Proceedings of the 9th U.S. National and 10th Canadian Conference on Earthquake Engineering
Compte Rendu de la 9ième Conférence Nationale Américaine et
10ième Conférence Canadienne de Génie Parasismique
July 25-29, 2010, Toronto, Ontario, Canada • Paper No 86

IDENTIFICATION OF DYNAMIC PROPERTIES OF DAMS USING SYSTEM IDENTIFICATION TECHNIQUES AND REAL EARTHQUAKE EXCITATIONS

A.Bakhshi[1], A.Meshkat[2]

ABSTRACT

Investigation of the dynamic characteristics of an existing structure system based on field tests is of practical significance, since it can be employed in the field of examining construction quality, validating or improving analytical finite element structural models, or conducting damage assessment. The main purpose of this study is to investigate the dynamic characteristics of dams using the seismic response data. Whereas, in the process of frequency domain identification of dam properties, frequency response function (FRF) plays the most important role, evaluating the corresponding parameters with better accuracy has always been desirable. Therefore, a comparison was made between some modified methods of calculating the FRF to reduce its variance. Furthermore, a software is introduced which is able to use all of these methods plus an Auto-Regressive model in order to calculate Power Spectral Density (PSD) of input and output data separately so that, modal parameters of the system are identified. In order to validating the results of this AR model, the software is employed to identify dynamic characteristics of Factor building in UCLA. Later, it is also exploited to identify the dynamic properties of Shahid Rajaee concrete arch dam (which is located in the north of Iran) and compared with the prior studies.

Introduction

The problem of system identification has become important in structural engineering, particularly in connection with the prediction of structural responses to adverse environmental loads, such as earthquakes, wind and wave forces, and also with respect to the estimation of existing conditions of structures for the assessment of damage and deterioration. [1] Investigation of the dynamic characteristics of an existing structure system based on field tests is a necessary and important task in the course of checking the construction quality, validating or improving analytical finite element structural models, or conducting damage assessment. To accomplish this task, the well-known field tests used are ambient vibration tests, forced vibration tests, free vibration tests, and earthquake response measurement.

Among these field tests, ambient vibration experiments are the most popular ones because they are portable and easy to set-up. However, because the input in an ambient vibration test is usually too complicated to be known or measured, one has to determine the modal parameters from the output data only [2, 3]. In contrast for the case of forced vibrations different methods are applicable because the input data could manage easily, but strong enough excitation which could excite all mode shapes will be expensive and hazardous. Structures subjected to seismic excitations can exhibit non-linear behaviours that cannot be accurately represented using linear, time-invariant models. For this reason, non-linear

[1] Department of Civil Engineering Sharif University of Technology, Iran e-mail: bakhshi@sharif.edu
[2] Department of Civil Engineering Sharif University of Technology, Iran e-mail: meshkat@alum.sharif.edu

modelling and system identification are critical to the design, health monitoring, and damage detection of such structures, although input data could be measured in this cases but still methods are limited [4, 5].

Modelling and identification techniques can be categorized as being either parametric or non-parametric. Parametric system identification seeks to determine the "optimal" parameters for an assumed structural model such that modelled response closely matches the recorded response of the structure. Non-parametric techniques attempt to identify the "optimal" functional representation of the structure without any a priori assumptions about the model's structure. However, most of these techniques inherently involve complicated search processes and are thus computationally inefficient. They may even be numerically unstable for large-scale infrastructures that have a significant number of degrees of freedom.

Due to the large scale and distributed nature of civil infrastructures and buildings, almost all the local and global identification techniques involve the use of the measured structural responses under dynamic excitations, such as earthquake, wind, or traffic loads, also material deterioration and damage result in change in structural parameters, for example, the stiffness of a structural member or a substructure and the damping coefficient of a structure [6-8]. These changes will modify the dynamic properties, such as the natural frequencies and mode shapes of a structure [9, 10].

State of the Problem

Frequency response functions (FRFs) are the most fundamental data for the frequency domain identifications of dynamic characteristics of structural systems. Accurate evaluations of FRFs are essential to many problems, such as experimental modal analysis, structural system identification, and input identification. In the case which the system encounters with noises on input and output data the system behaviour could be considered with equation (1):

$$y(t) = \sum_{k=-\infty}^{+\infty} h(k) \times u(t-k) + v(t) \tag{1}$$

In which *v(t)* is a white noise with zero average so to find *h(k)* we need to remove *v(t),* by multiplying *u(t-τ)* to both sides and calculating the Expected value we have:

$$r_{yu}(\tau) = \sum_{k=-\infty}^{+\infty} h(k) \times r_{uu}(\tau) + E\{v(t)\} \times E\{u(t-\tau)\} \tag{2}$$

r_{yu} and r_{uu} are stand for cross-correlation and auto-correlation respectively, and because the expected values of a zero mean signal, like v is zero the equation (2) leads to:

$$r_{yu}(\tau) = r_{uu}(\tau) * h(\tau) \tag{3}$$

Or in the other way by operating Fourier transform we have:

$$\varphi_{yu}(\omega) = \varphi_{uu}(\omega).H(\omega) \tag{4}$$

φ_{yu} and φ_{uu} are two well known cross spectrum and power spectral density, from the recent equation (4) the problem of finding $H(\omega)$ (FRF) seems to be solved, but from the definition of expected value:

$$E(x) = lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i \tag{5}$$

It's obvious that all the above equations could be accepted when the length of signal is infinite, so in the case of finite signals the exact equations will be estimation of real values. If the estimation is made to calculating correlation functions and then transformed using Fourier

the estimator will called correlogram but if the spectrums estimated separately we have a periodogram. The bias and variance of an estimator are two measures often used to characterize its performance. A primary motivation is that the total squared error of the estimate is the sum of the bias squared and the variance.

$$MSE = var\{\hat{\theta}\} + |bias\{\hat{\theta}\}|^2 \tag{6}$$

Both correlogram and periodogram are biased estimators, and as it's obvious from the above equation (6) it's impossible to reduce bias and variance of such estimators simultaneously, so from the study of their bias value we found that the bias value works like a weighting function on the power spectral density (PSD):

$$E\{\hat{\Phi}_B(\omega)\} = \sum_\tau [w_B(\tau) r(\tau)] e^{-j\omega\tau} \tag{7}$$

$$w_B(\tau) = \begin{cases} 1 - \frac{|\tau|}{N}, & \tau = 0, \pm 1, \pm 2, \pm 3, \ldots, \pm(N-1) \\ 0, & otherwise \end{cases} \tag{8}$$

By considering the shape these weighting functions in the frequency domain - Fig 1, we found that they are poor approximations of Dirac impulse. In addition, unlike the Dirac delta function, they have a large number of side lobes. It follows that the bias of estimators can basically be divided into two components, which correspond respectively to the non-zero main lobe width and the non-zero side lobes height of the window function [11].

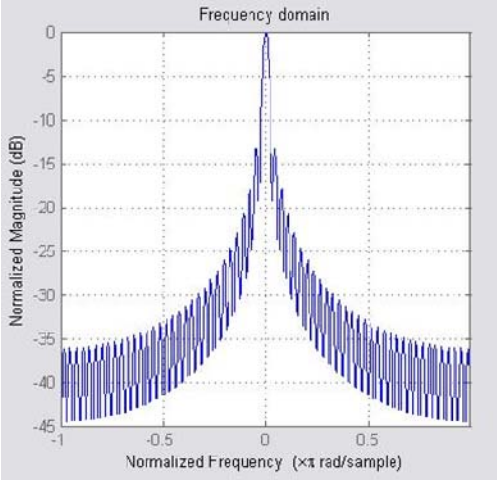

Figure 1: Rectangular window in frequency domain.

The principal effect of the main lobe is to smear or smooth the estimated spectrum. Assume, for instance, that $\varphi(\omega)$ has two peaks separated in frequency f by less than $1/N$ then these two peaks appear as a single broader peak, since the "response" of the "system" corresponding to the first peak does not get the time die out before the "response" to the second peak starts. This kind of effect of the main lobe on the estimated spectrum is called *smearing*. Owing to smearing, the periodogram-based methods cannot resolve details in the studied spectrum that are separated by less than $1/N$ in cycles per sampling interval. For this reason, $1/N$ is called the *spectral resolution limit* of the periodogram method.

The principal effect of side lobes on the estimated spectrum consists of transferring power from the frequency bands that concentrate most of the power in the signal to bands that

contain less or no power. This effect is called leakage. For instance, a dominant peak in $\varphi(\omega)$ may through convolution with the side lobes lead to an estimated spectrum that contains power in frequency bands where $\varphi(\omega)$ is zero. Note that the smearing effect associated with the main lobe can also be interpreted as a form of leakage from a local peak of $\varphi(\omega)$ to neighbouring frequency bands.

It follows from the previous discussion that smearing and leakage are particularly critical for spectra with large amplitude ranges, such as peaky spectra. For smooth spectra, these effects are less important. The bias of the periodogram estimator even though it might be severe for spectra with large dynamic ranges when the sample length is small, does not constitute the main limitation of this spectral estimator. In fact, if the bias where the only problem, then by increasing N (assuming this is possible) the bias would be eliminated. Hence, the periodogram is an *asymptotically unbiased spectral estimator*. The main problem of the periodogram method lies in its large variance.

Solution Procedure

As we have seen, the main problem with the periodogram is the high statistical variability of this spectral estimator, even for very large sample lengths. This effect may be reduced by truncating the sum in its definition equation (7), following this idea leads to some modified periodogram methods [12].

Blackman-Tukey Method

Blackman and Tukey (BT) estimator is given by the following equation:

$$\hat{\Phi}_{BT}(\omega) = \sum_{k=-(M-1)}^{M-1} w(k)\hat{r}(k)e^{-j\omega k} \qquad (9)$$

Where w(k) is an even function which is such that w(0)=1 and decays smoothly to zero with k, and where M<N. since it weights the lags of the sample covariance sequence, it's called a lag window. If w(k) is selected as a rectangular window then we simply obtain a truncated version of correlogram. However, the window may be chosen in many ways, the flexibility may be employed to improve the accuracy of the Blackman-Tukey spectral estimator.

Bartlett Method

The basic idea of Bartlett method is simple: to reduce the large fluctuations of the periodogram, split up the available sample of N observations to $L=N/M$ subsamples of M observations each, and then average the periodograms obtained from the subsamples for each value of ω. Since Bartlett method operates on data segments of length M, the resolution afforded should be on the order of $1/M$. Hence, the spectral resolution of the Bartlett method is reduced by a factor L (Number of segments), compared to the resolution of the original periodogram method. In return for this reduction in resolution, we can expect that the Bartlett method has a reduced variance. It can, in fact, be shown that the Bartlett method reduces the variance of the periodogram by the same factor L. The compromise between resolution and variance when selecting M (or L) is thus evident.

Welch Method

The Welch method is obtained by refining the Bartlett method in two respects. First, the data segments in the Welch method are allowed to overlap. Second, each data segment is windowed prior to computing the periodogram. By allowing overlap between the data

segments and hence by getting more periodograms to be averaged, we hope to decrease the variance of the estimated PSD. Additionally, the temporal window may be used to give less weight to the data samples at the ends of each subsample, hence making the consecutive subsample sequences less correlated to one another, even though they are overlapping. The principal effect of this "decorrelation" should be a more effective reduction of variance via the averaging. The use of windowed periodograms in the Welch method, as contrasted to the unwindowed periodograms in the Bartlett method, indeed offers more flexibility in controlling the bias properties of the estimated spectrum. The variance of the Welch spectral estimator is more difficult to analyze. However, there is empirical evidence that the Welch method can offer lower variance than the Bartlett method but the difference in the variances corresponding to the two methods is not dramatic.

Auto-Regressive Model

A rational PSD could be written in the following form:

$$\phi(\omega) = \frac{\sum_{k=-m}^{m} \gamma_k e^{-i\omega k}}{\sum_{k=-n}^{n} \rho_k e^{-i\omega k}} \tag{10}$$

The Weierstrass Theorem from calculus asserts that any continuous PSD can be approximated arbitrarily closely by a rational PSD of the form (10), provided the degrees *m* and *n* in (10) are chosen sufficiently large; that is, the rational PSDs form a *dense set in the class of all continuous spectra*. Since $\varphi(\omega)>0$, the rational spectral density can be factored as follows:

$$\phi(\omega) = \left|\frac{B(\omega)}{A(\omega)}\right|^2 \sigma^2 \tag{11}$$

So the arbitrary rational PSD can be associated with a signal obtained by filtering white noise of power σ^2 through the rational filter with transfer function $H(\omega) = B(\omega)/A(\omega)$. The filtering can be written in the time domain as

$$y(t) = \frac{B(q)}{A(q)} e(t) \tag{12}$$

Or, alternatively,

$$A(q)y(t) = B(q)e(t) \tag{13}$$

Where *y(t)* is the filter output, *e(t)* is white noise of variance equal to σ^2, and q^{-1} is the unit delay operator:

$$q^{-k}y(t) = y(t-k) \tag{14}$$

A signal *y(t)* satisfying this equation is called an *autoregressive moving average* (ARMA) signal. If m=0 then *y(t)* is an *autoregressive* (AR) signal; and *y(t)* is a *moving average* (MA) signal if n=0. In the ARMA class, the *autoregressive* or *all-pole* signals constitute the type that is most frequently used in applications. The AR equation may model spectra with narrow peaks; this is an important feature since narrowband spectra are quite common in practice. In addition, the estimation of parameters in AR signal models is a well-established topic; the estimates are found by solving a system of linear equations, and stability of the estimated AR polynomial can be guaranteed.

Special cases

The Factor Building: A permanently instrumented 15-story steel frame building

The UCLA Louis and Doris Factor building, was instrumented by the U.S. Geological Survey with an embedded 72 channel accelerometer network following the 1994 Northridge earthquake. The accelerometer network is distributed throughout the building and continuously recording building vibrations. In December 2003, the sensor network was upgraded by installing state-of-art data logging equipment and fibre optical network cables. To date, substantial data have been collected from ambient vibrations under different environmental conditions, as well as from low-amplitude vibrations from several earthquakes.

The Factor building, with its embedded sensor network, provides a unique platform for the identification of dynamic characteristics, structural performance monitoring, and damage detection. Establishing a reliable three-dimensional finite element model which accurately represents the stiffness of the structural system is an important step in assessing the structural performance and detecting damage under more significant shaking [13].

Four accelerometers exist at each floor above grade, oriented to record translational motions near the perimeter of the floor (two in each direction). Each of the two basement levels has an accelerometer to record translation in two directions, as well as two accelerometers to record vertical responses. The upgrade consisted of converting all 72 channels to a 24-bit network continuously recording data using Kinemetrics Antelope seismic software. The streaming data is currently viewable via the internet in near real-time. In February 2005, a 100m deep borehole seismometer was installed approximately 50m away from the building. The level of instrumentation provided in and around the Factor building makes it one of the most (if not the most) densely permanently instrumented buildings in North America.

Data collected on the Parkfield, CA earthquake (Mw = 6.0) occurred on 9/28/2004 17:15:24 UTC. Peak accelerations of 0.0025g were recorded at the roof of the building. In regards to this project, the unknown system is the Factor building structure, the measured outputs are the story accelerations, and the measured input is the ground motion. The inputs and outputs are recorded by the embedded sensor network. The results of system identification according to (3) and the results which calculated using Auto-Regressive model by the developed software can be seen below. The numeric results and finite element results are shown in table 1.

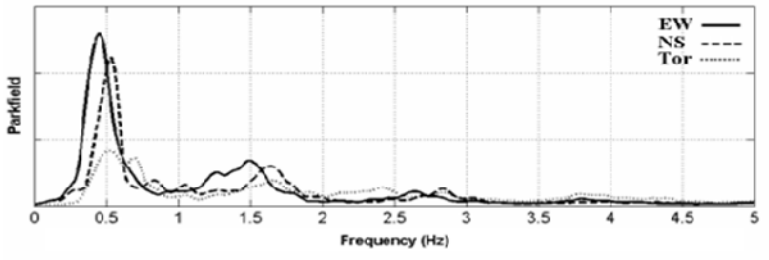

Figure 2: Factor building FRFs according to [13].

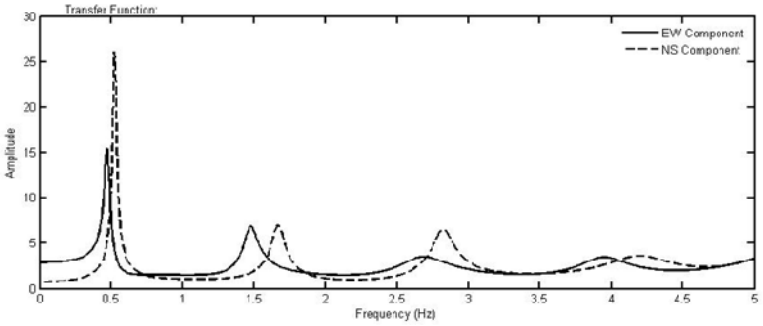

Figure 3: Factor building FRFs calculated by AR model.

Table 1: Numerical results of Factor building identification.

	Calculated by AR model		*According to [13]*		*Finite Element*
Mode Shape	*Freq (Hz)*	*Damping ratio %*	*Freq (Hz)*	*Damping ratio %*	*Freq (Hz)*
1st EW	*0.474*	*4.6*	*0.467*	*4.8*	*0.513*
1st NS	*0.523*	*2.3*	*0.506*	*4.7*	*0.511*
2nd EW	*1.480*	*3.8*	*1.488*	*5.4*	*1.507*
2nd NS	*1.670*	*2.9*	*1.665*	*4.9*	*1.445*
3rd EW	*2.690*	*7.6*	*2.677*	*4.4*	*2.534*
3rd NS	*2.822*	*2.9*	*2.862*	*4.9*	*2.386*

Shahid Rajaee Concrete Arch Dam

The Shahid Rajaee Dam is a 130m high and 420m long concrete arch dam over Tajan River in the south east of Sari the centre of Mazandaran province. Considering the fact that the cross section of the valley of the dam site is in the shape of a wide 'V' and that topographically the abutments spread out slightly toward the downstream, a three-centred double curvature with variable thickness arch dam layout was adopted.

The dam body is divided into 16 blocks and each of the blocks is about 26m long. This dam site is composed of massive and indurate sandstone and siltstone with thin alternated beds in parts. The bedding strike is generally parallel to the river channel and dips approximately 40 ° toward the right bank, thereby creating a dip slope left abutment and a scarp slope right abutment. The rock states at the dam foundation are very regular in their strike and dip and are essentially not offset by fault. Both the sandstone and siltstone are strong and hard. This dam was built in 1997.

By using the PADAP program, a dynamic response analysis was performed using finite element modelling and considering the Dam-Water interaction [15]. The first five mode frequencies of Shahid Rajaee arch dam were obtained and are shown in Table 2.

Table 2: Shahid Rajaee Finite Element results [15].

Finite Element	
Mode Shape	*Freq (Hz)*

1st	*2.022*
2nd	*2.359*
3rd	*3.002*
4th	*3.475*
5th	*4.079*

However, a strong-motion instrumentation program was developed at the Shahid Rajaee dam for the purpose of studying the post-seismic safety assessment of the dam. There are 5 SSA- 1 type (three components) strong motion seismometers on the dam. Three of them are at the crest level and two others are near the centre of the dam, but data are available only from four of them. Two earthquakes have been recorded recently. All the records are all in digital signal with a sampling rate of 0.02 s. Table 3, lists the recorded earthquakes from the dam instrumentation.

Table 3: Shahid Rajaee Earthquakes specifications [16].

Earthquake	*Date*	*Time (UTC)*	*Epicentre*		*Ml*	*Report*
Kojoor	*2004/5/28*	*12:38:44*	*51.56 E*	*36.30 N*	*6.1*	*BHRC*
Agh-ghola	*2005/1/10*	*18:47:30*	*54/54 E*	*37.12 N*	*6.1*	*BHRC*

Based on the data collected from the strong-motion instrumentation of the dam, system identification of the Shahid Rajaee dam was performed to identify the natural frequencies and damping ratios of the system. The response of the upstream-downstream direction of station ST5 (north-south direction) is considered as an output. Based on the AR model the identified FRF between each pair of input and output is discussed.

From different seismic events the identified frequency response functions, as shown in Figure 4, show very similar in shapes from different seismic event. The dominant frequency of FRF slightly changes from event to event. This can be explained as a result of the changes in water level in the reservoir. It is found that the dominant frequency of the dam reservoir system decreases with the increase in water level in the reservoir.

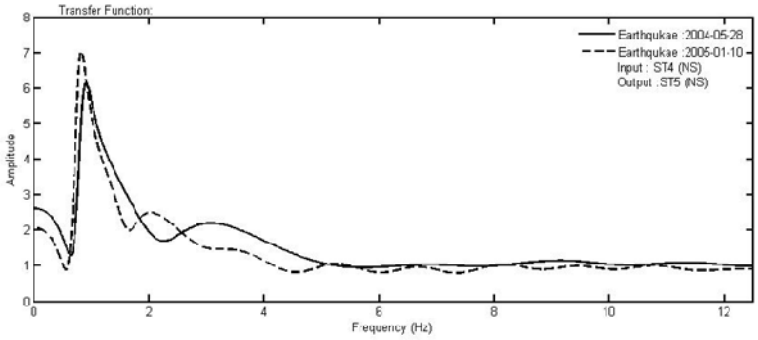

Figure 4: FRFs from different seismic events.

From the seismic event of 10 January 2005 the estimated FRF from same input (station ST4) but with different output (station ST5 or station ST3) are discussed. Figure 5 shows the comparison of the estimated FRF. It was found that the dominant frequency of the FRF shows little difference. This can be explained as being due to the motion of two different

blocks of the dam body during the same earthquake excitation. Generally, the identified FRF using different output data is quite similar.

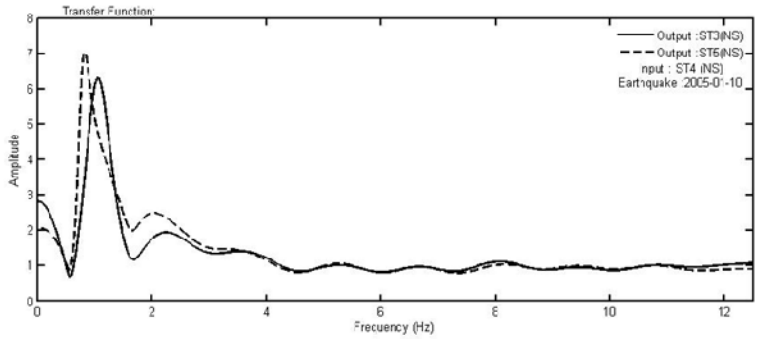

Figure 5: FRFs from same input but with different outputs

Discuss the FRF from the response data of 10 January 2005 earthquake, instead of using the output data in the north-south direction, the motion along the east-west direction is used as an output. Figure 6 shows the identified frequency response function between the output and input from station ST4. The identified natural frequencies are listed in Table 4. Comparison on the estimated FRF by using the two horizontal motions of stations ST3 and ST5 as an output shows that the amplitude of FRF in the north-south direction is smaller than the amplitude in the east-west direction.

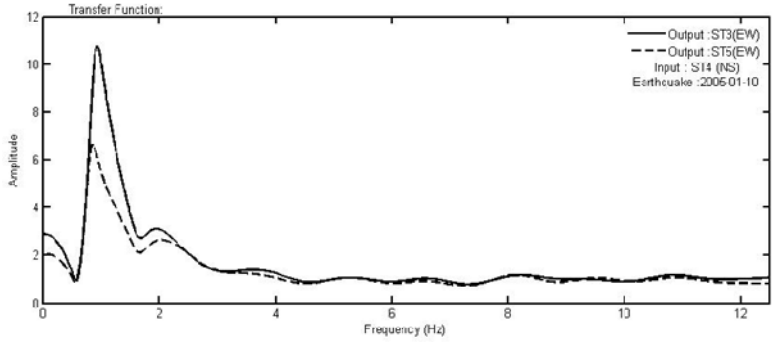

Figure 6: FRFs along the east-west direction

The order of the output and the inputs is shown in Table 4. Table 4 also lists the eight identified modal frequencies from the earthquake of 10 January 2005. The order of the AR model must be carefully selected in order to avoid pole-zero-cancellation. To verify the accuracy of the identification result, comparison between the estimated and recorded response was made and the cross correlation function of the residuals (the difference between the predicted response and the measured response) and input data was examined. The correlation function must be within a 95% confidence interval.

Table 4: Shahid Rajaee Frequency Responses.

North-South Component			*East-West Component*		
Mode	*Freq. (Hz)*		*Mode*	*Freq. (Hz)*	
shape	*ST3*	*ST5*	*shape*	*ST3*	*ST5*

1st	*0.85*	*0.94*	*1st*	*0.86*	*0.89*
2nd	*1.91*	*2.20*	*2nd*	*2.02*	*2.00*
3rd	*3.68*	*...*	*3rd*	*...*	*3.60*
4th	*5.27*	*5.36*	*4th*	*5.26*	*5.33*
5th	*6.80*	*6.56*	*5th*	*6.69*	*6.53*
6th	*8.08*	*8.17*	*6th*	*8.32*	*8.15*
7th	*9.34*	*9.41*	*7th*	*9.33*	*9.54*
8th	*10.73*	*10.79*	*8th*	*10.74*	*10.75*

ST3 orders for component NS and EW, respectively: 31,30
ST5 orders for component NS and EW, respectively: 25,18
ST4 orders for component NS, EW and VT, respectively: 25,35,21

CONCLUSIONS

The problem with the DFT of a finite signal is that it is equivalent to a convolution of an infinite signal with a weighting function in the frequency domain. In the other hand the auto-regressive model, which is capable of estimating the frequency content of a finite signal as if it were an infinite signal, overcomes this limitation. However, the additional step of calculating the model parameters may add noise to the prediction.

It must be pointed out that the identified natural frequencies from the seismic response data did not coincide with the analytical finite element results. The reason is that the mode shapes calculated from numerical analysis are developed under some limitations and assumptions of the computer program. It is necessary to develop a more sophisticated computer program for the dam-reservoir system.

The results of identification can be applied to the post-earthquake safety evaluation of dam-reservoir system. One of the tests is to predict the dam responses from the next seismic excitation using the identified model from previous results.

REFERENCES:

1. **Skolnik, D.** The Factor Building. *nees@UCLA.* [Online] 2005. http://www.seas.ucla.edu/~skolnik.
2. **Zhang, Y., et al.** Modal parameter identification using response data only. *Journal of Sound and Vibration.* 2005, Vol. 282, pp. 367-380.
3. **Yao, J. T.P.** Identification of structural damage in civil engineering. *CISM courses and lectures No.296.* Wien, New York : Springer-Verlag, 1982.
4. **Stoica, P. and Moses, R.** *Introduction to Spectral Analysis.* Upper Saddle River, N.J. : Prentice-Hall, Inc., 1997.
5. **Poor-Sartip, B. and Lotfi, V.** Modal analysis of arch dams including dam-water interaction in time domain. *Journal of Faculty of Engineering, Tehran University.* 2007, Vol. 41, 6, pp. 683-697. [Persian].
6. **Newland, D. E.** *An introduction to Random vibrations and spectral analysis.* New York : John Willey & Sons, Inc., 1987.
7. **Neild, S. A., McFadden, P. D. and Williams, M. S.** A review of time-frequency methods for structural vibration analysis. *Engineering Structures.* 2003, Vol. 25, pp. 713-728.
8. **Natke, H. G.** Identification of vibrating structures. *CISM courses and lectures No.272,.* Wien, New York : Springer-Verlag, 1982.
9. **McVerry, G. H.** Structural identification in the frequency domain from earthquake records. *Earthquake engineering and Structural dynamics.* 1980, Vol. 8, pp. 161-180.
10. **Ljung, L.** *System Identification - Theory for the User.* 2nd edition. Upper Saddle River, N.J. : Prentice-Hall, Inc. , 1999.
11. **Imai, H., et al.** *Fundamentals of system identification in structural dynamics.* s.l. : Technical Report NCEER 89-0008, 1989.
12. **Huang, C. S.** Structural Identification from ambient vibration measurement using the multivariate AR model. *Journal of Sound and Vibration.* 2001, Vol. 241, 3, pp. 337-359.
13. **Hong, K. S. and Yun, C. B.** Improved method for frequency domain identification of structures. *Engineering Structures.* 1993, Vol. 15, 3, pp. 179-188.
14. **Benzoni, G. and Gentile, C.** Two approaches to identify equivalent structural models from earthquake responses. *Soil dynamics and Earthquake engineering.* 1993, Vol. 12, pp. 113-125.
15. **Beck, J. L. and Jennings, P. C.** Structural identification using linear models and earthquake records. *Earthquake engineering and Structural dynamics.* 1980, Vol. 8, pp. 145-160.
16. *Signal Processing Toolbox For Use with MATLAB.* s.l. : The MathWorks, Inc., 2002. User's Guide. Ver. 6.0.

17. Iran Strong Motion Network (ISMN). *Building & Housing Research Center (BHRC).* [Online] http://www.bhrc.ac.ir/ISMN.