4th IASPEI / IAEE International Symposium:

Effects of Surface Geology on Seismic Motion

August 23–26, 2011 · University of California Santa Barbara

RAYLEIGH-WAVE DISPERSION CURVE: A PROXY FOR SITE EFFECT ESTIMATION?

Héloïse Cadet
ISTerre, CNRS,
University of Grenoble
Grenoble
France

Giovanna Cultrera
Istituto Nazionale di
Geofisica e Vulcanologia
Roma
Italy

Valerio De Rubeis
Istituto Nazionale di
Geofisica e Vulcanologia
Roma
Italy

Pierre-Yves Bard
ISTerre, CNRS, IFSTAR
University of Grenoble
Grenoble
France

ABSTRACT

One of the open issues on the effects of surface geology regards the estimation of site response when limited resources are available. In that restrictive context, one solution is to use soil characteristics as proxy. Despite its extensive use, the most common proxy, Vs30, is presently criticized because it cannot carry alone the main physics of site response. We propose here a statistical investigation of the capabilities of another proxy, the Rayleigh-wave dispersion curve, DC. When considered over a broad enough frequency band, it can provide deeper information missing in the single Vs30 parameter.
A set of shear-wave velocity profiles measured for more than 600 Japanese KiK-net stations is used to compute theoretical dispersion curves (DC) and theoretical SH transfer functions (SH), while instrumental surface/downhole spectral ratios were calculated in a previous work (Cadet *et al.*, 2011a). Canonical correlation techniques are applied to this large data set to analyze the relationship between DC and theoretical or empirical site responses. The results indicate very encouraging qualitative statistical relationships between DC and site amplification for numerically derived SH transfer functions, showing significant canonical couples with correlations up to 0.95. Results for instrumental surface/downhole transfer functions correspond to lower correlations (up to 0.73) but still allow the development of quantitative relationships.

INTRODUCTION

Different proxies are used to define the site effect for site with few information. The most common one is the time-averaged shear-wave velocity over the first 30 meters, Vs30. It is widely used in code provisions (in Europe with the Eurocode EC8, and in the USA with the National Earthquake Hazards Reduction Program NEHRP) and in ground motion prediction equation GMPE (Power et al. 2008). However its capability of carrying site effects information is drastically criticized (Mucciarelli and Gallipoli 2006; Castellaro et al. 2008; Lee and Trifunac 2010). Indeed it is a very restrictive parameter that does not carry any information about frequency content and concerns only a shallow depth.
Nowadays, other proxies are proposed. For example, in some GMPEs the combination between Vs30 and the depth to the bedrock is introduced for site effect consideration (e.g. Field 2000). Cadet et al. (2011a) proposed a site amplification prediction equation (SAPE) using the combination of Vs30 and the fundamental resonance frequency f_0. In these recent propositions Vs30 is still needed.
The geophysical campaign to get Vs30 might be either with active source (usually not easy to generate especially at large distances) or with passive source. In this last case, the ambient seismic noise is recorded and analyzed to compute the dispersion curve, DC (velocity of the surface waves according to their frequency). Then, this dispersion curve is inverted to obtain shear-wave velocity profiles, Vs(z). The inversion step is a research topic by itself (in the noise analysis context: Wathelet 2008; Wathelet et al. 2004). Some efforts have been done to provide a step-by-step inversion procedure (Renalier et al. 2009, Savvaidis et al. 2009). However despite these guidelines, the process is time-consuming and might be user dependent (Cornou et al. 2006). Having, in one hand, the difficulty of the inversion step and, in the other hand, the need of a site effect proxy carrying frequency information, we decided to test the capability of the DC to estimate the site effect. Thus the goal of this work is to investigate in a statistical way the ability of the dispersion curve to capture horizontal site response phenomena. Moreover, the dispersion curve can be defined in different frequency

band width depending on the survey. Large array of sensors will allow to characterize low frequencies whereas small arrays aperture will characterize high frequencies. The desired frequency band width is constrained by the survey cost and the natural filter that usually does not allow to describe the dispersion curve below the resonance frequency of the site (Scherbaum et al. 2003). Thus the fundamental resonance frequency is a reference limit of the frequency band described by the dispersion curve.
We applied the well-known multivariate statistical analyses techniques, namely, the canonical correlation, as already used by Theodulidis et al. (2008), Cara et al. (2008) and De Rubeis et al. (this issue) to quantify the correlation between physical parameters, called proxies, and numerical site effect estimation. Different canonical correlation tests were driven to identify the more adapted proxy for a proper site effect estimation. The proxies tested were the classical Vs30 combined with the fundamental resonance frequency, compared to different frequency bands of the dispersion curve.

DATA AND METHODS

In this paper the dispersion curves and numerical transfer functions are computed for the statistical analysis. Both computations are detailed in this chapter. A comparison with empirical data is also carried out. The main steps of the empirical spectral ratios estimation are described. Then the canonical correlation technique is briefly detailed.
The velocity profiles and the empirical data are extracted from the Japanese KiK-Net accelerometric network (http://www.kik.bosai.go.jp). This network is devoted to combine outcrop and down-hole measurements (up to 3500m deep, with a mean depth around 222 m). Moreover, each of the 689 stations has been characterized with P- and S-waves velocity profiles from down-hole measurements (Kinoshita, 1998; Aoi et al., 2000; Fujiwara et al., 2004; Pousse, 2005). This network constitutes an abundant source of information and thus has been the topic of numerous studies on the complete database (among others: Oth et al., 2011; Cadet et al., 2011a). In this study, a pre-selected set of 622 stations has been used (Pousse, 2005).

Dispersion Curve computation

The theoretical DC computation is based on the eigen values problem described by Thomson (1950) and Haskell (1953), then modified by Herrmann (1994). For the Love and Rayleigh surface waves, the motion equation can be reduced to a first order differential equations system. In the case of tabular layers (one dimensional, 1D, case with horizontal and isotropic layers), this problem can be solved with the propagation matrix method (Aki and Richards, 2002), as in the GEOPSY software (http://www.geopsy.org/) used for the DC computation. The input parameters are defined at each depth: P-waves velocity, Vp, S-waves velocity, Vs, and the density with a minor influence on the DC according to Whatelet (2004). The DCs are computed between 0.1 and 100 Hz for hundred frequency samplings in logarithmic scale. Reminding that the natural filter usually does not allow to describe the DC below the resonance frequency of the site, the fundamental resonance frequency might be a limit of the frequency band described by the dispersion curve. Considering this characteristic, it seems pertinent to work in a dimensionless frequency space, that is to say normalized the frequencies by the fundamental resonance frequency, f_0. This frequency has been automatically selected on the numerical transfer function. In that case, the "ad" suffix is used. For the statistical analysis the DCs are resampled with only six samplings from $0.3*f_0$ to $6*f_0$ in dimensionless frequency. This is done both for purpose of simplicity and also because each sample has to be independent.

Theoretical transfer function computation

The SH-waves transfer functions (SH) are performed for each velocity profiles. SH is computed numerically with a 1D reflectivity model (Kennett, 1974) reproducing the response of horizontally stratified layers excited by a vertically incident SH plane wave. The original software written by JC Gariel and P.Y. Bard, was used previously in a large number of investigations (Bard and Gariel, 1986; Theodulidis and Bard, 1995; among others). The velocity profiles have been exploited for both the DC computation and the SH computation with the velocity of the bedrock taken as the last measured velocity of the profile.
In addition to the shear-wave velocity profile derived from downhole measurements, densities and damping values are also needed to compute the site response. The density is assumed to increase with increasing S-wave velocity and their values was set as follows:

$$\rho = \begin{cases} 1.7 g/cm^3 \; for \; V_s(z) \le 180 \, m/s \\ 2.0 g/cm^3 \; for \; 180 \le V_s(z) \le 360 \, m/s \\ 2.2 g/cm^3 \; for \; 360 \le V_s(z) \le 1500 \, m/s \\ 2.5 g/cm^3 \; for \; V_s(z) \ge 1500 \, m/s \end{cases}$$

The material damping is introduced through a quality factor for shear waves, assumed to be independent of frequency: Qs=Vs/10, with

Vs in m/s. This assumption provides Q values close to the one chosen by Raptakis et al. (2000), Chavez-Garcia et al. (2002) and Marrara and Suhadolc (2001) for 1D simulations. The SHs are computed for 2048 frequency samples regularly spaced from 0 to 25 Hz. They were resampled in dimensionless frequency space with six samplings for the statistical analysis as mentioned for the DC. The SH and DC are both completely defined over the dimensionless frequency band from $0.3*f_0$ to $6*f_0$ for 447 sites.
The SH is computed with respect to an outcropping rock site having the velocity of the bottom layer of the site under consideration. To be consistent with the empirical data processing, a Konno-Ohmachi (Konno and Ohmachi 1998) smoothing was applied to the SH. From Cadet et al. (2011b) numerical data need a stronger smoothing and the selected smoothing parameter in the Konno-Ohmachi procedure is here chosen as b=10.

Empirical transfer functions (Fourier horizontal)

Using the KiK-Net data, borehole spectral ratios (BSR) are derived using the outcrop over the down-hole recordings of each site (details can be found in Cadet *et al.*, 2011b). However, it is much more common to reefer to an outcrop rock site than to a deep rock site, thus Cadet *et al.* (2011b) proposed a correction procedure. This procedure corrects from the depth effects (effects due to the depth of the down-hole sensor that gives a higher amplification than with an outcrop reference) and normalizes toward a standard reference rock site (Cadet *et al.* 2010). The corrected normalized ratios, BSRcn, were used as empirical transfer functions for canonical correlations. In that case, the database is reduced to 207 sites for which the BSRcn is defined in the dimensionless frequency from 0.3 to 6.1.

Canonical correlation

This technique is applied in order to compare the SH (variables to explain) to the explicative variables, either the DC or Vs30, in the dimensionless frequency space that implicitly includes f_0 information.
From Davis (2002), the canonical correlation analysis underlines the relationships between an explicative data group X and a group of data to explain Y. This analysis optimizes the linear combinations of the variables X, so called X_{can} and of the variables Y, called Y_{can}, in order to get the bets correlation value between X_{can} and Y_{can}. Those linear combinations are called canonical variables and are defined by their weights:

$$X_{can} = \sum_{i=1}^{n_x} (a_i \times X_i) \tag{1}$$

With a_i the weight associated to X_i and n_x the number of variables X. Similarly the canonical variable Ycan is defined:

$$Y_{can} = \sum_{j=1}^{n_y} (b_j \times Y_j) \tag{2}$$

With b_j the weight associated to Y_j and n_y the number of variables Y.
The vectors $A=(a_1,\ldots,a_{nx})$ and $B=(b_1,..b_{ny})$ are the weights, respectively, of the explicative variables and of the variable to explain. The variables X_{can} and Y_{can} are linked by a linear relation:

$$Y_{can} = \alpha \times X_{can} + \beta \tag{3}$$

with α and β real scalar. The linear relation is represented here in the (X_{can}, Y_{can}) plane (***fig. 1, first line, 2nd, 4th and 6th graphs***).
To understand the relations between X and Y, correlations between the original variables X_i and the canonical variables X_{can} (or between Y_j and Y_{can}) are computed. The result of such correlations are presented for each variable in a bar plot (***fig. 1, first line, 1st, 3rd, and 5th graphs for the Y and Y_{can} correlation and 2nd line for the X and X_{can} correlation***). For example, the variable X_i with the highest correlation is the one that better explain the variables Y. To illustrate this case, referring to fig. 1.B, we see that X variable n°3 is highly correlated with its Xcan and it decisively contribute to the canonical significant correlation with Ycan. Ycan is well represented by variable Y n°1. Highly correlated original variables are the source for correct interpretation of canonical couples Xcan and Ycan.
The relationships between X and Y depend on n combinations of the weights A and B, with n the smallest value between n_x and n_y. The first possibility is called the first canonical couple and has the highest correlation value. For example the first, second and third canonical couples are represented respectively in fig. 1.A, 1.B and 1.C.

In this study the X variables are either the DC samples or the Vs30 and Y variables are the SH samples.
Two main hypotheses have to be respected for the canonical correlation analysis: (i) a normal distribution of each variable and (ii) no or few outliers. These requirements have to be fulfilled in order to avoid to be influenced by a tail or an asymmetric shape.
To be closed to the normal distribution (first requirement) a Box Cox transform (Chen et al. 2002) has been used in both final X and Y families. The Box-Cox transformation of a set of variables y is basically defined by a coefficient λ such as:

$$BoxCox(y) = \begin{cases} \dfrac{y^{\lambda} - 1}{\lambda}, \; if \; \lambda \neq 0 \\ \log(y), \; if \; \lambda = 0 \end{cases} \quad (4)$$

λ is chosen such as BoxCox(y) is the closest to a normal distribution. If λ equals to zero then the transformation is equivalent to a natural logarithm transformation. A high value of λ will traduce a variable distribution far from a normal distribution, thus these variables have to be either rejected or taken into account carefully. For all the following studies the Box-Cox transformation is used for all parameters. The four families (SH, DC, Vs30 and f_0) used here are BoxCox transformed with λ values in the table 1 for each sample of the SH and DC, λ=0.048 for Vs30 and λ=0.373 for f_0.

Table 1. λ values of the Box Cox transformed for SH, DC and BSRcn in the dimensionless frequency space

frequency	0.3	0.55	1	1.8	3.3	6.1
SH ad	-1.34	-1.04	0.36	-0.35	-0.09	0.29
DC ad	0.67	0.65	0.55	-0.20	-0.19	-0.28
BSRcn ad	-0.97	-0.58	-0.93	-0.12	0.07	0.26

To avoid the presence of outliers (second requirement), we define them as sites having more than 10% of its variables (in log10) that are smaller or higher than the mean of the considered variable +/- two standard deviations. The outliers determination is computed separately on both the X and Y datasets. The outliers of both datasets are disregarded to build-up the final X and Y families on which statistical analysis has been derived.

DATA PROCESSING

The goals are (i) to show that it is possible to evaluate the site effect represented by the transfer function (SH theoretical or empirical) by using the dispersion curve (DC) as a proxy; and (ii) to compare this DC proxy with the classical Vs30 proxy. Indeed, one of the main practical issues is to choose easily available proxies that are the most relevant for site effect estimation.
To fulfil these goals, we studied the correlation between the transfer functions SH and the dispersion curve DC defined in different frequency ranges. We then correlated the SH with the Vs30 parameter to compare its capability to approximate the site transfer function.
Most of the analyses have been carried out using SH and DC in the dimensionless frequency space, that is by dividing the frequency range by the resonance frequency f_0 evaluated at the site. In this way we are able to propose site effect proxies which are site specific because it depends on the fundamental resonance frequency of the site. Moreover it is more adapted to give advices on the DC frequency band in dimensionless space. Notice that working in dimensionless frequency implies that f_0 is a part of the proxies.

Theoretical SH versus DC (0.3* f_0 to 6* f_0): ideal case

The 6 samplings of the DC from 0.3 to 6 in dimensionless frequency space are used as explicative parameters of the theoretical transfer function SH (*fig. 1*). The first 3 canonical couples of the canonical correlation have a correlation of 0.97, 0.86 and 0.78 respectively. The associated probabilities are less than 1/10000, this very low value reinforces the idea that the correlation, behind r values, have good significant values.
In the first couple (*fig. 1.A*), the DC samples below f_0 (black bar plot at the bottom, *fig. 1.A*) are correlated mainly with the SH around f_0 (grey bar plot at the top, *fig. 1.A*). This relation is physically understandable: the high DC velocity at low frequency is linked to a hard bedrock at depth, thus probably to an elevated velocity contrast value; in that case we expect a high SH amplification. On the other hand, the last 3 samples of the DC (black bar plot at the bottom, *fig. 1.A*) are anti-correlated with SH around f_0 (grey bar plot at the top, *fig. 1.A*): when the high frequency DC velocity is high, it is most probably the case of a stiff superficial soil that may correspond to a relatively small amplification in the SH.

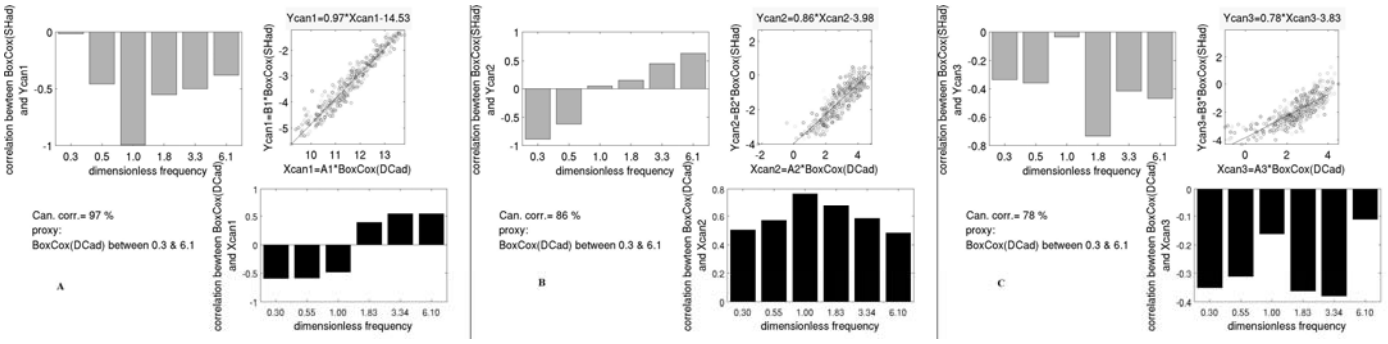

*Fig. 1. Canonical correlation between the explicative variables DC from 0.3*f_0 to 6*f_0 and the variables to explain SH from 0.3*f_0 to 6*f_0*

The second and third canonical couples (*fig. 1.B and C*) explain better the other frequency samples of the SH but the physical explanation is not as obvious as for the first canonical couple. In the 2nd canonical couple the all DC range, with particular relevance of 1*f_0, is anti-correlated with the SH at 0.3*f_0 and correlated with SH at 6*f_0. Finally the 3rd canonical couple underlines a correlation between the DC from 1.8*f_0 to 3.3*f_0 with the SH at 1.8*f_0. Both relations from the 2nd and 3rd canonical couples are difficult to explain with a simple physics.

Those results are promising because they show that the DC can explain up to 97% of the SH. However, it is difficult to obtain the DC from 0.3*f_0 to 6*f_0 in a real practical case. It is then useful to estimate how much we loose when using either less samplings of the DC or the Vs30 parameter.

Theoretical SH versus DC defined only over few samples: restrictive case

When the DC is reduced to the high frequency samples, it is still providing valuable information on SH. Indeed with a DC reduced to frequencies greater than 1*f_0, the correlation of the first canonical couple is very high (0.96). *Fig. 2A* shows that the DC value around f_0 is correlated with the SH value around f_0.: in case of high bedrock velocity, the DC velocity at low frequency will be high and the large velocity contrast might imply a large amplification on SH. A more restrictive DC dimensionless frequency band (f>1.8*f_0 or f>3.3*f_0) is still explaining the SH with 85% of correlation (*fig. 2.B and C*). Both frequency band restrictions are equivalent in terms of correlation values and explanation of the inter-relationships. This equivalence is crucial because it indicates that it is not necessary to put our computational and experimental efforts towards the low frequencies, usually more difficult to retrieve in a field experiment. These two last canonical correlations show that the DC around 3.3*f_0 is anti-correlated with the SH around 1*f_0: high DC velocity at relatively high frequency (around 3.3*f_0) may indicate a rock site, thus the SH amplification should be small. The correlation between the DC around 3.3*f_0 and the SH around 6*f_0 is less important than the first observation and there is not an obvious physical explanation of it.

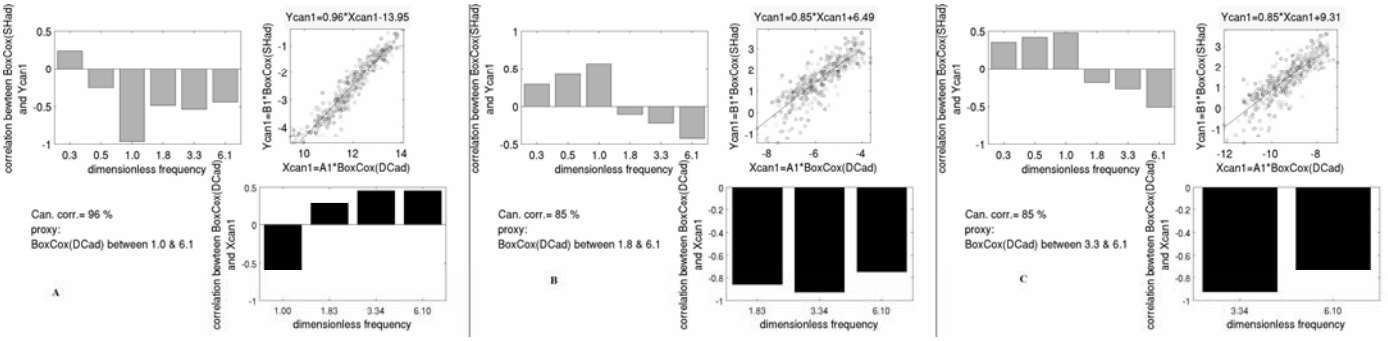

*Fig. 2. Canonical correlation between DC from (a)1*f_0 to 6*f_0; (b)1.8*f_0 to 6*f_0; (c) 3.3*f_0 to 6*f_0 and SH from 0.3*f_0 to 6*f_0*

Figure 3 represents the correlation value of canonical correlations between DC and SH, for different DC frequency samplings selection. The frequency sampling stays the same (defined in table 1), only the samples selection is changing. The samples selection of the DC starts from α*f_0 (abscise of *fig. 3*) and finishes after one, two or three samples (blue, red and black curves in *fig. 3*). The ordinate of *fig. 3* shows the correlation value using only one, two or three samples of the DC. For example, the round red point at 1*f_0

represents the correlation value (about 0.93) of the DC defined from $1*f_0$ to $1.8*f_0$ with SH defined from $0.3*f_0$ to $6.1*f_0$.
We can see that, with either one, two or three samples, the DC will better explain the site effect if the DC frequency selection starts from $1*f_0$ (correlation value around 0.96 with 3 samples, 0.93 with 2 samples and 0.84 with one sample). Thus there is no need to go below the fundamental resonance frequency as the correlation is lower when using low frequency content.
The 2-samples DC definition provides correlation values very close to the 3-samples DC definition (red curve versus black curve). As said in the previous paragraph, we get 0.85 of correlation using DC from $3.3*f_0$ to $6.1*f_0$ as for the DC defined over 3 samples from $1.8*f_0$ to $6.1*f_0$. Thus there is no need to extend to $1.8*f_0$ in that case, two samples are enough.
Finally, using only one sample, the best compromise between correlation value and facility to get the DC in a field work (that is to say above f_0) is to define the DC around $3.3*f_0$ (correlation of 0.82). In the restrictive situation where DC cannot be defined below $1*f_0$ (included), the difference between one, two and three samples is not so huge (*fig. 3*).
This study allows us to advice for the DC proxy even if a narrow frequency band is described.

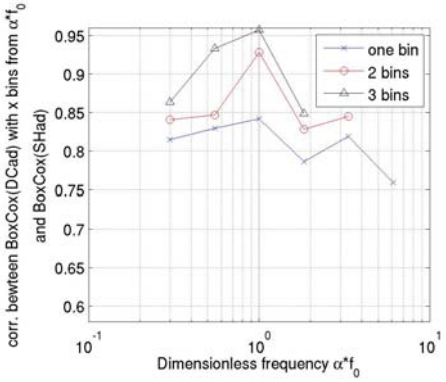

Fig. 3. Evolution of the correlation from the canonical correlation between DC described on few samples (one, two or three) and SH from $0.3*f_0$ to $6*f_0$

RESULTS AND DISCUSSION

Is DC a good proxy for empirical data?

In order to validate our proposition to use DC as a proxy, canonical correlations have been realized between the DC and the 207 empirical transfer functions (BSRcn) in the dimensionless frequency space. Empirical data should carry more variability (measurements accuracy, 2D-3D effects …) than SH that is a 1D model. As expected, the correlations values are relatively lower than using the numerical models. With restrictive frequency samplings of the DC ($f> 1*f_0$, $f>1.8*f_0$ and $f>3.3*f_0$) the correlations are 0.69, 0.64 and 0.64, respectively. We focus on the results for the shortest frequency range ($f>3.3*f_0$), because it was demonstrated in the previous section that the DC from $3.3*f_0$ to $6.1*f_0$ is able to explain the theoretical SH around $1*f_0$ with a correlation of 0.85. In the case of empirical transfer function we find similar results (*Fig. 4*) but the correlation value decrease to 0.64. The physical explanations are similar.
Figure 5 is equivalent to Fig. 3 but using empirical data. From Fig. 5, it seems useless to use the DC defined only over one sample at $1.8*f_0$ as the correlation drops significantly. This effect is also seen when using two samples: from 1.8 to 3.3 the correlation is not better than using DC from 3.3 to 6.1. The best one sample DC proxy (when information is not available below f_0) is at $3.3*f_0$ (correlation 0.63) as with the numerical transfer functions.
On the contrary of the numerical transfer function, the low DC frequency at $0.55*f_0$ explains in a better way the BSRcn than at $1*f_0$ (*fig. 5* compared to *fig. 3*). This observation underlines that there is different behavior between the numerical and the empirical transfer functions especially at frequency below f_0.

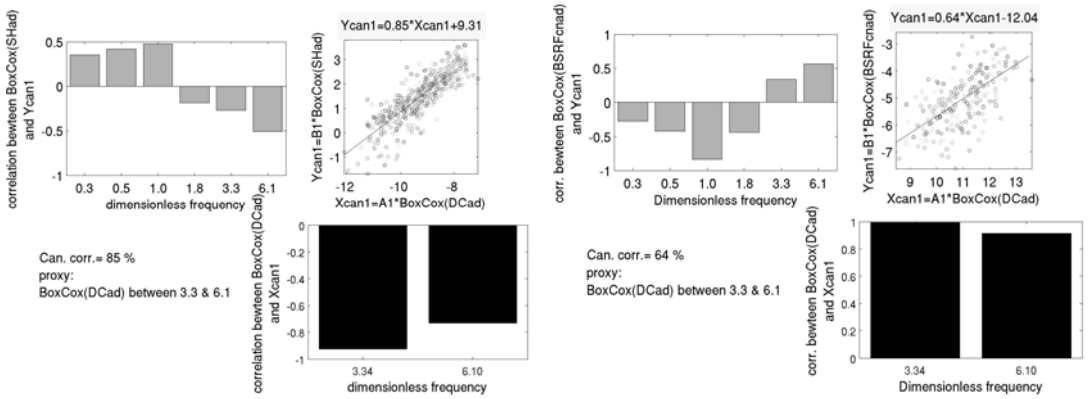

*Fig. 4. Canonical correlation between DC from 3.3*f_0 to 6*f_0 and SH (left) or BSRcn (right) from 0.3*f_0 to 6*f_0*

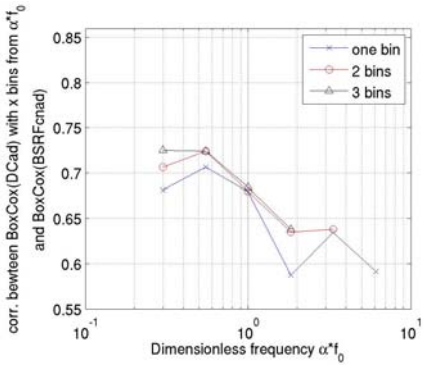

*Fig. 5. Evolution of the correlation from the canonical correlation between DC described on few samples and BSRcn from 0.3*f_0 to 6*f_0*

<u>Is DC proxy better than Vs30 proxy?</u>

In order to emphasis the proxy quality of the DC a comparison is conducted when using Vs30 alone with SH. Here the correlation is realized in real frequency domain in order not to introduce f_0 information (*fig. 6*). As all previous computations were performed in the dimensionless frequency, here, for purpose of comparison, canonical correlation between DC and the BSRcn is driven in the real frequency domain (*fig. 7*).

The correlation using Vs30 has a value of 0.69. It focuses on the middle frequencies from about 1 Hz to 3 Hz. The DC proxy from 3.3 Hz to 6.1 Hz in real frequency domain provides a better correlation of 0.82. This correlation underlines mainly a frequency band from 0.5 Hz to 2Hz (*fig. 7*), slightly lower than the underlined frequency band with Vs30. Using only two frequency samples of the DC, the DC proxy is better than the Vs30 proxy.

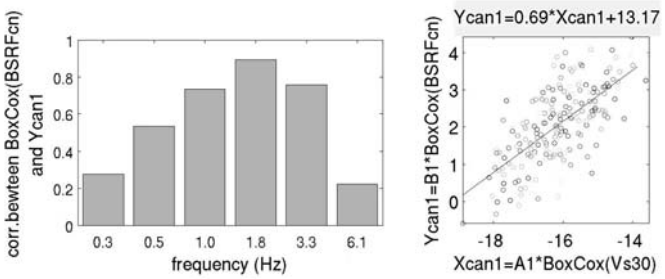

Fig. 6. Canonical correlation between Vs30 and BSRcn from 0.3 Hz to 6.1 Hz

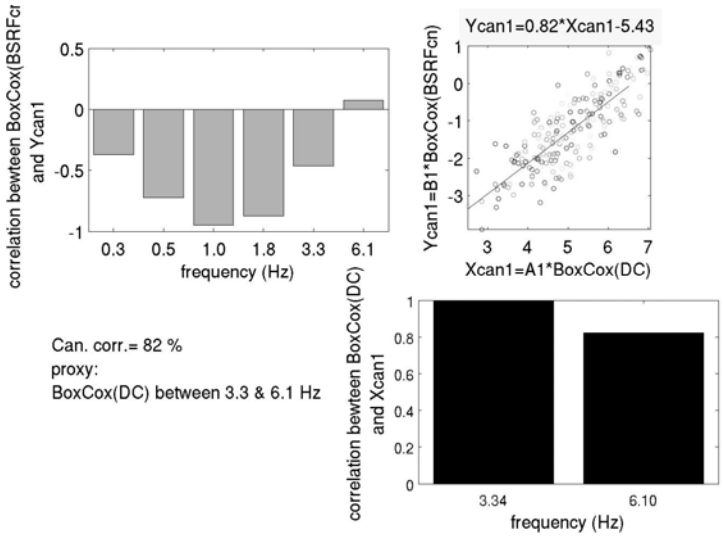

Fig. 7. Canonical correlation between DC from 3.3 Hz to 6.1 Hz and BSRcn from 0.3 Hz to 6.1 Hz

CONCLUSION

This study clearly showed that the dispersion curve carries a high potential for site effect estimation. Indeed, the DC is able to explain a large part of the transfer function, mainly around f_0 for both numerical and theoretical site effect estimations. Moreover, even if the DC is defined over a very narrow frequency band, its proxy capability for numerical transfer function is still high with a correlation value of 0.85 for a theoretical dispersion curve defined from $3.3*f_0$ to $6*f_0$ and even 0.82 with DC defined only around $3.3*f_0$. For the same frequency sampling selection, in the case of empirical transfer function, the correlation drops down to 0.64 and 0.63 respectively. This decrease might be link to the fact that empirical data also carry 2D and 3D effects together with non-linear effects, even though the selected data are not prone to such behavior. These complex effects are not included in the DC that comes from numerical modeling. For a practical utilization, we might advice to at least derive the DC around $3.3*f_0$ or better from $3.3*f_0$ to $6*f_0$. Indeed this "low cost" characterization allows a canonical correlation of 0.85 with the SH defined from $0.3*f_0$ to $6*f_0$ and 0.64 with empirical spectral ratios defined from $0.3*f_0$ to $6*f_0$.
Finally, this proxy seams to be much more appropriate than the Vs30 proxy. This has been proved using empirical transfer functions: the DC defined from 3.3 Hz to 6.1 Hz has a canonical correlation of 0.82 with the empirical transfer function in real frequency space, whereas the Vs30 has only 0.69 of correlation. Moreover the time of data analysis needed to obtain Vs30 is much longer than for the DC estimation, thus it is of great interest to develop relationships using the DC information.
Finally, all canonical correlations of this study have good significant values because they are always associated with a probability less than 1/10000. This enhances the capability of the tested proxies.
As done by De Rubeis *et al* (2011) the next step of this work will be to develop a quantitative analysis in order to propose site effect estimation based on the dispersion curve.

REFERENCES

Aki K., Richards P. G. [2002] "Quantitative seismology", Second Edition, University Science Books.

Bard P.-Y., Gariel J.-C. [1986] "The seismic response of two-dimensional sedimentary deposits with large vertical velocity gradients", Bulletin of the Seismological Society of America, Vol. 76, No. 2, pp. 343-346, April 1986.

Cadet H., Bard P.-Y., Rodriguez-Marek A. [2010] "Defining a Standard Rock Site: Propositions Based on the KiK-net Database", Bulletin of the Seismological Society of America, Vol. 100, No. 1, pp. 172–195, February 2010.

Cadet H., Bard P.-Y., Duval A.-M., Bertrand E. [2011a] "Site effect assessment using KiK-net data: part 2 - site amplification

prediction equation based on f0 and Vsz", Bulletin of Earthquake Engineering, accepted 20 June 2011, published one line.

Cadet H., Bard P.-Y., Rodriguez-Marek A. [2011b] "Site effect assessment using KiK-net data: part 1 - A simple correction procedure for surface/downhole spectral ratios", Bulletin of Earthquake Engineering, accepted 9 May 2011, published one line.

Castellaro S., Mulargia F., Luigi R.P. [2008] "Vs30: Proxy for seismic amplification?", seismological research letters, July/August 2008, vol. 79, number 4, pages 540-543.

Chávez-García F.J., Raptakis D., Makra K., Pitilakis K. [2002] "Importance of the reference station in modeling site effects up to larger frequencies. The case of Euroseistest", 12th European conference on earthquake engineering, paper reference 589, London, 2002.

Chen G., Lockhart R.A., Stephens M.A. [2002] "Box-Cox transformations in linear models: large sample theory and test of normality", the Canadian journal of statistics, Vol. 30, 2002.

Cornou C., Ohrnberger M., Boore D., Kudo K., Bard P.-Y. [2006] "Derivation of structural models from ambient vibration array recordings: results from an international blind test", Third International Symposium on the effects of surface geology on seismic motion, Grenoble, France, pp 92.

Davis J.C. [2002] "Statistics and data analysis in geology", 3rd edn. 2002, Wiley, New York.

De Rubeis V., Cultrera G., Cadet H., Bard P.-Y., Theodoulidis N. [2011] "Statistical estimation of earthquake site response from noise recordings", ESG4, this issue.

Eurocode 8 [1998] "Design of structures for earthquake resistance, part 1: General rules, seismic actions and rules for buildings", EN 1998-1, European Committee for Standardization (CEN), http://www.cen.eu/cenorm/homepage.htm.

Field E.H. [2000] "A modified ground motion attenuation relationship for southern California that accounts for detailed site classification and a basin depth effect", Bull. Seism. Soc. Am. 90, S209-S221.

Haskell N. A. [1953] "The dispersion of surface waves on multilayered media", Bull. Seism. Soc. Am. 1953, 43, 17–34.

Herrmann R. B. [1994] "Computer programs in seismology", vol IV, St Louis University.

Kennett B.L.N. [1974] "REFLECTIONS, RAYS, AND REVERBERATIONS". Bulletin of the Seismological Society of America. Vol. 64, No. 6, pp. 1685-1696. December 1974.

Konno K., Ohmachi T. [1998] "Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor", Bulletin of the Seismological Society of America, vol. 88, no. 1, 228–241.

Lee V.W., Trifunac M.D. [2010] "Should average shear-wave velocity in the top 30m of soil be used to describe seismic amplification?" Soil Dyn. Earthquake Eng. 30(11), 1250 – 1258, doi:DOI: 10.1016/j.soildyn.2010.05.007.

Marrara F., Suhadolc P. [2001] "2D modeling of site effect along the EURO-SEISTEST array (Volvi Graben, Greece)". Pure and applied geophysics, vol. 158, p. 2369-23688.

Mucciarelli M., Gallipoli M.R. [2006] "Comparison between Vs30 and other estimates of site amplification in Italy", Proc. 1st European Conference on Earthquake Engineering and Seismology, Cd-Rom edition, paper 270.

NEHRP [2001] Building Seismic Safety Council (BSSC) "National Earthquake Hazards, Reduction Program (NEHRP) Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, Part 1 – Provisions and Part 2 – Commentary", Reports No. FEMA-368 and FEMA-369, prepared by the Building Seismic Safety Council for the Federal Emergency Management Agency, Washington, D.C.

Power M., Chiou B., Abrahamson N., Bozorgnia Y., Shantz T., Roblee C. [2008] "An overview of the NGA Project", Earthquake Spectra, Volume 24, No. 1, pages 3–21, February 2008, DOI: 10.1193/1.2894833.

Raptakis D., Chavez-Garcia F.J., Makra K., Pitilakis K. [2000] "Site effects at Euroseistest—I. Determination of the valley structure

and confrontation of observations with 1D analysis". Soil Dynamics and Earthquake Engineering vol. 19, pp 1–22.

Renalier F., Jongmans D., Savvaidis A., Wathelet M., Endrun B., Cornou C. [2009] "Influence of parameterisation on inversion of surface wave dispersion curves and definition of an inversion strategy for sites with a strong Vs contrast", Geophysics, vol. 75, n°6, 197–209.

Savvaidis A., Ohrnberger M., Wathelet M., Cornou C., Bard P.-Y., Theodoulidis N. [2009] "Variability Analysis of Shallow Shear Wave Velocity Profiles Obtained from Dispersion Curve Inversion considering Multiple Model Parametrizations", SSA meeting, Poster#54, Monterey, USA.

Scherbaum F., Hinzen K.-G., Ohrnberger M. [2003] "Determination of shallow shear wave velocity profiles in the Cologne/Germany area using ambient vibrations" Geophysical Journal International, 152, 597–612.

Theodulidis N., Bard P.-Y. [1995] "Horizontal to vertical spectral ratio and geological conditions- An analysis of strong ground motion data from Greece and Taiwan (smart-1)". Soil dynamics and earthquake engineering, vol.14, n°3, pp. 177-197

Thomson W. T. [1950] "Transmission of elastic waves through a stratified solid medium", J Appl Phys, 1950, 21, 89–93.

Wathelet M. [2008] "An improved neighborhood algorithm: Parameter conditions and dynamic scaling" Geophysical Research Letters, 35, L09301. doi: 10.1029/2008GL033256

Wathelet M., Jongmans D., Ohrnberger M. [2004] "Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements" Near Surface Geophysics, 2:211–221.