Covariance Structure between Spectral Accelerations with Different Periods

K.TANAKA[1], M.WANG[2] and T.TAKADA[3]

[1] *Graduate student, T he University of Tokyo, Toyko, Japan*
[2] *Postdoctoral fellow, The University of Waterloo, Waterloo, Canada*
[3] *Professor, The University of Tokyo, Toyko, Japan*
Email: tanaka@load.arch.t.u-tokyo.ac.jp, m29wang@engmail.uwaterloo.ca, takada@load.arch.t.u-tokyo.ac.jp

ABSTRACT :

Ground motion attenuation relationships for spectral acceleration have been developed in Japan. Although these prediction models can usually provide a mean value and a standard deviation for each specified period, they do not characterize covariance property of spectral acceleration between two different periods. This leads to less precise prediction of the response of structural systems such as multi-degree-of-freedom systems and/or nonlinear systems. Some studies have paid attention to this shortcoming and tried to model the correlation of spectral accelerations by statistical methods based on the small number of records. In the present paper, a large amount of ground motion records, which have been observed from recent earthquakes by K-NET and KiK-NET, were used for the covariance analysis. The covariance structure of two spectral values was investigated through principal component analysis (PCA). The covariance characteristics were physically interpreted by examining the effects of the different attenuation relationships, paths and soil conditions.

KEYWORDS: *attenuation relationship, covariance structure, soil condition, principal component analysis, correlation model, generation of time history*

1. INTRODUCTION

Ground motion attenuation relationships for spectral acceleration have been developed in Japan. Although these prediction models can usually provide a mean value and a standard deviation for each specified period, they do not characterize covariance property of spectral acceleration between two different periods. This leads to less precise prediction of response of structural systems such as multi-degree-of-freedom systems and nonlinear systems.

Attenuation relationships characterized in terms of covariance property can make prediction more realistic. For example, response spectra simulated using correlation models can be utilized for a seismic design. Bazzurro and Cornell (2002) reported that the joint knowledge of two or more of ground motion intensity measures could improve the accuracy of structural damage prediction. If two spectral values at different periods are chosen as intensity measures, the correlation coefficient between the two spectral values is required. The authors (2007) already evaluated the joint probability of simultaneous failure of two arbitrary buildings.

Some studies have attempted to model the correlation of spectral accelerations by statistical methods based on the small number of records. Inoue and Cornell (1991) suggested that the correlation coefficient be a linear function in terms of the difference of logarithmic spectral acceleration values at two periods. The suggested model fitted over a period ranging between 0.1sec.and 4.0sec., using 127 records observed from 12 earthquakes. Ishida (1994) proposed a similar correlation model with periods ranging between 0.02 to 5.0 sec., using 104 records from 21 earthquakes observed at four sites. Baker and Cornell (2006) proposed a complicated model with periods ranging between 0.05 and 5.0 sec., using 267 records. Baker and Jayaram (2008) used four NGA (New Generation Attenuation) attenuation relationships and the NGA ground motion database and the proposed correlation model with periods ranging between 0.01 and 10 sec.

The above mentioned correlation models represent the difference of covariance structure only in terms of the coefficient of regression, however it cannot physically, explain the difference of covariance structure. The deviation spectrum which is calculated to subtract the mean attenuation relationships from an observed relationships is also, in general, nonstationary, so correlation coefficient does not always depend on the difference between two periods. In the present paper, the covariance structure of two spectral values will be investigated through the principal component analysis.

Past studies have produced a correlation model based on processing all records observed in different events. But a large amount of ground motion records, which were observed from the recent earthquakes by K-NET and KiK-NET, can be obtained even in a single earthquake event. So this study analyzes the covariance structure for each event, different attenuation relationships and so on. As a result, the covariance characteristics were interpreted on physical basis by examining the effects of the different attenuation relationships, paths and soil conditions.

2. Analysis Procedure for Explanation of Covariance Structure

2.1.Model of Correlation of Spectral Acceleration Values

In this paper, the main objective is the physical explanation of covariance structure. Correlation model is used as the indicator of the difference of covariance structure. Similar to the previous studies, the correlation model was calculated in the following. Calculated are the residuals $\varepsilon_{k,j}(T_i)$ of all N sites for all 25 different periods,

$$\varepsilon_{k,j}(T_i) = \ln(Sa_{k,j}(T_i)) - \ln(\overline{Sa}(T_i, X, Mw, \theta)) \tag{2-1}$$

where k denotes k-th earthquake, j denotes j-th site (1~N), i denotes i-th periods. Sa is the observed spectral acceleration with 5% damping. $\overline{Sa}$ is the attenuation relationships of natural log of spectral acceleration at a specified period, fault distance (X), a moment magnitude (Mw) and other parameters (θ). In this study, 25 discretized periods are between 0.03 and 5.0 sec. with equal interval in log scale. Then the residuals $\varepsilon_{k,j}(T_i)$ corrected by the mean of $\varepsilon_k(T_i)$ at a period T_i, $\mu_{k,\varepsilon(T_i)}$ are computed. The sample mean and variance of $\varepsilon_{k,j}(T_i)$ are

$$\mu_{k,\varepsilon(T_i)} = E[\varepsilon_{k,j}(T_i)] = \frac{1}{N}\sum_{j=1}^{N}\varepsilon_{k,j}(T_i) \tag{2-2}$$

$$\sigma^2{}_{k,\varepsilon(T_i)} = E[(\varepsilon_{k,j}(T_i) - \mu_{k,\varepsilon(T_i)})^2] = \frac{1}{N-1}\sum_{j=1}^{N}[\varepsilon_{k,j}(T_i) - \mu_{k,\varepsilon(T_i)}]^2 \tag{2-3}$$

Next calculated are the covariance, $Cov(T_k, T_l)$, and the correlation coefficient $\rho(T_k, T_l)$ of the corrected residuals at different periods, T_k and T_l, by definition.

$$Cov(T_k, T_l) = \sum_{j=1}^{N}(\varepsilon_{k,j}(T_k) - \mu_{k,\varepsilon(T_k)})(\varepsilon_{k,j}(T_l) - \mu_{k,\varepsilon(T_l)}) \tag{2-4}$$

$$\rho(T_k, T_l) = \frac{\sum_{j=1}^{N}(\varepsilon_{k,j}(T_k) - \mu_{k,\varepsilon(T_k)})(\varepsilon_{k,j}(T_l) - \mu_{k,\varepsilon(T_l)})}{\sqrt{\sum_{j=1}^{N}(\varepsilon_{k,j}(T_k) - \mu_{k,\varepsilon(T_k)})^2 \sum_{j=1}^{N}(\varepsilon_{k,j}(T_l) - \mu_{k,\varepsilon(T_l)})^2}} \tag{2-5}$$

Recently, many models of the correlation of spectral acceleration values have been proposed. Inoue and Cornell (1991) suggested the correlation coefficient as a linear function in terms of the difference of the log Sa values at two periods. This model is simple and fits to the observed data well. However, it was fit over a periods ranging between 0.1 sec. and 4.0 sec. and could not apply to the short period ranging which is shorter than 0.1sec. Ishida (1991) proposed similar correlation model with periods ranging between 0.1sec and 5.0 sec. and constant correlation coefficient at the short period ranging. However, our result indicates that correlation coefficient at the short period ranging depends on the period and is not a constant value. Baker and Jayaram (2008) proposed complicated model with periods ranging between 0.01sec and 10sec. Described above, correlation model is used only as the indicator of the difference of covariance structure in this paper. So, the accuracy of the model is of minor interest. Wang (2007) suggested a linear function which has a different slope in the short period ranging and the periods ranging between 0.1s and 5.0s.This paper used Wang's model as a correlation model in eq.2-6.

$$\rho(T_k, T_l) = 1 - a_1 \cdot \ln(T_{\max}/T_{\min}) + a_2 \cdot I_{0.1} \cdot \ln(0.1/T_{\min}) \tag{2-6}$$

Where $I_{0.1}$ is an indictator equal to 1 if $T_{\min} < 0.1$sec. and equal 0 otherwise, $T_{\max}$ and $T_{\min}$ are the maximum and minimum of T_k and T_l, respectively, a_1 and a_2 is the slope of the regression equation. The larger the value of a_1, the larger the correlation coefficient through all periods ranging.

2.2. Principal Component Analysis (PCA)

The principal component analysis (PCA) is adopted to investigate the effect of the difference of covariance structure. PCA can be used for dimensionality reduction in multivariate data by retaining the characteristics of the data set that contributes most to its variance, by keeping lower-order principal components and ignoring higher-order principal. Lower-order principal components express the characteristics of the data set. In this paper, the dominant variation mode of the deviation from the mean attenuation relationships is examined by PCA. Unlike the model of correlation coefficient, this methodology can facilitate physical interpretation. The principal components were calculated in the following step.
Compute the matrix Φ of the eigenvector ϕ_i which diagonalizes the correlation matrix R in eq.2-7.

$$\Phi^{-1} \mathrm{R} \Phi = \Lambda \tag{2-7}$$

Where Λ is a diagonal matrix. These elements, the singular values λ_i, are the squared roots of the eigenvalues of the correlation matrix R. The value λ_1 is associated with the score on the first principal component, PC$_1$ ϕ_1, for each object and is related to the amount of variance explained by PC$_1$. The contributing rate (Cr_i) which indicates the amount of variance explained by PC$_i$ is defined in the following ,

$$Cr_i = \lambda_i \Big/ \sum_{m=1}^{25} \lambda_m \cdot 100 \quad (\%) \tag{2-8}$$

The tendency of the deviation spectrum in data set was explained by the lower-order principal components. On the other hand, an observed deviation spectrum is expressed by the lower-order principal components. Now, let an observed deviation spectrum **X** be expressed by linear sum of PC$_1$ ϕ_1 and PC$_2$ ϕ_2.

$$[t_1, t_2] = X \cdot [\phi_1, \phi_2] = [\varepsilon_{k,i}(T_1), \varepsilon_{k,i}(T_2), \cdots, \varepsilon_{k,i}(T_{25})] \cdot [\phi_1, \phi_2] \tag{2-9}$$

Where t_1 and t_2 are the coefficients of linear sum of PC$_1$-ϕ_1 and PC$_2$-ϕ_2. And they indicate the amount of PC$_1$'s and PC$_2$'s information which the observed deviation spectrum **X** possesses.

2.3.Record Selection

In this paper, the record of observation is a strong ground motion taken from K-NET and KiK-net. Records were selected based on the following criteria.

1. Fault distance is not larger than 200km.

2. Moment magnitude is not smaller than 5.0.
3. PGA of EW or NS is larger than 10 gal.
4. Vs30 value can be calculated from the soil information of the site.

4410 records from 38 earthquakes are selected, which occurred from October 1996 to July 2007. These records were classified by the Table1. Soil effect of the records was accounted for by the amplification factor proposed by Uchiyama and Midorikawa (2004) and the records were normalized to the value at site class C1.

Table1 Soil classification

site class		Vs30 (m/s)	The number of data
A		Vs30 > 1500	0
B		1500 ≧Vs30＞ 760	175
C	C1	760 ≧Vs30＞ 460	962
	C2	460 ≧Vs30＞ 360	852
D	D1	360 ≧Vs30＞ 250	1349
	D2	250 ≧Vs30＞ 180	736
E		180 ≧ Vs30	336

3. Covariance Structure

3.1. Result of All Data

PCA was adopted to characterize the correlation matrix calculated by all data. 25 eigenvalues and 25 principal components were obtained. Eigenvalues are plotted in Figure 1. The contributing rate of PC_1 and PC_2, Cr_1 and Cr_2, are 53% and 32%, so two principal components can explain about 90% of data set. Eigenvector is illustrated in Figure 2. It shows that the deviation spectrum expresses two dominant variation mode, PC_1 which is parallel shift of the mean attenuation relationship, PC_2 which falls below the mean attenuation relationship in the short period ranging and exceeds in a long-period ranging. The variation mode which vibrates with the short period around attenuation relationship, like PC_{10}, is diminishing.

The covariance structure can be estimated by Cr_1 and Cr_2. If a data set is fully explained by PC_1, the correlation coefficient becomes 1 for all periods ranging. If a data which is explained by PC_2 is added to the data set, the correlation coefficient becomes smaller than 1, depending on Cr_2.

Coefficient a_1 which was obtained by the regression of all data is illustrated in Figure 3. Coefficients a_1 and a_2 are 0.306 and 0.308. There is large dispersion from the regression equation. It shows that the correlation model is not adequate for detailed interpretation about the difference of the covariance structure.

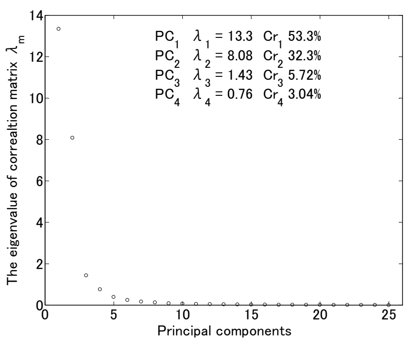

Figure 1 Eigenvalue λ_1 of correlation matrix in all data

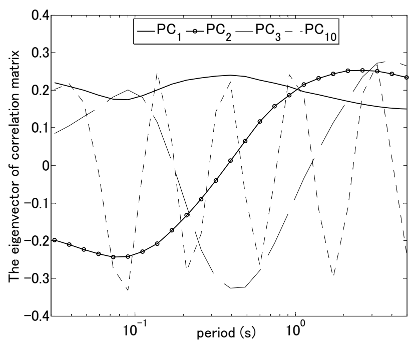

Figure 2 Eigenvector of correlation matrix (PC_1) in all data

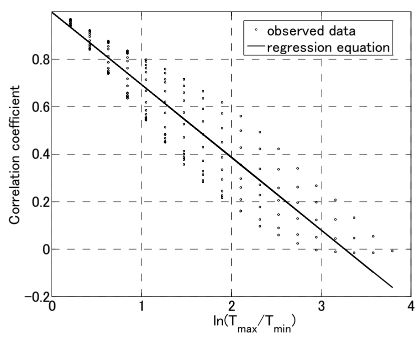

Figure 3 Regression of correlation coefficient in all data

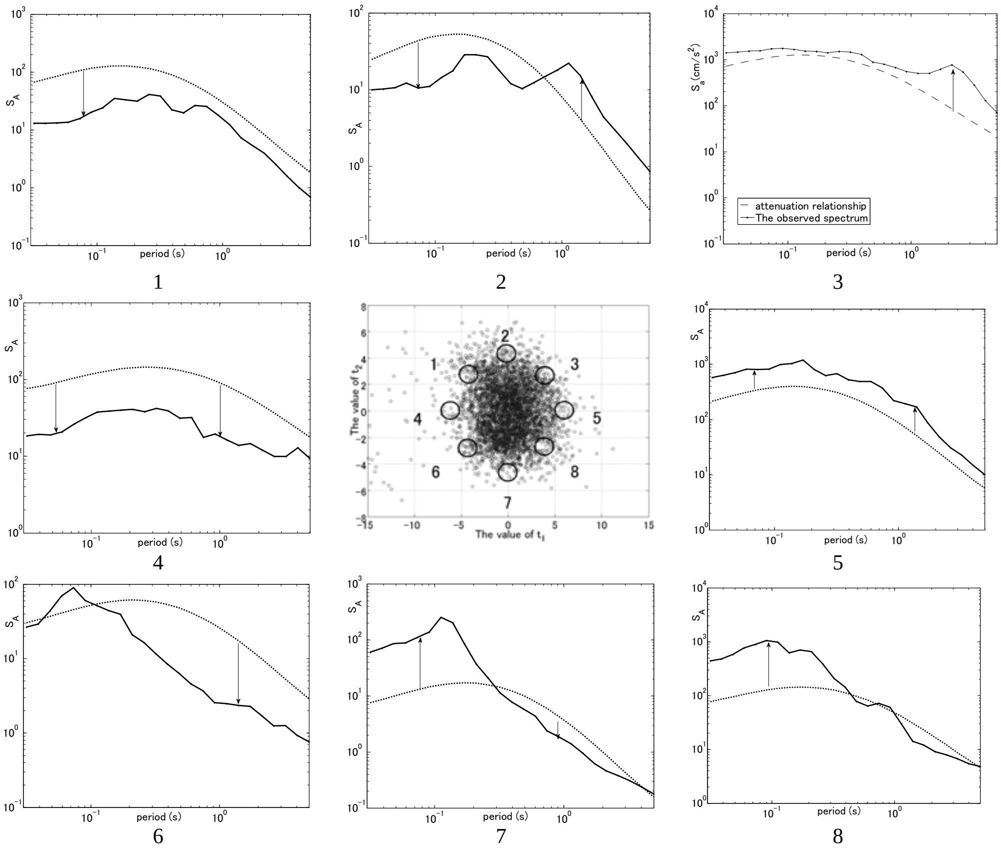

Figure 4 The plot of t_1 and t_2 in all data

t_1 and t_2 calculated by the PC_1 and PC_2 of all data are plotted in Figure 4. The spectrum around the scatter plot is a typical observed spectrum which is included in the circle of the scatter plot. The shape of these spectra is coincident with PC_1 or PC_2 or a superposition of PC_1 and PC_2. For example, the spectrum at upper right is coincident with a superposition of PC_1 and PC_2, so the deviation at a long period is larger than it at the short period.

3.2.Dependency of Each Attenuation Relationship

The deviation spectrum from the mean attenuation relationship is expected to depend on the prediction model. In this section, the Uchiyama and Midorikawa model (2006) and the Annaka model (1997) were used to calculate the covariance matrix. The regression coefficients, a_1 and a_2, of the correlation matrix in each earthquake are illustrated in Figures 5 and 6. The value a_2 does not show a distinctive trend and varies widely from one earthquake to the next. Figure 5 shows that the coefficients a_1 of the Uchiyama model was smaller than that of the Annaka model. The eigenvalue of PC_1, λ_1, in each earthquake is illustrated in Figure 7. It shows that λ_1 of the Uchiyama model was larger than that of the Annaka model. Described in 3.1, the larger the value of λ_1, the more the amount of information explained by PC_1. So the correlation coefficient of the Uchiyama model was larger than the Annaka one through all the ranging. And it is consistent that the coefficient a_1 of the Uchiyama model was smaller than that of the Annaka model. The variation of PC_1 was a parallel shift of mean of attenuation relationship. Since the Uchiyama model is improved by introducing the focal depth term and estimating the amplification factor of subsurface ground, the shape of the Uchiyama model is fit over the data set well. In this paper, the Uchiyama model was selected as an attenuation model.

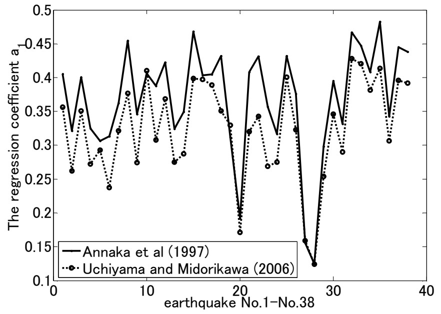

Figure 5 Plot of a_1 in each earthquake

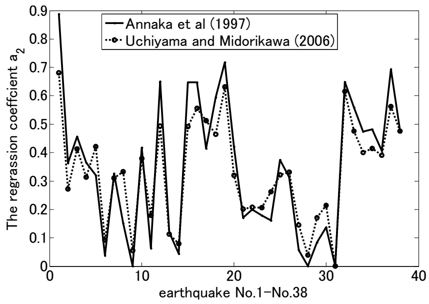

Figure 6 Plot of a_2 in each earthquake

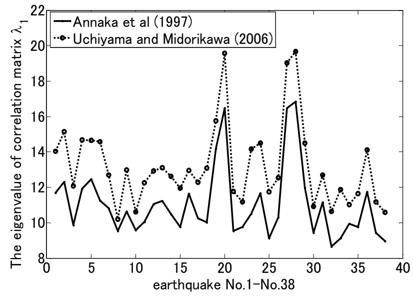

Figure 7 Plot of λ_1 in each earthquake

3.3. Effect of Soil Condition

As described in 2.1, the soil effect of the data was accounted for by introducing the amplification factor and the observed spectra transformed the spectra at the engineering bedrock. However, the soil effect which could not be eliminated by this simple method is examined in this section. The regression coefficient in each soil class is illustrated in Figure 8. The smaller the value of Vs30, the larger the value of a_1. Briefly, the softer the soil, the smaller the correlation coefficient through all period. It is consistent with Inoue (1997). PC_1 in each soil class is illustrated In Figure 9, The softer the soil, the larger the deviation at the short period ranging. The harder the soil, the larger the deviation at a long period ranging. However, Figure 10 which illustrates PC_2 for each soil class shows an opposite phenomenon.
To consider this result explicitly, t_1 and t_2 in class B and E are plotted in Figures 11 and 12. Figure 11 shows that the plot is biased on the lower right. Figure 12 shows that the plot is biased on the upper left. Judging from Figure 4, class B and class E are included the spectra illustrated in Figure 13. The spectra in Class B have a tendency to exceed the mean attenuation at the short period and Class C has a tendency to fall below. For the amplification factor at the short period was not evaluated adequately, Spectral value on the engineering bedrock at the short period is underestimation for soft soil and overestimation for hard soil.

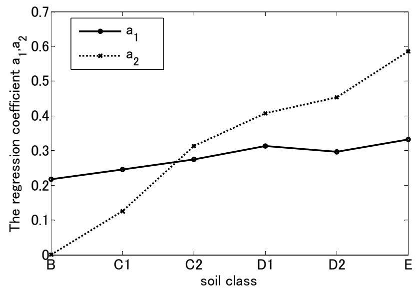

Figure 8 Plot of a_1 and a_2 for each soil class

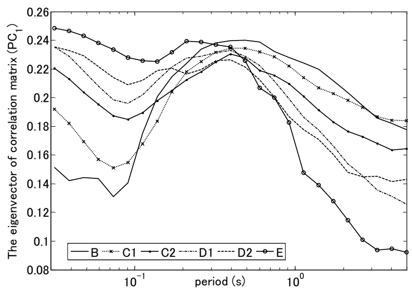

Figure 9 Eigenvalue (PC_1) for each soil class

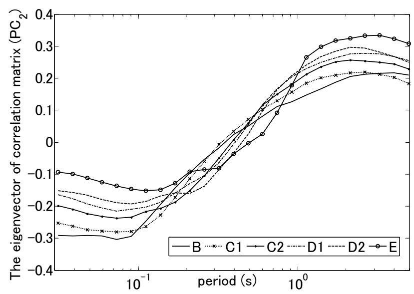

Figure 10 Eigenvalue (PC_2) for each soil class

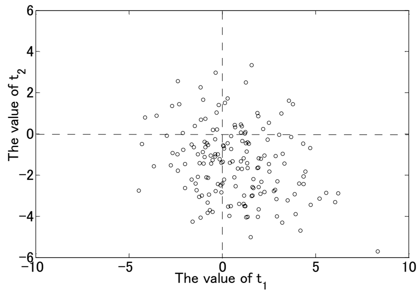

Figure 11 Plot of t_1 and t_2 for soil class B

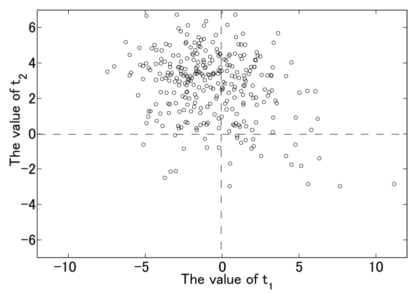

Figure 12 Plot of t_1 and t_2 for soil class E

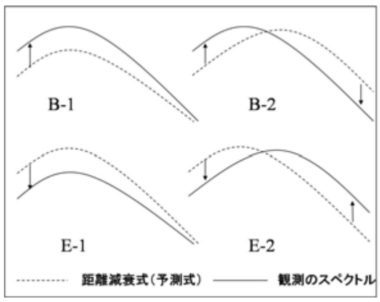

Figure 13 Shape of spectra for class B and class E

3.3. Path effect

3.3.1 Depending of fault distance

In this section, the difference of the covariance structure by the fault distance is investigated. The regression coefficient a_1 in each fault distance is illustrated in Figure 14. The regression coefficient a_1 in the fault distance between from 0 to 50 km (Class 0-50) is smaller than that in other fault distance. It is coincident that the value of Cr_1 in class 0-50 is over 70% (class 50-100 over 57%, class 100-150 over 45%, class 150-200 over 50%). The shape of PC_1 in each distance class is illustrated in Figure 15. The shape of PC_1 in class 0-50 is almost a constant value. It means that the spectrum which is a parallel shift of the mean attenuation relationship is well existed in class0-50. However, the longer the path to the site, the more complicated the path. So Figure15 shows that the short period ranging of PC_1 is relatively larger than the long period ranging in the class 100-150 and class 150-200. Additionally, Cr_1 value in these classes decreased as compared with class 0-50. It means that the spectra in long fault distance include various spectra which cannot be explained by the mean attenuation relationship.

3.3.2 Depending of Focal Depth

The difference of the covariance structure by the focal depth is similar to fault distance one. The smaller the focal depth, the smaller the path to the site. So the figure 16 shows that the shape of PC_1 in depth between 0 and 30 km is almost a constant value. t_1 and t_2 in over 90 km is plotted in Figure 17. Figure 17 shows that the plot is biased on the lower right and the upper left. Judging from Figure 4, it means that the deviation from the mean attenuation is large in the short period ranging.

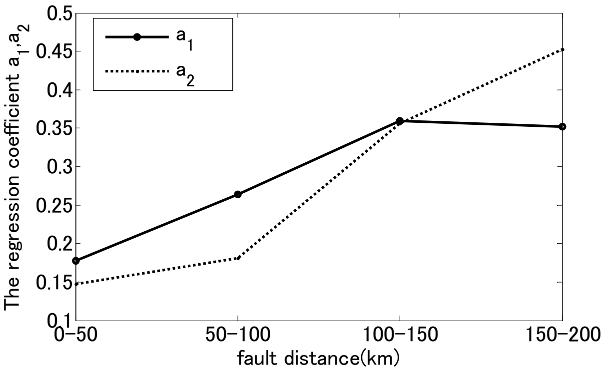

Figure 14 Plot of a_1 and a_2 in each fault distance class

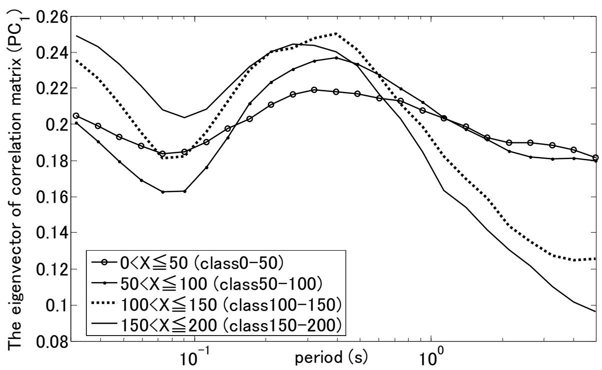

Figure 15 Eigenvalue (PC_1) in each fault distance class

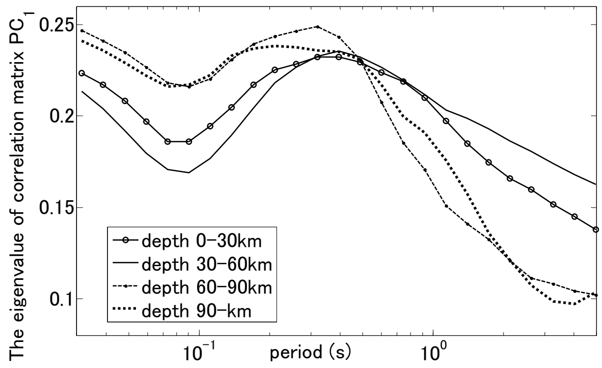

Figure 16 Eigenvalue (PC_1) of focal depth

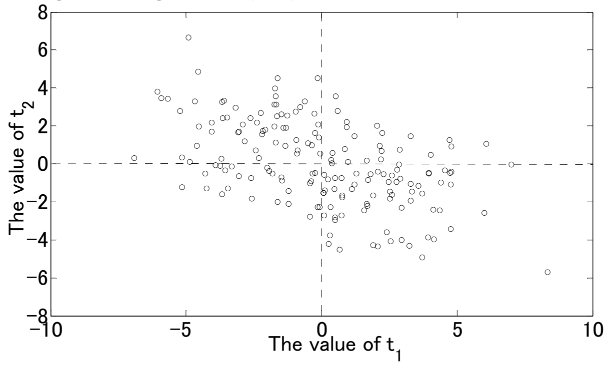

Figure 17 Plot of t_1 and t_2 in depth over 100 km

4. Conclusions

In this paper, the covariance structure of two spectral values was investigated through principal component analysis. The covariance characteristics were interpreted by examining the effects of the different paths and soil conditions. The conclusions can be drawn in the following.

1. The correlation coefficient model developed by Wang (2007) is adopted in this study. The coefficients of the same model (a_1 = 0.306 and a_2 = 0.308) are found out by re-regressing the model with the present data set. There is very little difference from the previous studies. Two principal components govern the dominant variation mode of the deviation spectrum.

2. The first two principal components govern the dominant variation mode of the deviation spectrum.
3. The deviation spectrum from the mean attenuation relationships is nonstationary, correlation coefficient doesn’t always depend on the difference of two periods.
4. The prediction accuracy of two attenuation relationship was examined quantitatively. The Uchiyama and Midorikawa (2006) model fits over this data set well.
5. Soil effect of the records was accounted for by the amplification factor proposed by Uchiyama and Midorikawa (2004). However, the amplification factor at the short period was not evaluated adequately, Spectral value on the engineering bedrock at the short period is underestimation for the soft soil and overestimation for the hard soil
6. The deviation spectrum from the mean attenuation relationships at the short period is large. Because the longer the fault distance and the focal depth, the more complicated the path to the site.

ACKNOWLEDGEMENT

This study is a part of a research program on the establishment and application of probabilistic inter-related fragility assessment methodology for urban structures against earthquakes. The support provided by the MEXT (Ministry of Education, Culture, Sport, Science and technology) under Grant Scientific Research A 17201034 is also gratefully acknowledged. The ground motion data used in this study are provided by K-NET and KiK-NET.

REFERENCES

Sakamoto, S., Uchiyama, Y., and Midorikawa, S. (2006). Study on the evaluation variance of the attenuation relationship for response spectra. *12th Japan Association for Earthquake Engineering, 362-365.*

Baker, J. W., and Cornell, C. A. (2006). Correlation of response spectral values for multicomponent ground motion. *Bulletin of the seismological Society of America* **96:1**, 215-227.

Bazzuro, P., and Cornell, C. A. (2002). Vector-valued probabilistic seismic hazard analysis (VPSHA). *Proceedings of the 7th U.S. National Conference on Earthquake Engineering*, 1CD-ROM, Earthquake Engineering Research Institute, Boston, MA, July 21-25.

Tanaka, K., Wang, M.,and Takada, T. (2007). Simultaneous collapse of buildings with correlated response. *Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan*, **B-1**, 9-10.

Ishida, H. (1993). Probabilistic evaluation of earthquake response spectrum and its application to response analysis,*Proceedings of ICOSSAR’93*, **3**, 2131~2138.

Inoue, T., and Cornell, C. A. (1991). Seismic hazard analysis of multi-degree-of freedom structures, *in proceedings of the 2nd Japan Conference on Structural Safety and reliability*, 105-112.

Inoue, T., (1997). Correlation of response spectral values. *Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan*, **B-1**, 89-90.

Baker, J. W. and Jayaram, N. (2008). Correlation of spectral acceleration values from NGA ground motion models, *Earthquake Spectra (in press)*, http://www.stanford.edu/~bakerjw/publications.html (accessed 2008-06-04).

Wang, M. (2007). *New Perspectives for Probabilistic Prediction of Seismic Ground Motion*, PhD dissertation, Department of Architecture, the University of Tokyo, 433-493.

Annaka, T., Yamazaki, F., and Katahira, F., (1997). Proposal of attenuation equations for peak ground motion and response spectrum using JMA-87 type accelerogram, *in Proceedings of the 24th JSCE Earthquake Engineering Symposium*,161-164.

Uchiyama, Y., and Midorikawa, S. (2006). Attenuation relationship for response spectra on engineering bedrock considering effects on focal depth, *Journal of Structural and Construction Engineering, AIJ*, **606**, 81-88.

Uchiyama, Y., and Midorikawa, S. (2003). Evaluation of Amplification factor of site classes based on strong motion records and nonlinear response analysis, *Journal of Structural and Construction Engineering, AIJ*, **571**, 87-93.

Uchiyama, Y., and Midorikawa, S. (2003). Practical method to evaluation response spectra of bedrock motions using amplification factors for site classes, *Journal of Structural and Construction Engineering, AIJ*, **582**, 39-46

Hino, M. (1977). *Spectral analysis*, Asakura Publishing Co., Ltd, 97-101.