Assessment of liquefaction effects on ground motion frequency parameters for accelerogram-based liquefaction detection

Weiwei Zhan, S.M.ASCE,[1] and Qiushi Chen, A.M.ASCE[2]

[1]The Glenn Department of Civil Engineering, Clemson University, Clemson, SC. E-mail: wzhan@clemson.edu
[2]Associate Professor, The Glenn Department of Civil Engineering, Clemson University, Clemson, SC, (corresponding author). E-mail: qiushi@clemson.edu

ABSTRACT

An accelerogram-based method has been recently developed to quickly assess the liquefaction occurrence using only the ground motion records. In this method, two frequency-related ground motion parameters, the richness of the low-frequency components (termed RL), and the temporal variation rate of the mean instantaneous frequency (termed MIFr), are extracted from accelerograms and used as explanatory variables of a logistic regression model that outputs a new probability indicator of liquefaction occurrence (termed LQI). This method achieves an overall accuracy of over 90% when classifying the liquefaction occurrence in both the training dataset and the validation dataset, and the method shows promising potential for applications in real-time disaster mitigation systems and rapid post-earthquake loss estimations. However, the success of detecting liquefaction from accelerograms relies on the assumption that the liquefaction effect on the site-specific ground motions is more significant than other effects. In this study, we select eight factors representing the source, path, and site effects of a ground motion record, and conduct correlation analyses between the eight independent factors and the proposed two frequency-related ground motion parameters (RL and MIFr). The results reveal that, in general, liquefaction has a more significant effect on the site-specific ground motion frequency parameters compared to other factors such as earthquake source, path and site effects, which is the basis of the accelerogram-based liquefaction assessment methods. This study will benefit the advancement of the accelerogram-based methods for rapid liquefaction assessment.

INTRODUCTION

Soil liquefaction is one of the secondary hazards induced by earthquakes and is continually threatening many urban areas. Liquefaction can not only induce ground deformations but also modify the amplitude and frequency characteristics of surface ground motions (Bonilla et al.2005; Youd and Carter 2005; Kramer et al. 2015; Gingery et al. 2015). Given the fact that liquefaction

modifies ground motions, we can utilize the existing seismic network for real-time liquefaction detection after an earthquake and rapid hazard response, an example of which is the SUPREME system in Japan (Shimizu et al. 2000).

Recently, we develop an accelerogram-based method to quickly assess the liquefaction occurrence based on the two frequency parameters easily extracted from ground motion records. The first parameter is the ratio of the low-frequency portion to the whole area of the Fourier amplitude spectrum (termed RL), which represents the relative richness of the low-frequency components in a ground motion record. The second parameter is the mean instantaneous frequency decrease rate (termed MIFr), which represents the temporal variation rate of the mean instantaneous frequency. A logistic regression model that outputs a new probability indicator of liquefaction occurrence (termed LQI) is developed and validated based on a worldwide database consisting of 167 pairs of ground motion and liquefaction observations (summary data in Table 1). The method achieves an overall accuracy of 92.8% on the training dataset consisting of 135 case histories from the original K-Y dataset (Kostadinov and Yamazaki 2001), 2010 Mw 7.1 Darfield earthquake, and 2011 Mw 6.2 Christchurch earthquake, and the method achieves an overall accuracy of 90.6% on the validation dataset consisting of 32 case histories from the 2001 Mw 6.8 Nisqually earthquake. Figure 1 shows the proposed model plotted against liquefaction data in the compiled datasets. The method shows promising potentials for applications in real-time disaster mitigation systems and rapid post-earthquake loss estimations. However, the success of detecting liquefaction from accelerograms relies on the assumption that liquefaction effect on the site-specific ground motions is more significant than other effects, such as earthquake source, propagation path, and general seismic site response. In this work, we aim to analyze and compare the correlations between the two strong motion frequency parameters, RL and MIFr, and the factors representing the earthquake source, path, and site effects on the site-specific ground motions.

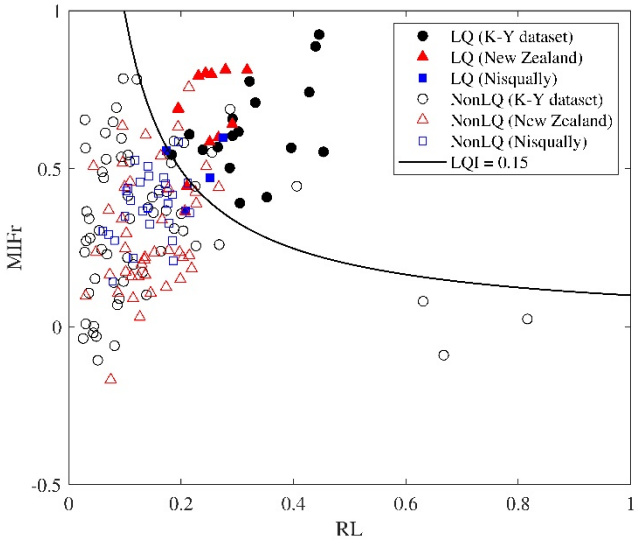

Figure 1. Liquefaction classification results of the proposed accelerogram-based method

Table 1. Ground motion recordings and associated earthquake data

ID	Name	UTC Date[a]	Country	Depth (km)	Mw	Fault type[b]	Event type[c]	LQ# [d]	NLQ# [e]
1	Niigata	6/16/1964	Japan	10	7.6	T	2	1	0
2	Tokachi-Oki	5/16/1968	Japan	13.2	8.2	T	0	1	2
3	Miyagiken-Oki	6/12/1978	Japan	48.5	7.6	R	0	0	1
4	Nihonkai-Chibu	5/26/1983	Japan	16.5	7.7	R	2	1	1
5	Michoacan	9/19/1985	Mexico	15.6	8	R	0	0	5
6	Superstition Hills	11/24/1987	USA	9	6.6	SS	2	1	1
7	Chibaken-Toho-Oki	12/17/1987	Japan	45.1	6.5	SS	1	0	5
8	Loma Prieta	10/18/1989	USA	17.5	7	RO	2	1	11
9	Kushiro-Oki	1/15/1993	Japan	92.3	7.6	N	1	1	7
10	Hokkaido-Nansei-Oki	7/12/1993	Japan	12	7.7	T	0	0	4
11	Northridge	1/17/1994	USA	17.5	6.7	R	2	0	10
12	Hokkaido-Toho-Oki	10/4/1994	Japan	29.2	8.2	T	1	0	4
13	Sanriku-Haruka-Oki	12/28/1994	Japan	27	7.7	T	0	0	1
14	Hyogoken-Nanbu	1/16/1995	Japan	17.9	6.9	SS	2	11	10
15	Kagoshimaken	5/13/1997	Japan	7.7	5.9	SS	2	0	4
16	Darfield	9/3/2010	New Zealand	11	7.2	R	2	2	31
17	Christchurch	2/21/2011	New Zealand	5	6.2	R	2	9	12
18	Nisqually	2/28/2001	USA	52.4	6.8	N	1	4	28

Notes:

[a] Earthquake origin time in Coordinated Universal Time (UTC).

[b] T = thrust; R = reverse; SS = strike slip; N = normal; RO = reverse oblique.

[c] 0 = interface; 1 = intraslab; 2 = shallow crustal.

[d] Number of liquefied-site records collected from corresponding earthquake in this work.

[e] Number of non-liquefied-site records collected from corresponding earthquake in this work.

FACTOR SELECTION AND PROCESSING

In this study, we compiled a total of 167 ground motion records from the 18 worldwide earthquakes (see Table 1), among which 32 records are from liquefied stations and 135 records are from non-liquefied stations. We compute RL and MIFr for each ground motion records. To analyze the correlations between the two frequency parameters with source, path, and site effects, we select eight factors representing these three effects based on two criteria: 1) the factor can affect the seismic wave propagation and hence the site-specific ground motion, and 2) the factor is acquirable. Liquefaction is considered a factor belonging to the site effect. To ensure the comparability of effects from different types of factors, we convert factors that are a continuous variable to a categorical variable (e.g., converting earthquake magnitude to magnitude class) using universally accepted criteria or new criteria inferred from the data. The eight factors and classification criteria are summarized in Table 2, and their selection and extraction processes are detailed in the following.

We take four factors to represent the earthquake source effect, including the event type, fault type, magnitude, and focal depth. The event type (the term used in the NGA-Subduction

project (Bozorgnia et al. 2020) is also called the tectonic environment, which is usually treated separately in ground-motion prediction models as it can affect the source characteristics and traveling paths of seismic waves (Van et al. 2017). The event type of the earthquakes in Table 1 has been grouped by three classes: 0 = Interface, 1 = Intraslab, 2 = Shallow crustal, as reported in Bozorgnia et al. (2020) and Gingery et al. (2015). The exceptions are the 1994 Northridge and 1997 Kagoshimaken earthquakes (No.11 and No.15 events in Table 1), which were not reported in the previous two studies, and are classified in this work following the same classification groups. The fault type is grouped into five classes: 0 = Normal, 1 = Reverse, 2 = Strike slip, 3 = Thrust, 4 = Reverse oblique based on reported information of the earthquake. The earthquake magnitude is grouped into 6 classes: 0 = Minor (M 3~3.9), 1 = Light (M 4~4.9), 2 = Moderate (M 5~5.9), 3 = Strong (M 6~6.9), 4 = Major (M 7~7.9), 5 = Great (M 8 or more). The focal depth (termed D) is grouped into three classes following a hybrid criterion considering the current data and the universal accepted criteria: 0 = Very shallow (0~20 km), 1 = Shallow (20~70 km), 2 = Intermediate (70~300 km). With the above classifications, we then obtained the factors classification information of 167 ground motion recordings.

We select the epicentral distance (termed Repi) measured between the station and corresponding earthquake epicenter to represent the path effect. The epicentral distance is grouped into 5 classes: 0 = (0~100 km), 1 = (100~200 km), 2 = (200~300 km), 3 = (300~400 km), 4 = (>400 km). We choose 3 factors to represent the site effect: site class, PGA class, and liquefaction occurrence. The site class is determined using the V_{s30} data generally following the NEHRP classification: 0 = NEHRP A class (hard rock, V_{S30}>1500 m/s), 1 = NEHRP B class (rock, V_{S30}=760~1500 m/s), 2 = NEHRP C class (very dense soil and soft rock, V_{S30}=360~760 m/s), 3 = NEHRP D class (stiff soil, V_{S30}=180~360 m/s), 4 = NEHRP E class (soft soil, V_{S30}<180 m/s). The V_{S30} data of our 167 record stations are determined from the global slope-based V_{S30} database (USGS,2020) and replaced by site-specific V_{S30} data if they are available in the literature. Ground motion PGA class is used to represent the possible effects of nonlinear soil behavior on site response and ground motion records. Regnier et al. (2013) found that the nonlinear soil behaviors tend to significantly modify the site responses characterized by weak motions when the site PGA exceeds a site-specific threshold value. Here, we group the site PGA into 5 classes: 0 = Weak (PGA<100 cm/s^2), 1 = Moderate (PGA=100~200 cm/s^2), 2 = Strong (PGA=200~300 cm/s^2), 3 = Very strong (PGA=300~400 cm/s^2), 4 = Severe (PGA> cm/s^2). The occurrence of liquefaction is considered as a special site effect here. We group the liquefaction occurrence of the 167 recordings from literature into two classes: 0 = Non-liquefied, 1 = Liquefied.

Table 2. Factors used in this work to represent earthquake source, path, site effects on site-specific ground motions, and their grouping criteria.

Group	Source							
Factor (Variable, unit)	EventType		FaultType		MagnitudeClass (Mw,)		DepthClass (Depth, km)	
Criteria	0	Interface	0	Normal	0	3~3.9	0	0~20
	1	Instraslab	1	Reverse	1	4~4.9	1	20~70
	2	Shallow	2	Strike Slip	2	5~5.9	2	70~300

	crustal	3	Thrust	3	6~6.9		
		4	Reverse	4	7~7.9		
			oblique	5	⩾ 8		
Group	Path			Site			
Factor (Variable, unit)	RepiClass (Repi, km)		SiteClass (V_{S30}, m/s)		PgaClass (PGA, cm/s2)		Liquefaction
Criteria	0	0 ~ 100	0	> 1500	0	<100	0 Non-liquefied
	1	100 ~ 200	1	760 ~ 1500	1	100 ~ 200	1 Liquefied
	2	200 ~ 300	2	360 ~ 760	2	200 ~ 300	
	3	300 ~ 400	3	180 ~ 360	3	300 ~ 400	
	4	>400	4	< 180	4	> 400	

CORRELATION ANALYSES

After pre-processing of the ground motion data summarized in Table 1 following the approach described in Table 2, we obtain tabular data with 167 samples and 8 independent factors and 2 response variables (RL and MIFr). In this section, we present the correlation analysis and the comparison between these categorical independent variables and numerical response variables. We first visualize the correlations between these dependent factors and RL, MIFr using paired box plots in Figure 2 and 3, respectively. We observe that, in general, RL and MIFr calculated from liquefied records are higher than those from the non-liquefied records. Earthquake source and path factors may have a similar effect on RL and/or MIFr as liquefaction does. For non-liquefied records, the interface earthquakes (EventType=0) correspond to higher RL values and lower MIFr values compared to the intraslab and shallow crustal earthquakes. The RL value does not exhibit a clear trend with the fault type, while the events induced by strike slip fault activity (FaultType=2) tend to correspond to larger MIFr values than events induced by other types of fault activity. Increasing the earthquake magnitude tends to correlate to an increasing RL value and a decreasing MIFr value. Deep events (DepthClass=2) tend to have lower RL values and larger MIFr values compared to other shallow events. Far-field sites (increasing RepiClass) tend to have higher RL values and lower MIFr values. However, these trends are not significant due to imbalanced data among different Repi classes. Softer sites (higher SiteClass values) tend to have higher RL and MIFr values. The RL values tend to increase with ground motion intensity (increasing PgaClass) while the MIFr value does not show a clear trend with PGA level.

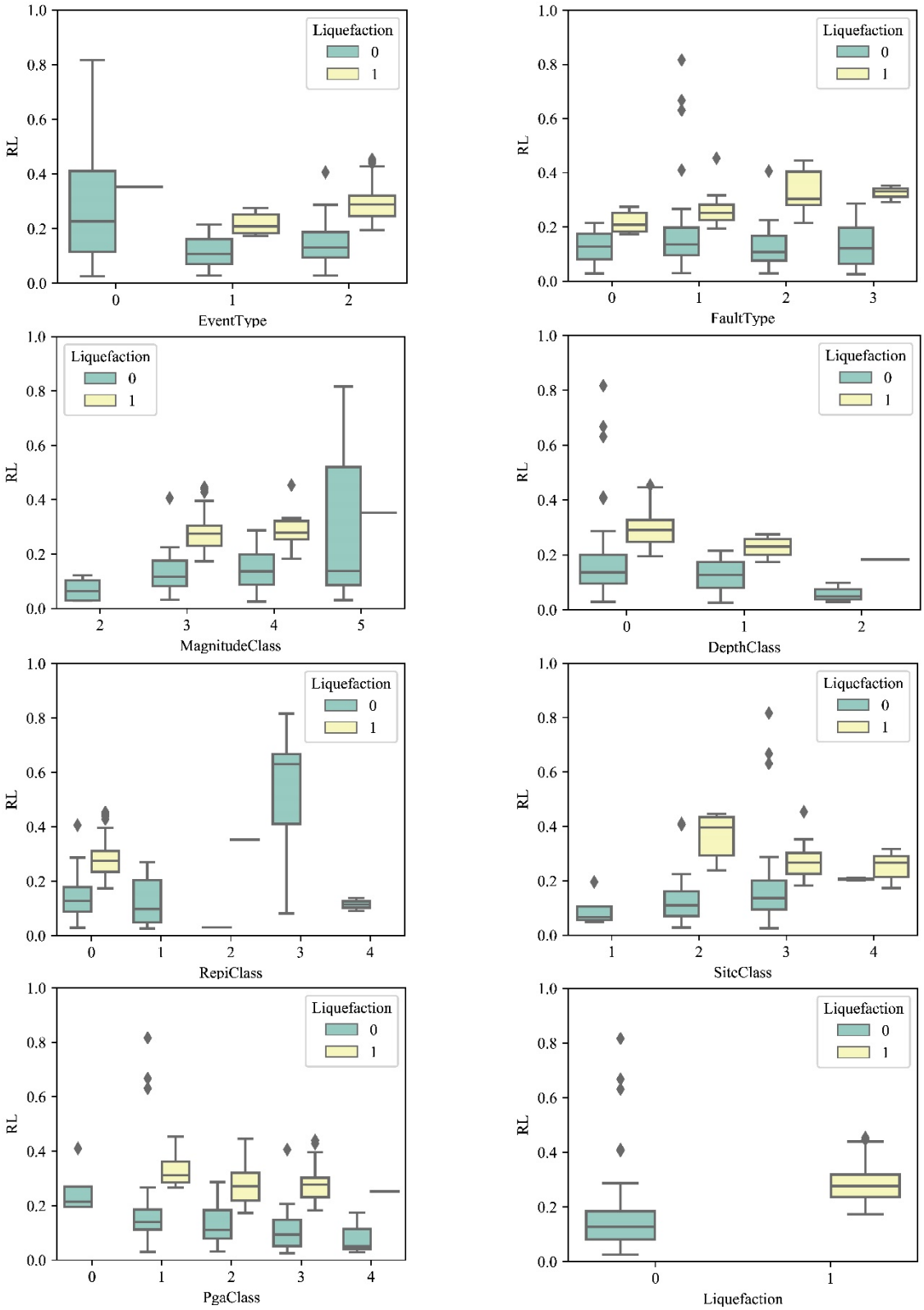

Figure 2. Box plots of RL parameter corresponding to the eight factors representing source, path, and site effects (factor classification criteria see Table. 2)

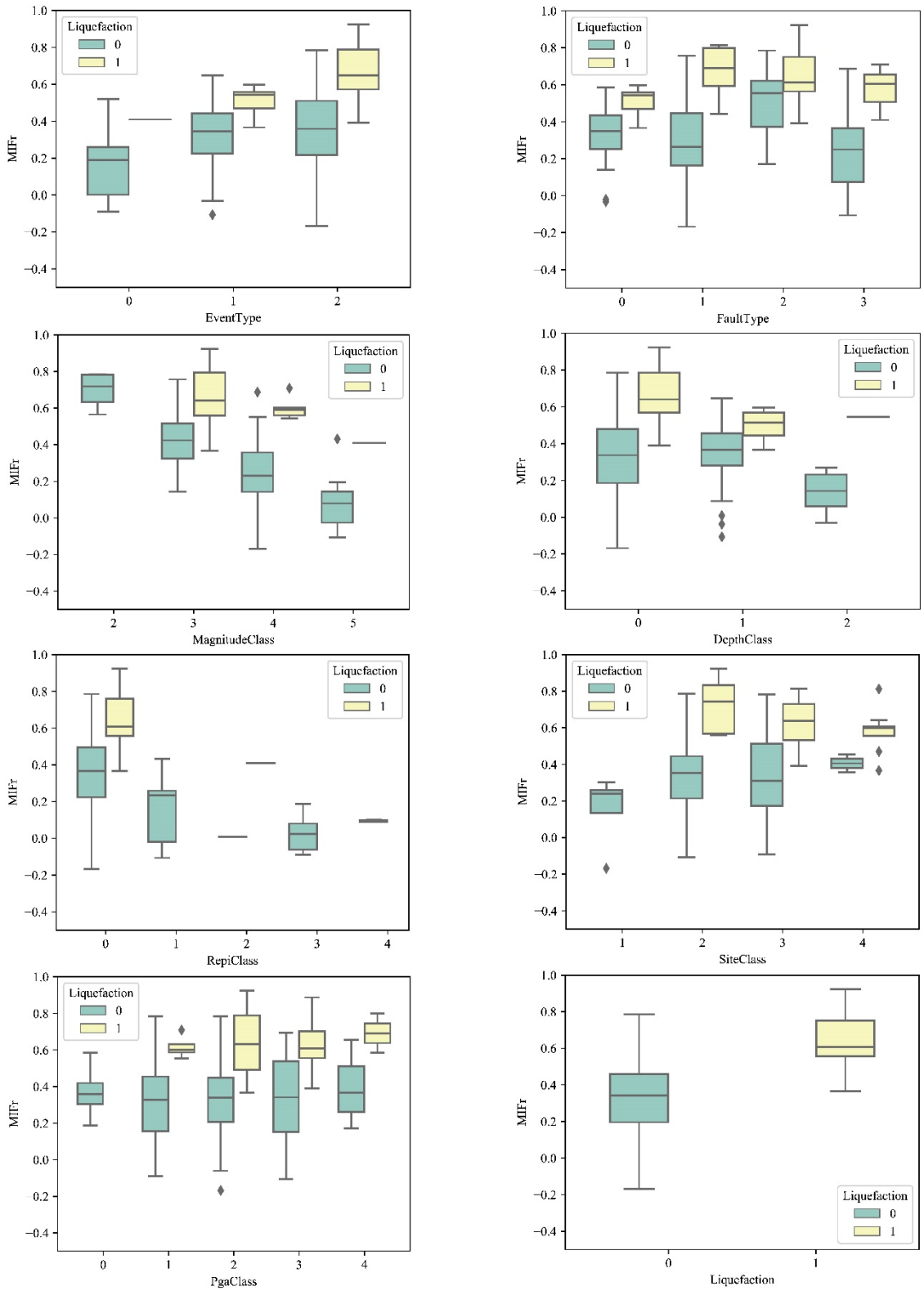

Figure 3. Box plots of MIFr parameter corresponding to the eight factors representing source, path, and site effects (factor classification criteria see Table. 2)

The above graphic analyses show a general trend of the RL and MIFr variations with different influencing factors. To quantitatively compare the correlations between RL and MIFr (two continuous variables) and the influencing factors (categorical variables), we adopt the Analysis of Variance (ANOVA) method to rank the effects of different factors on RL and MIFr separately. To ensure the comparability of different types of factors, we convert some continuous factors (i.e., earthquake magnitude, focal depth, epicentral distance, PGA) to categorical factors as detailed previously. The ANOVA method is a method of hypothesis testing and can be used to analyze the variance among different classes of categorical variables. Here, we use the one-way ANOVA that quantifies the correlation between one categorical independent factor and one numerical dependent variable at a time to analyze all pairs of the 8 factors and 2 dependent variables (i.e., 16 pairs in total). The null hypothesis is that means of population of all classes are equal, and the alternative hypothesis is that not all of the means are equal. We then calculate the variances due to different sources: between samples and within samples, and conduct the F test to assess whether a significant difference exists in the ground motion frequency parameters across the different classes of a categorical independent factor (Ott and Longnecker, 2015). According to Ott and Longnecker (2015), the j^{th} sample observation from the population (i.e., class) i can be expressed as the sum of three terms:

$$y_{ij} = \mu + \alpha_i + \xi_{ij}$$

where μ is the overall mean that is an unknown constant, a_i is the effect due to population i and is an unknown constant, ξ_{ij} is a random error associated with the j^{th} observation from population i. Usually, ξ_{ij} is independent and normally distributed with a mean of 0 and the variance of σ^2 (Ott and Longnecker, 2015).

The variance of the data is expressed as three types of sum of squares (SS) from different sources: the total sum of squares (*TSS*), the sum of squares between samples (*SSB*), and the sum of squares within samples (*SSW*), and their calculations are following:

$$TSS = \sum_{ij} y_{ij}^2 - \frac{Y^2}{n_T}$$

$$SSB = \sum_{i}^{t} \frac{Y_i^2}{n_i} - \frac{Y^2}{n_T}$$

$$SSW = TSS - SSB$$

where n_T is the total sample size, t is the number of population, n_i is the sample size selected from population i, Y is the sum of all sample observations, and Y_i is the sum of sample measurements from population i. The F distribution has two degrees of freedom: $df_1 = t - 1$ and $df_2 = n_T - t$, and the F value is calculated as:

$$F_{obs} = \frac{SSB / (t-1)}{SSW / (n_T - t)}$$

The p value decided by the F distribution and F_{obs} needs to be smaller than 0.05 (i.e. 95% significance level) to reject the null hypothesis and to conclude that the categorical variable has a significant effect on the dependent variable. We can get the importance ranking by comparing the p values of different categorical variables. Using this strategy, we calculate and rank the correlation (in terms of p value) between the eight factors and the two frequency parameters in Table 3. We find the liquefaction factor ranks the second strongest correlation with both RL and

MIFr, and liquefaction has *p* values that are several orders smaller than those of the other factors. Hence, the assumption that liquefaction effects on the two strong motion frequency parameters are stronger than a common source, path, and site effects on site-specific ground motions is still valid. Special modification or restriction may be applied in the accelerogram-based liquefaction assessment method to limit the confusing effects of epicentral distance (rank No.1 in RL correlations) and earthquake magnitude (rank No.1 in MIFr correlations).

Table 3. The ANOVA test results for the pairs between the eight factors and two ground motion frequency parameters and their correlation ranking

Group	Factors	RL			MIFr		
		F	p	Rank	F	p	Rank
Source	EventType	14.753	1.27E-06	3	10.191	6.69E-05	5
	FaultType	2.188	9.14E-02	n.a.	10.891	1.45E-06	4
	MagnitudeClass	6.171	5.30E-04	4	29.516	2.49E-15	1
	DepthClass	7.636	6.72E-04	5	4.057	1.90E-02	6
Path	RepiClass	15.595	7.82E-11	1	10.074	2.58E-07	3
Site	SiteClass	3.059	2.98E-02	6	3.829	1.10E-02	5
	PgaClass	1.924	1.09E-01	n.a.	1.214	3.07E-01	n.a.
	Liquefaction	46.942	1.34E-10	2	68.305	4.15E-14	2

Note: n.a. denotes the effect is not significant, and therefore, is excluded from ranking.

CONCLUSION

In this work, we selected eight factors representing the earthquake source, path, and the site effects for the 167 ground motion records used in a recently developed accelerogram-based liquefaction assessment method. We analyze the correlations between the two ground motion frequency parameters, RL and MIFr, and the eight factors using both paired box plots and the one-way ANOVA method. The results reveal that, in general, liquefaction has a more significant effect on the site-specific ground motion frequency parameters compared to other factors such as earthquake source, path, and site effects, which is the basis of the accelerogram-based liquefaction assessment methods. However, earthquake source and path factors may have similar effect on RL and/or MIFr as liquefaction does, and they need to be accounted for in the assessment method to improve the accuracy and effectiveness of the method. Potential limitations of the current work are that a limited number of ground motion records are used, and that the one-way ANOVA analysis does not consider the interaction between different factors. For future work, we will continue our efforts to expand the database of ground motion records with the corresponding liquefaction observations and to consider additional factors to study the effect of the earthquake source, path, and site effects on the ground motion frequency characteristics. Moreover, advanced data analysis methods will be explored to quantify the correlations between various factors and frequency characteristics.

REFERENCES

Bonilla, L. F., Archuleta, R. J., and Lavallée, D. (2005). "Hysteretic and dilatant behavior of cohesionless soils and their effects on nonlinear site response: Field data observations and modeling." *Bulletin of the Seismological Society of America*, 95(6), 2373–2395.

Bozorgnia, Y., Stewart, J., and Abrahamson, N. (2020). "Data resources for NGA subduction project." *Pacific Earthquake Engineering Research Center*, USGS.
Gingery, J. R., Elgamal, A., and Bray, J. D. (2015). "Response spectra at liquefaction sites during shallow crustal earthquakes." *Earthquake Spectra*, 31(4), 2325–2349.
Kostadinov, M. V. and Yamazaki, F. (2001). "Detection of soil liquefaction from strong motion records." *Earthquake Engineering and Structural Dynamics*, 30(2), 173–193.
Kramer, S. L., Asl, B. A., Ozener, P., and Sideras, S. S. (2015). "Effects of liquefaction on ground surface motions." *Perspectives on Earthquake Geotechnical Engineering*, Springer, 285–309.
Ott, R. L. and Longnecker, M. T. (2015). *An introduction to statistical methods and data analysis*. Nelson Education.
Régnier, J., Cadet, H., Bonilla, L. F., Bertrand, E., and Semblat, J.-F. (2013). "Assessing nonlinear behavior of soils in seismic site response: Statistical analysis on kik-net strong-motion data." *Bulletin of the Seismological Society of America*, 103(3), 1750–1770.
Shimizu, Y., Watanabe, A., Koganemaru, K., Nakayama, W., and Yamazaki, F. (2000). "Super high-density realtime disaster mitigation system." *Proceedings of the 12th World Conference on Earthquake Engineering*, number 2345, New Zealand Society for Earthquake Engineering, Auckland, New Zealand (30 January-4 February).
USGS (2020). "Vs30 Models and Data" USGS website <https://www.usgs.gov> (Oct. 2, 2020).
Van Houtte, C., Bannister, S., Holden, C., Bourguignon, S., and McVerry, G. (2017). "The new zealand strong motion database." *Bulletin of the New Zealand Society for Earthquake Engineering*, 50(1), 1–20.
Youd, T. L. and Carter, B. L. (2005). "Influence of soil softening and liquefaction on spectral acceleration." *Journal of Geotechnical and Geoenvironmental Engineering*, 131(7), 811–825.

ACKNOWLEDGMENTS

The authors would like to acknowledge Prof. Qiong Zhang for the discussions and suggestions regarding the statistical analysis presented in this work.