Data verification for ambient vibration tests

Valeri V. Lenchine

Environmental Management, GHD Pty Ltd, L9, 180 Lonsdale Str., Melbourne VIC 3000, Australia

ABSTRACT
Ambient vibration tests for large structures have become more popular as a tool to identify fundamental or lowest natural frequencies. Ambient vibration analysis is pertained to "response only" methods where the characteristics of excitation are not known. Different techniques are utilised to post-process data. The quality of the ambient monitoring data used for subsequent post-process is critical for further analysis, independently of the method used for structural dynamics analysis. A data pre-processing procedure was suggested for the rectification of data collected during ambient vibration tests. It is suggested to make the decision on the validity of data acquired during particular time periods based on the comparison of descriptive statistics such as skewness and kurtosis with reference magnitudes. A high deviation of the skewness and kurtosis from the reference magnitudes can be used as an indication that the data may be invalid. An example in the paper utilised the Jarque-Bera test for normal distribution to rectify data for subsequent fundamental frequency identification. The method was successfully implemented for the analysis of a large amount of data collected during an ambient vibration test of a dam structure.

1 INTRODUCTION

Several reasons, such as the complexity of forced excitation and the potential unreliability of structures after forced excitation tests contribute towards the wider use of ambient vibration data acquisition.

Ambient vibration analysis is pertained to "response only" methods where characteristics of excitation are not known. Different techniques are utilised to post- process data. Some of the methods that are frequently used for identification of possible natural frequencies are limited to the analysis of amplitude information acquired during a random excitation of a structure (Hattingh et al., 2019; Zhang et al., 2013). These methods include (Brinker et al.,2000; Calcina et al., 2014; Castro et al., 1998; Gul and Catbas, 2008; Ditomasso et al., 2013; Fat-Helbary et al., 2019):

- Site – to - reference spectral ratio
- Extension of Horizontal - to - Vertical Spectral Ratio (HSVR) method for structural dynamics (also known as Nakamura method); and
- Analysis of local maximums in acceleration spectrums acquired at the monitoring points.

If vibration data was acquired by a synchronous data acquisition at multiple points of a structure, other techniques that also use phase information can be utilised. Operating modal analysis (OMA) may provide more information regarding possible modal characteristics of a structure.

The quality of the ambient monitoring data used for consequent post-process is important for further analysis, independently of the method used for structural dynamics analysis.

2 DATA PRE- PROCESSING

Depending on the data acquisition system, the initial analysis of data collected during ambient vibration monitoring typically involves few steps. Many seismic and structural vibration monitoring data acquisition instruments introduce trends or bias in the vibration time traces. Elimination of these trends by different detrending techniques or high pass filtering of the data is typically the first step in data analysis (Srbulov, 2010). Band-pass or low-pass filtering is also used to concentrate further analysis on the frequency span of interest.

These pre-processing steps still do not answer the question of whether particular blocks of data can be used for extracting useful information about dynamics of a structure. Suitability of data for consequent analysis is addressed from qualitative rather than from quantitative perspective (Wenzel and Pichler, 2005). The assumptions underlying ambient vibration analysis include the measured signal not being affected by anthropogenic or other extraneous vibration sources and the spectral characteristic of ambient excitation being relatively flat within the frequency span of interest. Initial data processing does not indicate which of the data meet these requirements and can be taken as a basis for consequent structural dynamics analysis. Simple thresholds of magnitude of the signal or the presence of multiple peaks in the acquired data may not always allow for accurate identification of valid ambient vibration records. This is especially true if extraneous vibration is mixed with ambient excitation and does not manifest itself as peaks in the signal or does not lead to noticeable change in the temporal or spectral structure of the measured signal.

2.1 Assumptions about valid ambient vibration data

A measured vibration microtremor generally represents a random process. Ambient vibration acquired on large structures such as bridges, dams or tall buildings results from different kinds of excitation, such as wind, ground transmitted geo processes and excitation from a number of anthropogenic or natural sources. It is difficult to derive correct conclusions about the validity of collected data based on the level of excitation only. For example, higher wind load can result in a higher level of vibration response that is greater than during other periods, but it can still be classified as valid data. Not all of the data collection periods can be classified as quasi-stationary. Most of the relatively long-term monitoring data will represent a mix of stationary and transient responses of a structure under random excitation. Therefore, a direct comparison of the magnitudes of signal during different monitoring periods or spectral magnitudes and frequency content may not necessarily provide a good basis for the rectification of data. Greater vibration may correspond to periods with more significant natural excitation unless it is known that the augmented response is affected by particular vibration sources.

It is suggested to address the problem of the validity of data from a statistical perspective. The statistical descriptors used should have features:

- Be unsensitive to scaling of the signal (assuming that a greater or lower excitation may cause a lower or greater response which is proportional to the level of excitation)
- Does not depend on shifts of transient processes in a signal (since they may happen randomly with respect to the beginning or end of data acquisition).

Simple statistical parameters like the mean or standard deviation are not deemed to be the right candidates. Furthermore, the correlation or coherence between different blocks of data acquired under random excitation may not be appropriate statistical metrics.

Statistics based on higher order moments may be more appropriate for drawing conclusions about the similarity of data. The skewness of data is based on the third order statistical momentum:

$$S = \frac{E(x-\mu)^3}{\sigma^3}, \tag{1}$$

where E is the expected value of the quantity, μ is the mean and σ is the standard deviation. For an array of discrete data, the skewness can also be presented as follows:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^3}, \tag{2}$$

where $\bar{x}$ is the mean value and n is the number of sample points. If the skew is negative, the left tail of the probability distribution is longer and the mass of the distribution is concentrated on the right. The reverse is true for positive skewness.

Another statistical metric that is based on the fourth order statistical momentum is called kurtosis:

$$\beta = \frac{E(x-\mu)^4}{\sigma^4}. \tag{3}$$

This can be represented for a finite array of data:

$$\beta = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^2}. \tag{4}$$

It should be noted that slightly different definitions are sometimes used for skewness and kurtosis in the literature. The formulas given above correspond to ISO 3534-1: 2006.

The skewness considers the extremes of the data set rather than focusing solely on the average and indicates if the probability distribution for the data set is different from the normal distribution. ISO 20816 considers this as an indicator of asymmetry of the vibration signal. The interpretation of kurtosis includes the number of transients in the vibration signal or the characterisation of the "number of pulses" in the signal; comparison of the signal with white noise or other types of signal with known probability density distributions. This statistical descriptor is also linked to the effective duration of impulses in a signal.

2.2 Skewness and kurtosis for ambient vibration measurements

If the vibration signal is normalised to the zero mean, then the formulas in the previous section can be simplified. Estimates of skewness and kurtosis for a vibration signal *v(n)* for sufficiently high number of sample points can be defined as follows:

$$S(n) \approx \frac{\sqrt{n}\sum_{i=1}^{n} {x_i}^3}{\left(\sqrt{\sum_{i=1}^{n} {x_i}^2}\right)^3}, \tag{5}$$

$$\beta(n) \approx \frac{n\sum_{i=1}^{n} {x_i}^4}{\left(\sum_{i=1}^{n} {x_i}^2\right)^2}. \tag{6}$$

Measured ambient vibration is not necessarily statistically stationary. The independent block of samples may contain a combination of random quasi stationary and transient processes. However, kurtosis may still be useful in characterising such signals. For example, it was shown in work (Müller et al., 2020) that the coefficient of kurtosis converges for a sequence of pulses or transient processes to a greater magnitude. Therefore greater than expected magnitudes of kurtosis may indicate presence of significant number of impulsive or intermittent events in a data record and hence unsuitability for ambient vibration analysis.

3 DATA FROM AMBIENT VIBRATION TESTS

Data acquisition during ambient vibration tests typically covers a long period of time, following which the data is analysed by blocks over different periods of time (Giraldo et al., 2009; Abdel-Ghaffar and Scanlan, 1985; Guler et al., 2008). Substantial differences in the spectral content or vibration magnitudes are sometimes used to segregate vibration time histories in "normal" periods or periods where the data may be affected by anthropogenic or other non-typical vibration sources. However, a simple increase of vibration levels may be caused by increased natural excitation. It is suggested to analyse the validity of data blocks based on the similarity of their statistical characteristics such as skewness and kurtosis. It can be interpreted as if a random signal in different data blocks can contain scaled or shifted transient processes but they are still expected to be similar from a statistical perspective. The segregation of data blocks into valid and not valid periods should be done for microtremors normalised to the zero mean. Pre-processing techniques like detrending or filtering may also be required depending on the expected outcomes of consequent data analysis. It is not expected that the time series of a random vibration will have identical statistical characteristics. Certain deviation from reference kurtosis and skewness is to be used as an indicator of validity of the vibration data set.

Examples in Figure 1 and Figure 2 show skewness and kurtosis based on 10 min blocks of data. The microtremor was measured on a dam crest in vertical and horizontal plains. The data was collected over a period of more than 50 hours, and the time in the figures corresponds to the start of the measurement period. Data collected during different periods show substantial differences in the magnitudes and major spectral components. It is difficult to tell if some of the data blocks are affected by extraneous noise sources and point out reference periods of ambient vibration.

One can see that skewness of the data fluctuates around zero magnitude. It is known that skewness of a normal distribution or symmetrical dataset equals to zero. It is difficult to expect that ambient vibration testing will result in a data set that is characterised by an exact normal distribution and zero skewness. However, it is deemed to be that an almost Gaussian distribution corresponds to periods with relatively steady vibration signals and the absence of substantial pulses during the acquisition period. Skewness and kurtosis for these periods is close to the magnitudes of a normal distribution (the kurtosis of a normal distribution is 3).

One can see that some of the magnitudes in Figure 1 and Figure 2 are significantly different from the parameters expected for a normal distribution. This raises concerns about other excitation sources that may affect microtremor measurements. High deviations for skewness and kurtosis from the values of a normal distribution are not always synchronised. However, the greatest magnitudes of the skewness are detected at the same time for vertical and horizontal directions at approximately 36.8 hours.

3.1 Deviations of skewness and kurtosis

It is suggested to use a high deviation of the skewness and kurtosis from the "normal" magnitudes as an indication that the data may not be valid for consequent analysis. There is no general agreement on the acceptable deviation of the parameters for a distribution to be considered approximately normal. Some works suggest comparing skewness and kurtosis with twice of the standard errors for these parameters (West et al., 1995; Brown, 1997). The standard errors for these distribution parameters are dependent on the number of data points analysed and tend to be zero for a large number of samples (Tabachnick and Fidell, 1996), which makes this approach impracticable in many cases. Absolute deviations are suggested for significant arrays of data in some works. Absolute skewness greater than 2 and kurtosis greater than 7 can be considered as indicators of substantial non-normality of the data distribution (Kim,2013). It should be noted that these recommendations are different from some 'rules of thumb' that are generally used in data analysis. They assume that absolute skewness greater than 1 is substantial, while absolute skewness below 0.5 is considered small. Neither of these approaches is backed up by convincing evidence. Other recommendations involve the separate comparison of skewness and kurtosis with a threshold that may be chosen based on experience or other rationales.

Since both of these statistical parameters are considered together in this paper, the Jarque-Bera test is suggested for choosing valid data blocks in this case. This test is based on the calculation of the data array statistics (Jarque and Bera, 1987):

$$JB = \frac{n}{6}\left(S^2 + \frac{(\beta-3)^2}{4}\right), \tag{7}$$

where S is the sample skewness and β is the sample kurtosis. Following this test, the magnitude of the Chi-Square distribution with 2 degrees of freedom corresponding to the computed JB is compared with the specified level of significance (typically p- value of 0.05 is accepted) to decide on whether the parameters of the distribution correspond to a normal distribution.

The available data set of dam vibration measured at the crest of the dam was analysed using typical procedures, where data blocks with greater root mean square magnitudes and significant peaks were not used for subsequent analysis. The same set of data was rectified using the procedure described above. The implementation of the suggested method resulted in a significantly lower amount of valid data blocks in comparison with the typical procedure of data rectification, i.e., the number of valid blocks was 24 out of 301 using the skewness and kurtosis analysis. On the other hand, segregation of the data based on the typical post-process analysis resulted in about 50% of valid data blocks.

The shortcoming of too small a percentage of valid data can be alleviated by exploring data with shorter periods, for example 2 minute data blocks can be explored instead of 10 minute blocks that may potentially increase the number of valid data acquisition periods. Horizontal to vertical spectral ratios (HVSR) were used for the analysis of the fundamental frequency of the dam as an expansion of the Nakamura method (Nakamura 1989; Nakamura 2008) to structural dynamics (Srbulov 2010; Ditomasso et al. 2013; Mucciarelli 2004). An example of the calculated average HVSR for both of the rectified data sets is shown in Figure 3, where HVSR1 was computed using the technique suggested in this paper, while HVSR2 resulted from the typical data rectification process. There is an apparent similarity in both of the curves. However, the implementation of a similar technique for other monitoring points revealed that the fundamental frequency for the dam was better identified using data pre-processed using the suggested method. The estimate of the fundamental frequency is more consistent with the magnitudes computed by other methods and is supposed to be more accurate than the estimate of the fundamental frequency obtained using the typical data rectification procedure.

It should be noted that the suggested procedure based on deviation from skewness and kurtosis of normal distribution is not necessarily universal. In general, collected data may not tend to fit a normal distribution. In this case, other acceptable deviation should be explored based on similarity to other kinds of statistical distributions or distributions corresponding to known reference periods.

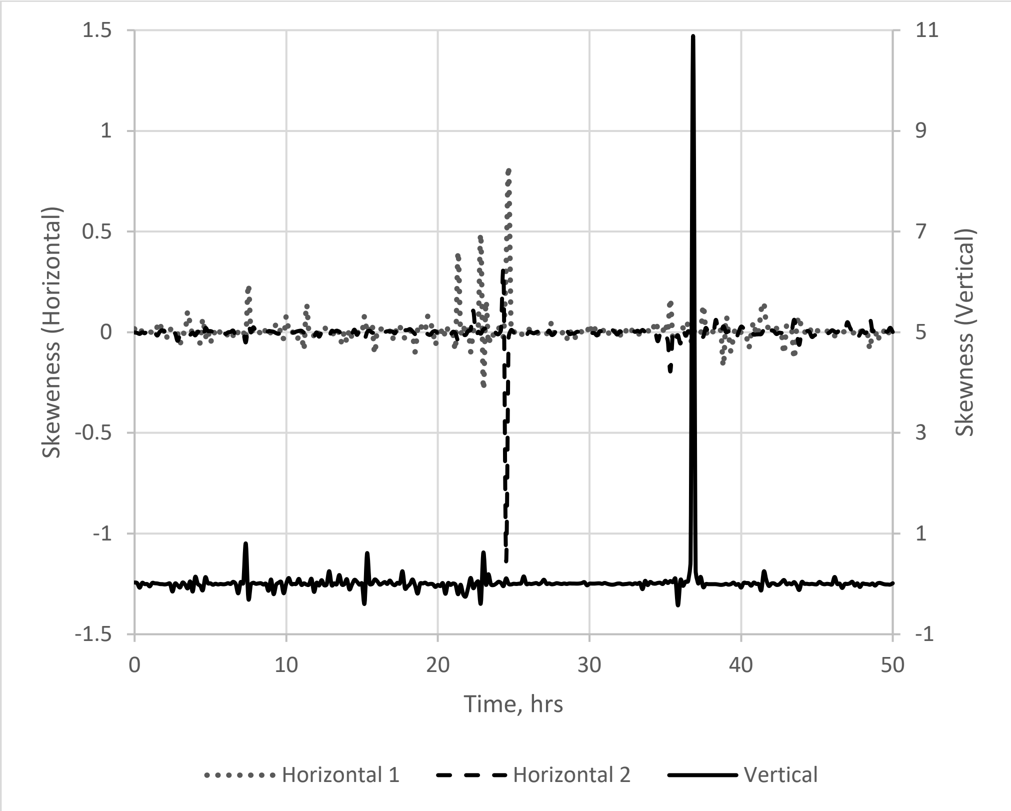

Figure 1: Skewness of vibro- acceleration data

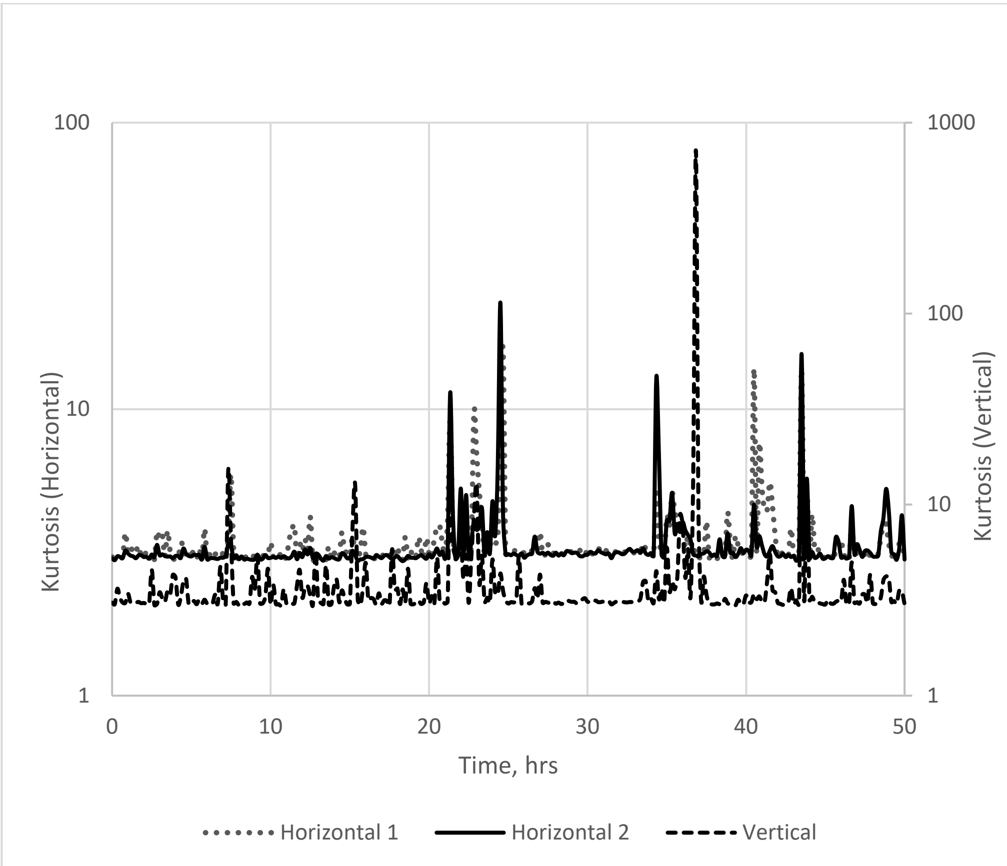

Figure 2: Kurtosis of vibro- acceleration

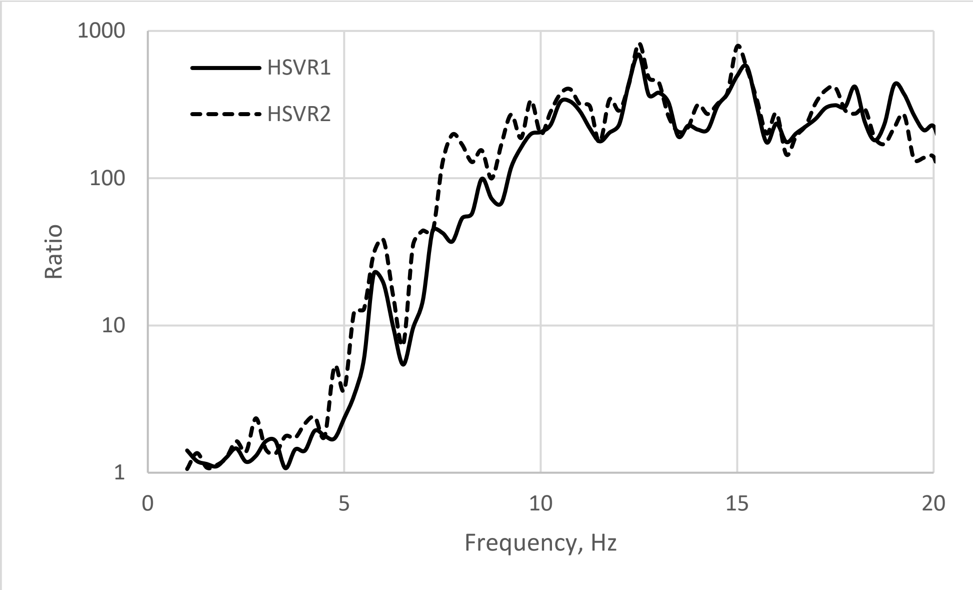

Figure 3: Horizontal to vertical spectral ratio for the dam crest

4 SUMMARY

A data pre-processing procedure was suggested for the rectification of data collected during ambient vibration tests. It is suggested to make the decision on the validity of data acquired during particular time periods based on the comparison of descriptive statistics such as skewness and kurtosis with reference magnitudes. A high deviation of the skewness and kurtosis from the reference magnitudes can be used as an indication that the data may not be valid for consequent analysis. In the considered example, the Jarque-Bera test for normal distribution was used to rectify data for consequent fundamental frequency analysis. The method was successfully implemented for the processing of a large amount of data collected during an ambient vibration test of a dam structure. The results of consequent fundamental frequency analysis were proven to be more accurate based on a comparison with the fundamental frequency estimates obtained by other methods.

REFERENCES

Abdel-Ghaffar, A.M., Scanlan, R.H. (1985). “Ambient vibration studies of Golden Gate bridge: I. Suspended structure”, *Journal of Engineering Mechanics,* vol. 111(4), pp. 463–482. https://doi.org/10.1061/(ASCE)0733-9399(1985)111:4(463)

Brincker, R., Zhang, L., Andersen, P. (2000). “Modal identification from ambient responses using frequency domain decomposition”, *Proceedings of 18th International Modal Analysis Conference (IMAC XVIII)*, San Antonio, USA.

Brown, J.D. (1997). “Skewness and kurtosis”, *JALT Testing & Evaluation SIG Newsletter,* vol. 1(1), pp. 20-23.

Giraldo, D.F., Song, W., Shirley J. Dyke, S.J., and Caicedo, J.M. (2009). “ Modal identification through ambient abration: comparative study”, *Journal of Engineering Mechanics,* vol. 135(8), pp. 759-770. https://doi.org/10.1061/(ASCE)0733-9399(2009)135:8(759)

Calcina, S.V., Eltrudis, L., Piroddi, L. and Ranieri G. (2014). “Ambient vibration tests of an arch dam with different reservoir water levels: experimental results and comparison with finite element modelling.” *The Scientific World Journal,* vol.2014, pp. 1-12. https://doi.org/10.1155/2014/692709

Castro, R.R., Mucciarelli, M., Pacor, F., Frederici, P.P., and Zaninetti A. (1998). “Determination of the characteristic frequency of two dams located in the region of Calabria, Italy”, *Bulletin of the Seismological society of America,* vol. 88(2), pp. 503-511.

Gul, M., Catbas, F.N. (2008). “Ambient vibration data analysis for structural identification and global condition assessment”, *Journal of Engineering Mechanics,* vol. 134(8), pp. 650-662. https://doi.org/10.1061/(ASCE)0733-9399(2008)134:8(650)

Ditommaso R, Vona M, M. R. Gallipoli MR, Mucciarelli M. (2013). “Evaluation and considerations about fundamental periods of damaged reinforced concrete buildings”, *Natural Hazards and Earth System Sciences*, vol. 13(7), pp. 1903–1912. http://dx.doi.org/10.5194/nhess-13-1903-2013

Fat-Helbary, R.E., El-Faragawy, K.O., Hamed A. (2019). “Application of HVSR technique in the site effects estimation at the south of Marsa Alam city, Egypt*”, Journal of African Earth Sciences,* vol. 154, pp. 89-100. https://doi.org/10.1016/j.jafrearsci.2019.03.015

Guler K., Yuksel, E., Kocak, A. (2008). “Estimation of the fundamental vibration period of existing RC buildings in Turkey utilizing ambient vibration records”, *Journal of Earthquake Engineering,* vol. 12(S2), pp. 140–150. https://doi.org/10.1080/13632460802013909

Hattingh, L., Moyo, P., Shaanika, S., Mutede, M., le Roux, B,, Muir C. (2019). “The use of Ambient Vibration Monitoring in the behavioral assessment of an arch dam with gravity flanks and limited surveillance records”, *Proceedings of ICOLD Symposium Sustainable and Safe Dams around the World*, Ottawa, Canada.

ISO 20816-1:2016: *Mechanical vibration — Measurement and evaluation of machine vibration — Part 1: General guidelines,* International Organization for Standardization, Geneva, Switzerland.

ISO 3534-1:2006: *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*, International Organization for Standardization, Geneva, Switzerland.

Jarque, C.M., Bera, A.K. (1987). “A test for normality of observations and regression residuals”, *International Statistical Review,* vol. 55(2), pp. 163–172. https://doi.org/10.2307/1403192

Kim, H.Y. (2013). “Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis”, *Restorative Dentistry & Endodontics,* vol.38(1), pp. 52-54. https://doi.org/10.5395/rde.2013.38.1.52

Mucciarelli, M., Masi, A., Gallipoli, M.R., Harabaglia, P., Vona, M., Ponzo, F., Dolce, M. (2004). “Analysis of RC building dynamic response and soil- building resonance based on data recorded during a damaging earthquake (Molise, Italy,2002)”, *Bulletin of the Seismological Society of America*, vol. 94(5), pp. 1943-1953. https://doi.org/10.1785/012003186

Müller, R.A.J., von Benda-Beckmann, A.M., Halvorsen, M.B., and Ainslie, M.A. (2020). “Application of kurtosis to underwater sound”, *Journal of Acoustical Society of America*, vol. 148 (2), pp. 780-792. https://doi.org/10.1121/10.0001631

Nakamura, Y. (1989). *A method for dynamic characteristic estimation of subsurface using microtremor on the ground surface,* Quarterly Report, Railway Technical Research Institute. vol. 30(1), pp. 25-33.

Nakamura, Y. (2008). “On the H/V spectrum”, *Proceedings of the 14th World Conference on Earthquake engineering,* Beijing, China.

Srbulov, M. (2010). *Ground vibration engineering. Simplified analyses with case studies and examples*. Springer, London.

Tabachnick, B.G., Fidell, L.S. (2019). *Using multivariate statistics,* 7th ed. Pearson Education UK, Harlow.

Wenzel, H.,Pichler D. (2005). *Ambient vibration monitoring*, John Wiley & Sons, Chichester.

West, S.G., Finch, J.F., and Curran P.J. (1995). ‘Structural equation models with nonnormal variables: problems and remedies.” In RH Hoyle, R.H. (Ed.). *Structural equation modeling: Concepts, issues and applications*: 56-75. SAGE Publications.

Zhang, J., Prader, J., Grimmelsman, K.A., Moon, F. (2013). “Experimental vibration analysis for structural identification of a long-span suspension bridge”, *Journal of Engineering Mechanics*, vol. 139(6), pp. 748-759. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000416