# Epistemic Uncertainty *versus* Aleatory Variability in Seismic Response of Soil Profiles

**P. Labbé**
*EDF, Saint-Denis, France*

**SUMMARY**
Considering the seismic response of a conventional homogeneous soil profile, a relationship is established that derives site natural frequency coefficient of variation from shear wave velocity coefficient of variation. Outputs vary dramatically depending on whether this coefficient of variation is regarded as representative of epistemic uncertainty or as representative of spatial variability. In the later case variability of shear wave velocity should be amended before it is considered as an input for computing the soil profile response.

*Keywords: soil profile, epistemic uncertainty, spatial variability, shear wave velocity, natural frequency*

## 1. INTRODUCTION

When dealing with seismic hazard assessment for a dangerous facility site, and especially for a nuclear facility, it is more and more frequent that a site specific probabilistic seismic hazard assessment (site specific PSHA) is implemented. Except for hard rock sites, this approach generally includes a site response analysis. For this purpose site mechanical features are in principle derived from dedicated site investigations. However an issue is that these observed site features include significant uncertainties. This paper discusses how such uncertainties should be treated when computing the site response.

For practical reasons, in most of the cases, the current state of the art considers only the case of vertically propagating shear waves. Concurrently, and consequently, description of the site mechanical features takes usually the simplified form of a soil profile of shear wave velocity, usually designated by $V_s$. A typical output of in-field shear wave velocity investigation is presented in figure 1 (left). It can be observed that $V_s$ fluctuates rapidly with depth. For practical reasons, description of $V_s$ is often limited to its mean (or median) value and standard deviation, both taking the form of slowly fluctuating functions of depth.

On this basis, the current practice considers a series of soil profiles (practically $V_s$ profiles) that are appropriately selected so that site feature uncertainties are reasonably taken into account. These soil profiles take also the form of slowly fluctuating functions. Such a series of $V_s$ profiles is presented in figure 1 (right), derived from observed Vs values presented on the left. This selection of profiles looks very sensible, however we are going to make evident in this paper that it is not appropriate for the purpose of assessing site response uncertainty.

This evidence is based on mathematical analysis of an academic case. The case consists of a soil layer, the $V_s$ mean and standard deviation of which are constant with depth. We are going to examine the coefficient of variation of the first eignenfrequency of this soil layer as a function of the coefficient of variation of $V_s$.

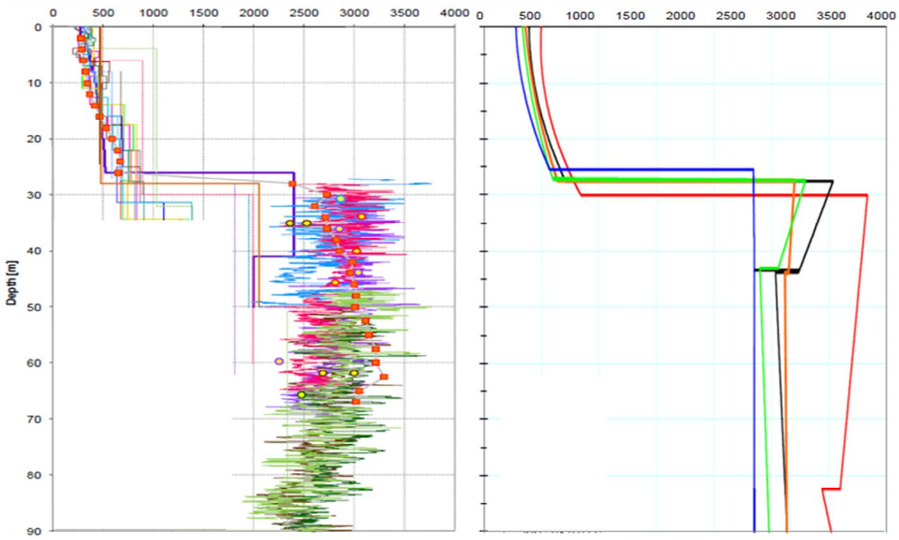


**Figure 1.** Typical in-situ record of shear wave velocity and corresponding derived soil profiles

## 2. ACADEMIC CASE STUDY

The academic case consists of a homogeneous soil profile lying on a rigid bed rock, excited by vertically propagating shear waves. The profile depth is H and the shear wave velocity is c. For the sake of simplicity of some calculations, we introduce also C=1/c. Under these assumptions, the wave propagation duration and the profile first eignenfrequency read respectively:

$$t = H/c = HC\,,\ \text{ and }\ f = c/4H. \tag{1}$$

It is considered in this exercise that there is no uncertainty on H, while c is known only by its mean value $c_m$ and its standard deviation $\sigma_c$. The coefficient of variation (COV) of c is denoted $\delta_c$ (it is reminded that, by definition, $\delta_c=\sigma_c/c_m$.). The purpose of this paper is to derive the COV of t, $\delta_t$ , and/or the COV of f, $\delta_f$ , as functions of $\delta_c$ or $\delta_C$.

There are two possible approaches for dealing with this problem:

Approach A philosophy is that $\delta_c$ is representative of an epistemic uncertainty. In this interpretation, the profile is regarded as actually homogeneous; however its c value is only statistically known. Conceptually, in terms of random fields, we could numerically create as many as desired samples of constant c value profiles, the mean and standard deviation of which would be $c_m$ and $\sigma_c$.

Approach B philosophy is that $\delta_c$ is representative of an aleatory variability. In this interpretation, c is not considered as constant along the profile. The profile is regarded as a sample of a random process, stationary in z (z lies between 0 and H), with a mean value $c_m$ and a standard deviation $\sigma_c$.

Consequences of these two approaches are discussed hereunder. Depending on the adopted approach, $\delta_t$ [$\delta_f$] is designated by $\delta_{tA}$ or $\delta_{tB}$ [$\delta_{fA}$ or $\delta_{fB}$]. Depending on circumstances, calculating $\delta_t$ might be easier than calculating $\delta_f$, or the opposite. Anyway, both t and f are equally valid and legitimate for characterizing the profile response.

## 3. EPISTEMIC UNCERTAINTY

In Approach A, calculations of profile eignenfrequency [propagation duration], mean, $f_{mA}$ [$t_{mA}$], and standard deviation, $\sigma_{fA}$ [$\sigma_{tA}$ ], are immediate, leading to:

$$\delta_{fA}.= \delta_c \text{ and } \delta_{tA}.= \delta_C. \tag{2}$$

The above formula is obtained without any assumption on c (or C) distribution function. In order to establish a direct comparison between $\delta_{fA}$ and $\delta_{tA}$ it is possible to assume, as widely accepted, that c is log-normally distributed, with a $c_0$ median value and a $\beta_c$ dispersion, (it means that Ln(c) is normally distributed with a Ln($c_0$) mean value and a $\beta_c$ standard deviation.). Under this assumption, it is possible to establish that C is also log-normally distributed with $C_0$=1/$c_0$ as median value and $\beta_c$ as dispersion.

We introduce now the non dimensional random variable e, log-normally distributed with 1 as median value and $\beta_c$ as dispersion. This e random variable is so that c=$c_0$ e and C=$C_0$ e.

The assumption of c being log-normally distributed leads then to:

$$\delta_{tA}.= \delta_{fA}.= \delta_C = \delta_c.= \delta_e\ . \tag{3}$$

For information of the reader, it is reminded that $\delta_e = \sqrt{e^{\beta_c^2} - 1}$ .

## 4. ALEATORY VARIABILITY

### 4.1. Description of soil profile variability

Opposite to the previous section, c is now regarded as a function of z. More precisely c is a random process that depends on z. Consequently, C and e are also random processes. (It is not necessary to assume here that c, C and e are log-normally distributed. Hereunder developments are valid without this assumption). C and e are linked by C(z)= $C_0$ e(z). These processes are assumed to be stationary, meaning that their mean and standard deviation are constant versus z. Obviously $\delta_e$ is equal to $\delta_C$.

Mean and standard deviation of e(z) are denoted $e_m$ and $\sigma_e$. We consider the zero mean random process $\varepsilon(z)=e(z)-e_m$ , ($\varepsilon_m=0$ ; $\sigma_\varepsilon=\sigma_e$). By definition, its autocorrelation function reads:

$$\Gamma_\varepsilon(z_1,z_2) = E[\varepsilon(z_1)\ \varepsilon(z_2)] \quad \text{, where E[y] is the expectation of y.} \tag{4}$$

Stationarity implies that $\Gamma_\varepsilon(z_1,z_2)$ does not depend separately on $z_1$ and $z_2$, but only on $\Delta z=|z_1-z_2|$ . We select the classical, simple and widely used following autocorrelation function:

$$\Gamma_\varepsilon(z_1,z_2) = R_\varepsilon(\Delta z) = \sigma_e^2\ e^{-\frac{\Delta z}{h}}\ . \tag{5}$$

In this formula h is a correlation length of mechanical soil properties in the vertical direction. It is useful to consider also the non dimensional correlation length, introduced here as $\alpha$=h/H. Three samples of c/$c_0$ profiles are presented in figure 2, corresponding to three different $\alpha$ values. For the purpose of plotting these samples c was taken as log-normally distributed with a 0.35 $\beta$ value (corresponding to $\delta_c$=0.36).

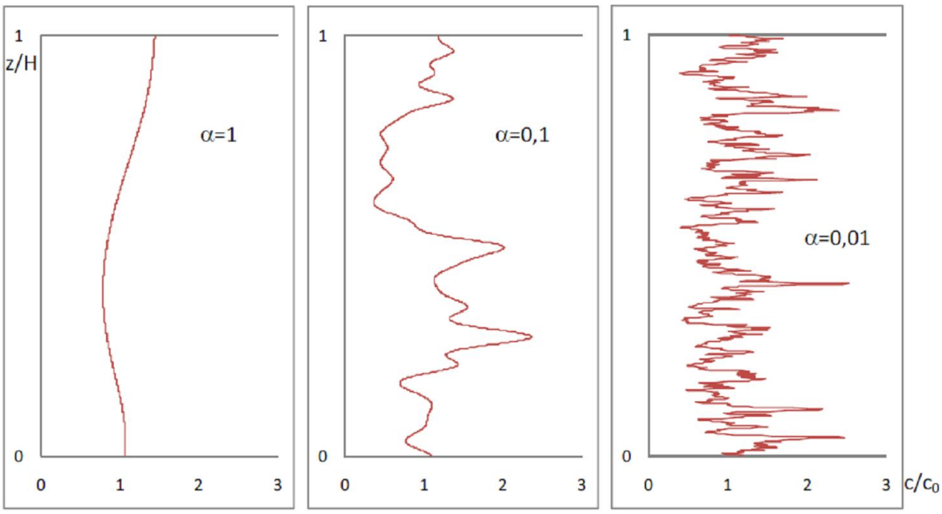


**Figure 2.** Examples of $c/c_0$ fluctuations with depth for 3 different $\alpha$ values

### 4.2. Mean and standard deviation of propagation duration

Duration of shear wave propagation is now a random variable that reads:

$$t = \int_0^H \frac{1}{c(z)} dz \quad \text{, and can be written as:} \tag{6}$$

$$t = t_m + \Delta t \quad \text{with} \quad t_m = C_0 \, H \, e_m \quad \text{and} \quad \Delta t = C_0 \int_0^H \varepsilon(z) dz \, . \tag{7}$$

In the above formula, $t_m$ is the mean value of t, while $\Delta t$ is a zero-mean random variable. Both variables t and $\Delta t$ have a common standard deviation $\sigma_t$ that reads:

$$\sigma_t^2 = C_0^2 \;\; E[(\int_0^H \varepsilon(z)dz)^2] \, . \tag{8}$$

Mathematical tools for calculation of $\sigma_t$ are presented for instance by Vanmarke (1983) and Preumont (1990). It is convenient to put the expectation in the form of a double integral of the autocorrelation function, resulting in:

$$\sigma_t^2 = C_0^2 \int_0^H \int_0^H \Gamma_\varepsilon(z_1, z_2) dz_1 dz_2 \, . \tag{9}$$

Developing further is not possible without selecting a specific autocorrelation function. With the function introduced under Eqn. 4., the following formula is derived:

$$\sigma_t^2 = 2 \, C_0^2 \, H^2 \sigma_e^2 \, (\frac{H}{h} + e^{-\frac{H}{h}} - 1) / (\frac{H}{h})^2 \, . \tag{10}$$

### 4.3. Interprétation

As introduced at the beginning of this paper, the quantity of interest is the coefficient of variation (COV) of t, $\delta_{tB}=\sigma_t/t_m$, to be calculated as a function of $\delta_c$ or $\delta_C$ . This COV is derived from Eqn. 7. for $t_m$ and Eqn. 10. for $\sigma_t$ . It can be expressed in the form of a function of the non dimensional correlation length $\alpha$=h/H as follows:

$$\delta_{tB} = \delta_C\ D(\alpha)\text{, with } D(\alpha) = \sqrt{2\alpha(1+\alpha(e^{-1/\alpha}-1))}\ , \qquad (11)$$

And additionally with $\delta_C = \delta_c$ in case c is log-normally distributed.

It can be observed that

- $D(\alpha)$ approaches 1 as $\alpha$ approaches $\infty$.
- $D(\alpha)$ approaches 0 as $\alpha$ approaches 0, and more precisely $D(\alpha) \approx \sqrt{2\alpha}$ .

These $D(\alpha)$ features can be interpreted as follows: As exemplified in Fig. 2., in case h=H every sample of e(z) fluctuates very slowly. In case h is much larger that H every sample is practically a constant profile. It means that the case under consideration reaches Approach A. This is confirmed by the fact that $\delta_{tB}$ tends to $\delta_{tA}$ .

On the opposite, and also as exemplified in Fig. 2., in case h is much smaller than H, every sample of e(z) fluctuates rapidly. This fast fluctuation creates an averaging effect resulting in the fact that integrations of e(z) result in numbers that are very close together. As $\alpha$ approaches zero, all these integral numbers become closer and closer, resulting in the fact that $\delta_{tB}$ approaches 0.

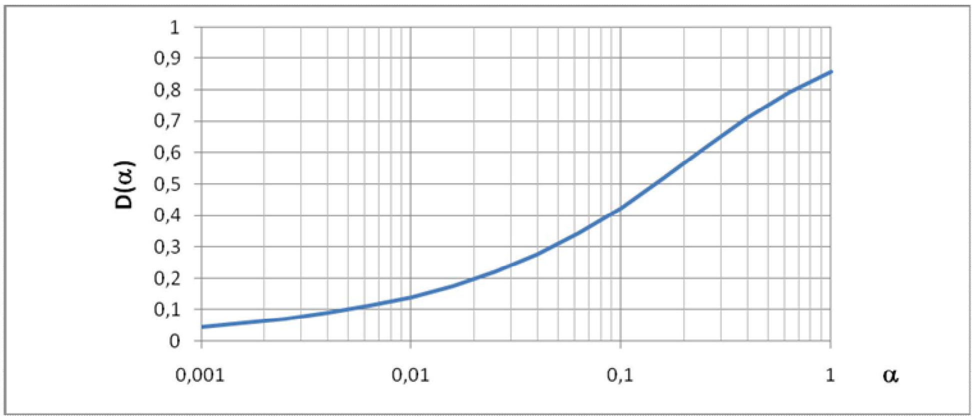


**Figure 3.** *D(α) function*

This type of averaging effect was already investigated and clarified by Toubalem (1996) and by Toubalem *et al.* (1999) when analyzing soil-structure interaction phenomena taking into account fluctuations of soil mechanical characteristics. With this approach, Toubalem (1996) was able to elicit the unexpected in-situ observed azimuth-dependant variability of eigenfrequencies of an axisymetrical concrete building.

Practically, cases with h larger than or even comparable to H do not exist. In orders of magnitude, typical vertical correlation lengths in sedimentary soils are decimetric to metric, while profile depths are decametric to hectometric. Information about vertical correlation length is provided by Antoinet (1995), Phoon and Kulhawy (1999) and Badaoui (2008). Consequently $\alpha$ lies typically between 0.001 and 0.1, generally around 0.01, meaning that in practice $D(\alpha)$ lies between 0.1 and 0.2.

## 5. PRACTICAL CONSEQUENCES FOR SITE SPECIFIC PSHA IMPLEMENTATION

In most cases, site specific probabilistic seismic hazard assessment includes computation of the site response to a seismic input. In this framework, the issue of uncertainty of the site response should be addressed, and consequently uncertainty on the profile mechanical characteristics, principally on shear wave velocity. A point is then that data available from in situ reconnaissance investigations generally take the form of shear wave velocity records that exhibit large variability, rapidly fluctuating with depth.

In order to cope with this uncertainty, a current practice, exemplified in Fig. 1, is to adopt the Approach A when deciding the range of shear wave velocity to be considered in the analysis. Selection of Approach A is not explicitly formulated by the person in charge of establishing soil profiles, which are expected to be representative of the shear wave variability. However, the fact that the question of the possible value of the correlation length is not even asked makes clear that this Approach A is implicitly adopted.

Everybody understands easily that modeling a profile that realistically duplicates the rapidly fluctuating shear wave velocity such as resulting from soil investigations would not be reasonable, if even possible. On the contrary, it seems reasonable that the current practice of slowly fluctuating profiles is maintained. However it should be recognized that this practice introduces a bias in soil profile modeling, and that this bias should be corrected.

For instance, in the academic case presented above, it is understandable that, for $\alpha$=0.01, a shear wave profile such as presented in Fig. 2. is not adopted and replaced by a constant velocity profile. Then the point is that, instead of the observed shear wave variability, represented here by $\delta_c$= 0.36, a reduced variability should be adopted, namely $\delta_c$= 0.36 D($\alpha$) = 0.05.

Finally, when making decision about soil profiles to be considered in view of estimating the impact of soil variability on the response of the site, it should be recognized that the observed soil variability must be significantly reduced, by applying on it the above introduced D($\alpha$) factor. Typically for a 100 m deep soil profile, the in-situ observed variability of shear wave velocity must be divided by a factor between 5 and 10 before it is accepted as an input data for addressing uncertainties in the response on this soil profile.

## REFERENCES


Antoinet E. (1995). Intégration de la variabilité spatiale dans l'estimation probabiliste du tassement et du dévers de grands ouvrages. PhD, Université de Bordeaux I, France.

Badaoui M. (2008). Influence de l'hétérogénéité géologique et mécanique sur la réponse des sols multicouches, PhD, Ecole Nationale Polytechnique, Alger, Algeria, Université Paris-Est Marnes-la-Vallée, France.

Phoon Kok-Kwang, Kulhawy F.H. (1999). Characterization of geotechnical variability, *Canadian Geotechnical Journal* **36**: 612–624.

Preumont A. (1990). Vibrations aléatoires et analyse spectrale, Presses polytechniques et universitaires romandes, Switzerland.

Toubalem F. (1996). Interaction sol-structure en milieu stochastique. PhD, Ecole Centrale de Lyon, France.

Toubalem F, Zeldin B, Thouverez F, Labbé P. (1999) P. Vertical excitation of stochastic soil structure interaction systems. *J. Geotech. Eng., ASCE* **125**: 349-356.

Vanmarke E. (1983). Random fields, analysis and synthesis, MIT Press, Cambridge, Massasuchets, USA.