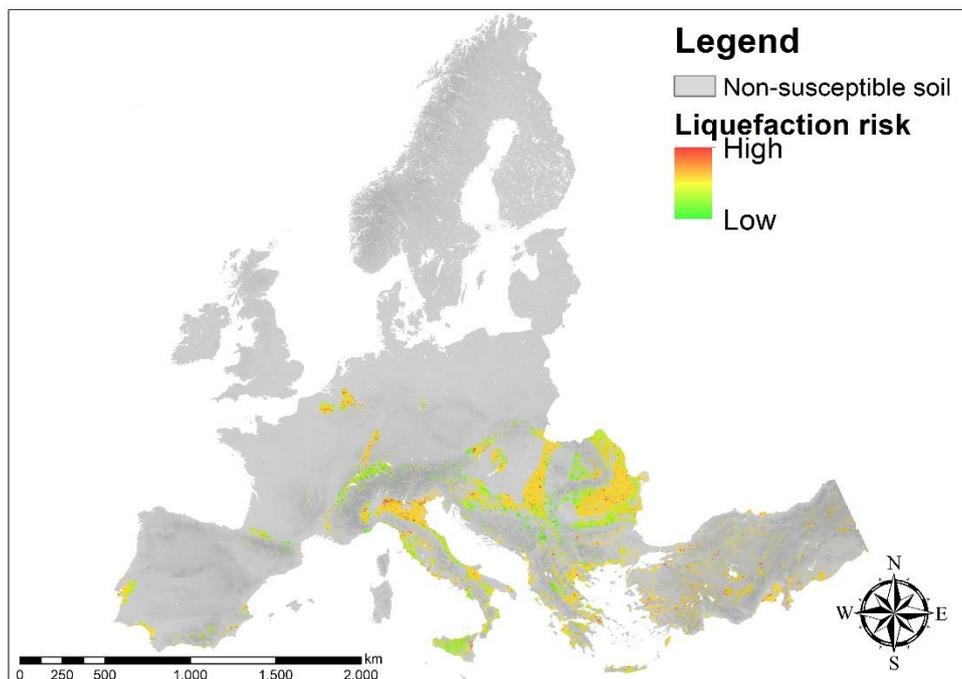
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

### 5.3 European maps of liquefaction risk

The AHP method was applied in order to merge the hazard and the exposure model. The hazard data, represented by the probability of liquefaction calculated with the ADASYN model, and the exposure data, represented by the combination of population density and CORINE land cover (see Section 4.5), where divided into 5 different classes (see Sections 4.5 and 5.1).

The AHP method was subsequently applied considering thus two variables, following the steps presented in Section 4.6. The hazard (represented by the liquefaction probability) was considered moderately dominant, referring to Table 4-5. This resulted in a weight of 0.75 assigned to hazard and a weight of 0.25 assigned to exposure.

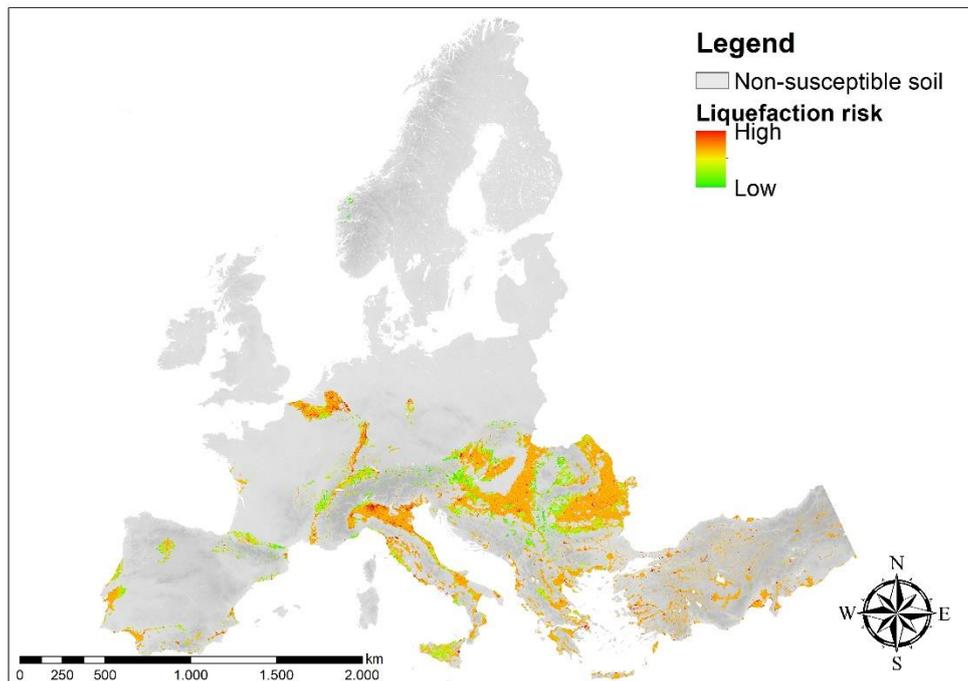
The final risk maps, for the three return periods, are shown in Figure 5-7.



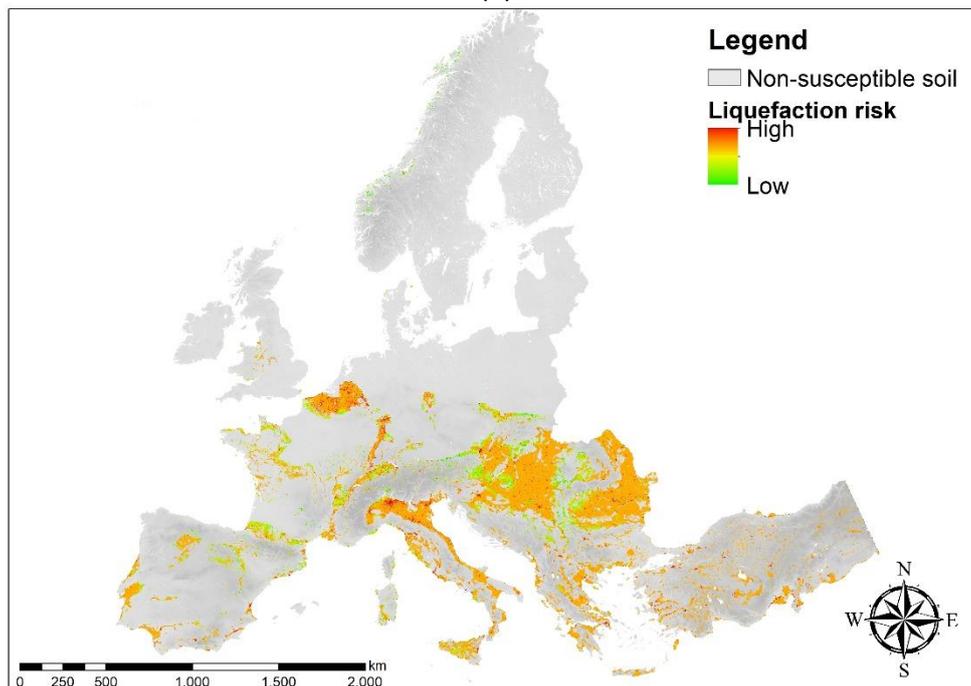
(a)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748



(b)



(c)

Figure 5-7: Maps showing European liquefaction risk for (a) 475 years, (b) 975 years, (c) 2475 years return period.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

LIQUEFACT  
Deliverable 2.6

*Report to describe the adopted procedure for the development of the European liquefaction hazard map*  
v. 1.0

## 6. CONCLUDING REMARKS

The Task 2.5 of the Work Package 2 of the LIQUEFACT project sets his aim in the development of a liquefaction risk macrozonation of the European territory. Being the liquefaction a strongly local phenomenon, the objective of the Task 2.5 showed an intrinsic problem that could only be overcome with the assumption of some simplifications. It is apparent, thus, that the final maps developed can't be employed to derive detailed information at small scale, but they find their purpose in providing an idea at a glance on which macro-regions could experience the phenomenon and could be affected in terms of potential losses.

The Task was structured in different stages and established its starting point on the geological, geomorphological and seismological data (the explanatory variables) collected within Task 2.2, and on the liquefaction events catalogue developed within Task 2.3.

In the first step a dataset was built, employing the data collected. The second stage of the task had the objective of establish which explanatory variables are best correlated with the liquefaction manifestations. This stage was accomplished applying the Luco & Cornell methodology.

Subsequently, with the best variables extracted from the analysis at the previous step, a liquefaction prediction model was calibrated. The variables selected were the natural logarithm of the PGAm, CTI and the natural logarithm of the Vs30. The statistical model chosen was the logistic regression. The model developed was applied, and three liquefaction hazard maps, for the return period of 475 years, 975 years and 2475 years, were produced.

Finally, to produce risk liquefaction maps, an exposure model was developed combining together population density and CORINE land cover. The model was subsequently combined with the hazard model and three liquefaction risk maps for the three return periods were developed.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

## REFERENCES

- Ambraseys, N. N.: The Kresna earthquake of 1904 in Bulgaria, *Annali di Geofisica*, 44(1), 95–117, 2001.
- Cetin, K. O., Youd, T. L., Seed, R. B., Bray, J. D., Stewart, J. P., Durgunoglu, H. T., ... & Yilmaz, M. T. (2004). Liquefaction-induced lateral spreading at Izmit Bay during the Kocaeli (Izmit)-Turkey earthquake. *Journal of geotechnical and geoenvironmental engineering*, 130(12), 1300-1313.
- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P. 2002. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Cornell, C.A., Luco, N. 2001. Ground motion intensity measures for structural performance assessment at near-fault sites. Proceedings U.S.-Japan Joint Workshop and Third Grantees Meeting, U.S.-Japan Coop. Res. on Urban EQ. Disaster Mit. Seattle, Aug. 15-16 2001, Univ. of Washington.
- Daja, S., Shkodrani, N., Lako, A., & Ago, B. (2013). Real liquefaction probability of non-cohesive soils in Semani area, in Albania. *Italian Journal of Geosciences*, 132(2), 213-219.
- Eurocode 8, 2003. Design of structures for earthquake resistance, Part 1: General rules, seismic actions and rules for buildings, Pr-EN1998-1, European Committee for Standardization (CEN), Brussels.
- Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, 339(6122), 940-943.
- Galli, P. (2000). New empirical relationships between magnitude and distance for liquefaction. *Tectonophysics*, 324(3), 169-187.
- Green, R.A., Bommer, J.J. 2018. Smallest Earthquake Magnitude that Can Trigger Liquefaction, Report of a study performed by the Virginia Tech Center for Geotechnical Practice and Research, March 2018, CGPR # 92.
- Györi, E., Tóth, L., & Mónus, P. (2015). Secondary effects generated by earthquakes: liquefaction occurrences in and around Hungary. *Acta Geodaetica et Geophysica*, 50(1), 79-95.
- He, H., Bai, Y., Garcia, E.A., Li, S. (2008). ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. 2008 International Joint Conference on Neural Networks (IJCNN 2008).
- Herak, D., & Herak, M. (2010). The Kupa Valley (Croatia) Earthquake of 8 October 1909—100 Years Later. *Seismological Research Letters*, 81(1), 30-36.
- Jorge, C. R., Coelho, A. G. (1994). A liquefaction potential zoning map of Portugal. Sociedade Portuguesa de Geotecnia | International Association of Engineering Geology, 1994.
- Karimzadeh S., Miyajima M., Hassanzadeh R., Amiraslanzadeh R., and Kamel B. A GIS-based seismic hazard, building vulnerability and human loss assessment for the earthquake scenario in Tabriz. *Soil Dynamics and Earthquake Engineering*, 66 (10), pp. 263-280; 2014.
- Kociu, S. (2004). Induced seismic impacts observed in coastal area of Albania: case studies.
- Lanfredi Sofia, C., Oliveira, S. C., Pereira, S., Zêzere, J. L., & Corsini, A. (2018). A comparison between bivariate and multivariate methods to assess susceptibility to liquefaction-related coseismic surface effects in the Po Plain (Northern Italy). *Geomatics, Natural Hazards and Risk*, 9(1), 108-126.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700748

- Luco, N., Cornell, C.A. (2007). Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earth. Spectra*, 23(2), 357-392.
- Meybeck, M., Green, P., & Vörösmarty, C. (2001). A new typology for mountains and other relief classes. *Mountain Research and Development*, 21(1), 34-46.
- Moustafa S.S.R. (2015). Application of the Analytic Hierarchy Process for Evaluating Geo-Hazards in the Greater Cairo Area, Egypt, *EJGE*, Vol. 20 [2015], Bund. 6; 2015.
- NTC (2018). Norme Tecniche per le Costruzioni. Ministero delle Infrastrutture e dei Trasporti, Decreto Ministeriale del 17 gennaio 2018, Supplemento ordinario alla G.U. n. 8 del 20 febbraio 2018.
- Padgett, J.E., Nielson, B.G., DesRoches, R. (2008). Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earth. Eng. and Struct. Dyn.*, 37(5), 711-725.
- Panagos P., Van Liedekerke M., Jones A., Montanarella L. (2012). European Soil Data Centre: Response to European policy support and public data requirements. (2012) *Land Use Policy*, 29 (2), pp. 329-338. doi:10.1016/j.landusepol.2011.07.003.
- Panahi M., Rezaie F., Meshkani S. (2014). Seismic vulnerability assessment of school buildings in Tehran city based on AHP and GIS. *Nat. Hazards and Earth System Sciences Discus.*, 1, pp. 4511-4538; 2014.
- Papathanassiou, G., & Pavlides, S. (2011). GIS-based database of historical liquefaction occurrences in the broader Aegean region, DALO v1. 0. *Quaternary international*, 242(1), 115-125.
- Papathanassiou, G., Ganas, A., & Valkaniotis, S. (2016). Recurrent liquefaction-induced failures triggered by 2014 Cephalonia, Greece earthquakes: spatial distribution and quantitative analysis of liquefaction potential. *Engineering geology*, 200, 18-30.
- Papathanassiou, G., Valkaniotis, S., Chaztipetros, A., & Pavlides, S. (2010). Liquefaction susceptibility map of Greece. *Bulletin of the Geological Society of Greece*, 43(3), 1383-1392.
- Plant, J. A., Whittaker, A., Demetriades, A., De Vivo, B., & Lexa, J. (2003). The geological and tectonic framework of Europe. *Geochemical Atlas of Europe. Part, 1*.
- Plant, J.A., Reeder, S., Salminen, R., Smith, D.B., Tarvainen, T., De Vivo, B. & Petterson, M.G., (2003). The distribution of uranium over Europe: geological and environmental significance. *Transactions of the Institution of Mining and Metallurgy, Section B*, 112 (3), 221-238.
- Saaty, T.L. 1980. *The analytic hierarchy process*. McGraw-Hill, New York.
- Veinovic, Z., Kvasnicka, P., & Domitrovic, D. (2010). Preparatory works for liquefaction microzonation in Croatia. *STRATEGIES FOR REDUCTION OF THE SEISMIC RISK*, 45.
- Yilmaz, C., Silva, V., Weatherill, G., Rathje, E. (2018). Probabilistic seismic loss estimation due to ground failure. *Proceedings 16th European Conference on Earthquake Engineering*, 16ECEE, Thessaloniki, Greece, 18-21 June, 2018.
- Zhu, J., Baise, L., Thompson, E. (2017). An Updated Geospatial Liquefaction Model for Global Application. *Bulletin of the Seismological Society of America*, 107(3), 1365-1385.
- Zhu, J., Daley, D., Baise, L., Thompson, E., Wald, D., Knudsen, K. (2015). A Geospatial Liquefaction Model for Rapid Response and Loss Estimation. *Earthquake Spectra*, 31(3), 1813-18