Probability distributions of spectra and spectral ratios

H.Mogi
Saitama University, Urawa, Japan

H.Kawakami
Saitama University, Urawa, Japan

M.R.Ghayamghamian
Saitama University, Urawa, Japan

ABSTRACT : Distributions of the power spectral ratios of microtremors have been analytically investigated on the assumption that the microtremor can be expressed as a stationary Gaussian random process. The statistical aspects of the power spectral ratios are revealed. Then, these results are compared with those of the observed microtremors, and the statistical results are shown to be valid for both the soil/rock ratios and the QT spectra.

1. INTRODUCTION

From a statistical point of view, since a microtremor can be expressed as a stationary Gaussian random process, the estimated spectrum is also a random variable. Therefore, statistical attentions are required to estimate the power spectrum, and regarding to this problem several methods and criteria have been investigated (e.g.,Hino 1988).

In recent years, the ground characteristics are frequently investigated by the microtremor measurements. In these investigations, two kinds of spectral ratios are utilized. One is the ratio of the spectra measured at soil and rock sites, and the other is the ratio of the horizontal to vertical spectrum, named QT spectrum (Nakamura 1989).

In this study, we have analytically investigated the statistical characteristics of the spectra and the spectral ratios, and compared with those of the observed microtremors.

2. STOCHASTIC FORMULATION OF POWER SPECTRA

2.1 χ^2 *distribution of the power spectra*

Microtremors are idealized as stationary Gaussian random processes as follows:

$$v(t) = \sum_{j=1}^{\infty} A_j \cos(\omega_j t) + B_j \sin(\omega_j t) \tag{1}$$

where $v(t)$ is the time history record like a particle velocity and ω_j is the j-th angular frequency. The coefficients of the trigonometric functions are Gaussian random variables whose expectations are zero (Hoshiya 1971, Morikawa and Kameda 1991). The variances of these coefficients are related to the expectation of the power spectrum of the j-th frequency component as follows:

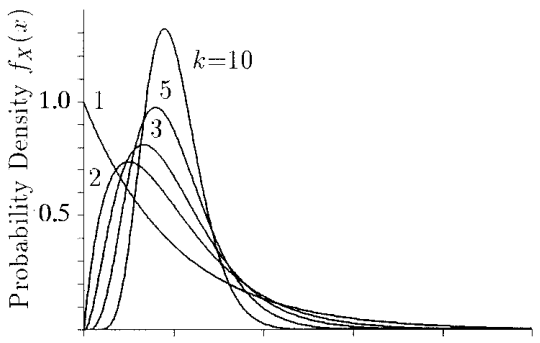

Fig.1 Probability density function of the power spectrum, $f_X(x)$ (μ=1, k=1,2,3,5,10).

$$\mathrm{E}\left[\frac{A_j^2 + B_j^2}{2}\right] = s_j \Delta\omega \tag{2}$$

where s_j is the expectation of the one-sided power spectrum and E indicates an ensemble average.

The estimated power spectrum X that is an average of k samples

$$X\,\Delta\omega = \mathop{\mathrm{E}}_{k}\left[\frac{A_j^2 + B_j^2}{2}\right] \tag{3}$$

is a statistical variable that has the χ^2 distribution with $2k$ degrees of freedom. The expectation of the distribution μ is equal to s_j, and from Eq.3, the probability density function of the power spectrum is given by

$$f_X(x) = \frac{k^k}{\mu^k \Gamma(k)} x^{k-1} \exp(-kx/\mu) \tag{4}$$

Table 1. Percentiles of the power spectrum (μ=1).

Number of Samples	Percentiles of the power spectrum(X)						
k	5%	10%	20%	50%	80%	90%	95%
1	0.051293	0.105361	0.223144	0.693147	1.609438	2.302585	2.995732
2	0.177681	0.265906	0.412194	0.839173	1.497154	1.944860	2.371932
3	0.272564	0.367355	0.511681	0.891353	1.426343	1.774107	2.098598
5	0.394030	0.486518	0.617908	0.934182	1.344196	1.598718	1.830704
7	0.469331	0.556395	0.676238	0.952805	1.296484	1.504582	1.691771
10	0.542541	0.622130	0.728922	0.966871	1.251875	1.420599	1.570522
12	0.577018	0.652445	0.752575	0.972364	1.231388	1.383177	1.517293
15	0.616422	0.686641	0.778804	0.977868	1.208340	1.341867	1.459099
18	0.646350	0.712314	0.798193	0.981544	1.191078	1.311449	1.416624
20	0.662733	0.726263	0.808624	0.983384	1.181713	1.295126	1.393962
236	0.895392	0.917538	0.944844	0.998588	1.054333	1.084277	1.109425

The variance of the distribution expressed in Eq.4 is given by

$$\sigma^2 = \mu^2/k \tag{5}$$

Fig.1 shows the probability density functions $f_X(x)$ in Eq.4 with numbers of samples k being 1,2,3,5,10. Table 1 shows 5,10,20,50,80,90,95 percentiles of the power spectrum in Eq.4.

2.2 *Distribution of the weighted average of the power spectra*

In analyses of microtremors' records, the power spectrum is often smoothed by a digital filter. In this case, the spectrum is averaged over neighboring frequencies. Although this operation causes a problem of resolution, in this study, we have investigated the probability density function and the variance of the spectrum.

If we assume that the number of samples is k=1 and the expectation is μ=1, Eq.4 can be transformed into

$$f_X(x) = \exp(-x) \tag{6}$$

Assuming that the distributions of the random variables X_1, X_2 have the probability density functions given by Eq.6, the probability density function of the sum $Z = w_1X_1 + w_2X_2$ is expressed as

$$f_Z(z) = \frac{1}{w_1}\int_0^{z/w_2} f_X\left(\frac{z-w_2\xi}{w_1}\right) f_X(\xi)d\xi \tag{7}$$

$$= \frac{\exp(-z/w_1) - \exp(-z/w_2)}{w_1 - w_2} \tag{8}$$

Applying this operation consecutively, we get the probability density function of the sum

$$Z = \sum_{i=1}^{I} w_i X_i \tag{9}$$

as follows:

$$f_Z(z) = \sum_{i=1}^{I} \frac{w_i^{I-2}\exp(-z/w_i)}{\prod_{\substack{j=1\\ j\neq i}}^{I}(w_i - w_j)} \tag{10}$$

The expectation and variance of this distribution are expressed as

$$\mu_Z = \sum_{i=1}^{I} w_i \tag{11}$$

$$\sigma_Z^2 = \sum_{i=1}^{I} w_i^2 \tag{12}$$

The basic concept of the digital filter is averaging the neighboring frequency components with certain weight factors. This operation also reduces the variance, like an ensemble average of different samples.

The digital filter that is in the simplest form and used frequently has three weight factors (w_{-1}, w_0, w_1). From its symmetry w_{-1}=w_1, we can express a smoothed power spectrum,

$$Z_i^{(N)} = \sum_{j=-N}^{N} {}_NW_j \; Z_{i+j}^{(0)} \tag{13}$$

$${}_NW_j = \sum_{m=0}^{M} {}_N\mathrm{C}_{N-j-2m} \cdot {}_{j+2m}\mathrm{C}_m \cdot w_0^{N-j-2m} \cdot w_1^{j+2m} \tag{14}$$

$$M = \begin{cases} \frac{N-j}{2} & N-j : \text{Even} \\ \frac{N-j-1}{2} & N-j : \text{Odd} \end{cases} \tag{15}$$

where $Z_i^{(N)}$ is the i-th frequency component of the smoothed spectrum after N times filtering operation, and ${}_NW_j$ is the weight factor for the $i \pm j$ frequency component. Table 2 shows the weight factors for each component and the variance in the case of the Hanning filter (0.25, 0.5, 0.25).

As the number of samples k does not appear in Eq.10, it is difficult to compare the expectation and variance of $Z^{(N)}$ in Eq.10 with those derived from Eq.4. Therefore, based on Eq.5, the equiva-

Table 2. Weight factors for the Hanning filter and the variance of the spectrum (μ=1).

Times(N)	${}_NW_0$	${}_NW_1$	${}_NW_2$	${}_NW_3$	${}_NW_4$	${}_NW_5$	${}_NW_6$	${}_NW_7$	${}_NW_8$	${}_NW_9$	${}_NW_{10}$	σ^2
1	.500+0	.250+0										.375000+0
2	.375+0	.250+0	.625-1									.273438+0
3	.313+0	.234+0	.938-1	.156-1				$\pm n$ indicates $\times 10^{\pm n}$				.225586+0
4	.273+0	.219+0	.109+0	.313-1	.391-2							.196381+0
5	.246+0	.205+0	.117+0	.439-1	.977-2	.977-0						.176197+0
6	.226+0	.193+0	.121+0	.537-1	.161-1	.293-2	.244-3					.161180+0
7	.209+0	.183+0	.122+0	.611-1	.222-1	.555-2	.854-3	.610-4				.149446+0
8	.196+0	.175+0	.122+0	.667-1	.278-1	.854-2	.183-2	.244-3	.153-4			.139950+0
9	.185+0	.167+0	.121+0	.708-1	.327-1	.117-1	.311-2	.584-3	.687-4	.381-5		.132061+0
10	.176+0	.160+0	.120+0	.739-1	.370-1	.148-1	.462-2	.109-2	.181-3	.191-4	.954-6	.125371+0

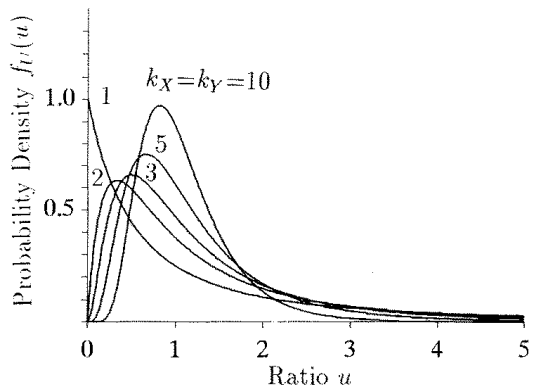

Fig.2 Probability density function of the spectral ratio, $f_U(u)$, in Eq.17 (μ_X=μ_Y,k_X=k_Y=1,2,3,5,10).

lent number of samples k_e is defined as

$$k_e = \frac{\mu_Z^2}{\sigma_Z^2} \tag{16}$$

where μ_Z is given by Eq.11 and σ_Z^2 by Eq.12.

3. STOCHASTIC FORMULATION OF POWER SPECTRAL RATIOS

3.1 *Distribution of the ratio of the ensemble averaged power spectra*

As the distributions of the probabilistic variables X, Y are expressed in Eq.4, the ratio $U = Y/X$ can be shown to have the following F distribution:

$$\begin{aligned} f_U(u) &= \int_0^\infty f_X(\frac{\xi}{u})\, f_Y(\xi)\, \frac{\xi}{u^2}\, d\xi \\ &= \frac{\Gamma(k_X+k_Y)\, k_X^{k_X}\, k_Y^{k_Y}\, \mu_X^{k_Y}\, \mu_Y^{k_X}}{\Gamma(k_X)\,\Gamma(k_Y)} \\ &\quad \cdot \frac{u^{k_Y-1}}{(k_X\mu_Y + k_Y\mu_X u)^{k_X+k_Y}} \end{aligned} \tag{17}$$

where the subscripts X and Y indicate parameters for each power spectrum. Fig.2 shows the probability density function $f_U(u)$ of the ratio U in Eq.17 in the case of k_X=k_Y=1,2,3,5,10, μ_X=μ_Y=1. From Eq.17, the expectation(μ_U), variances(σ_U^2 and $\sigma_{Y/X}^2$) and the cumulative distribution function ($F_U(u)$) are expressed as follows:

$$\mu_U = \frac{\mu_Y}{\mu_X} \cdot \frac{k_X}{k_X - 1} \tag{18}$$

$$\sigma_U^2 = \frac{\mu_Y^2}{\mu_X^2} \cdot \frac{k_X^2(k_X+k_Y-1)}{k_Y(k_X-1)^2(k_X-2)}$$

Variance around μ_U (19)

$$\sigma_{Y/X}^2 = \frac{\mu_Y^2}{\mu_X^2} \cdot \frac{k_X^2 + k_Xk_Y + 2k_Y}{k_Y(k_X-1)(k_X-2)}$$

Variance around μ_Y/μ_X (20)

$$\begin{aligned} F_U(u) &= \frac{\Gamma(k_X+k_Y)}{\Gamma(k_X)\,\Gamma(k_Y)} \cdot \sum_{r=0}^{k_Y-1} (-1)^r \frac{{}_{k_Y-1}\mathrm{C}_r}{k_X+r} \\ &\quad \cdot \left\{ 1 - \left(1 + \frac{k_Y\mu_X}{k_X\mu_Y} u \right)^{-k_X-r} \right\} \end{aligned} \tag{21}$$

Table 3 shows percentiles estimated numerically from Eq.21.

From Eqs.18-21 and Table 3, we can point out the followings:

1) The expectation of the spectral ratio is larger than the the ratio of the expected values(μ_Y/μ_X).

2) The expectation does not depend on the number of samples of the numerator spectrum(k_Y).

3) If the number of samples is unity, the expectation is infinity.

4) The variance is infinity when the number of samples k_X is 1 or 2.

5) The coefficient of variation(σ_U/μ_U) is larger than that of the power spectrum (see Eqs.5,12,19). Namely, the power spectral ratios are less stable than the power spectra.

6) The expectation is always larger than μ_Y/μ_X as described in 1), but the medium is equal to μ_Y/μ_X in the case of k_X=k_Y as shown in Table 3 and Fig.3(except Case F).

3.2 *Monte Calro Simulations of power spectral ratios*

In the case of the weighted average of the power spectra, the probability density of their ratio is complicated. The Monte Carlo simulations have been conducted to estimate the expectation, vari-

Table 3. Percentiles of the spectral ratio ($\mu_X = \mu_Y = 1$, $k_X = k_Y$).

Number of Samples	Percentiles						
$k_X = k_Y$	5%	10%	20%	50%	80%	90%	95%
2	0.156538	0.243472	0.402801	1.000000	2.482613	4.107250	6.388233
3	0.233434	0.327380	0.484997	1.000000	2.061869	3.054551	4.283866
4	0.290858	0.386197	0.538683	1.000000	1.856379	2.589349	3.438101
5	0.335769	0.430551	0.577502	1.000000	1.731595	2.322604	2.978237
6	0.372213	0.465671	0.607368	1.000000	1.646447	2.147437	2.686637
7	0.402621	0.494454	0.631328	1.000000	1.583963	2.022434	2.483726
8	0.428544	0.518651	0.651137	1.000000	1.535775	1.928079	2.333484
9	0.451020	0.539397	0.667892	1.000000	1.497247	1.853923	2.217197
10	0.470775	0.557462	0.682320	1.000000	1.465588	1.793843	2.124155
12	0.504093	0.587594	0.706063	1.000000	1.416303	1.701854	1.983760

Table 4. Simulations of the spectral ratios; the weighted average values of the spectrum are used.

Case	Hanning Window		μ_U	σ_U^2	k_e	
	X(times)	Y(times)			X	Y
A	1	1	1.543	3.678	2.667	2.667
B	2	2	1.346	1.498	3.657	3.657
C	3	3	1.334	1.353	4.433	4.433
D	5	5	1.203	0.648	5.675	5.675
E	7	7	1.176	0.502	6.691	6.691
F	9	1	1.158	0.780	7.572	2.667

ance and percentiles.

The Monte Carlo simulation have been executed using the pseudo Gaussian random variables; their expectation and variance are 0 and 1, respectively. A random variable that is obtained by the summation of the two squared Gaussian random variables corresponds to a raw power spectrum. We have taken the ensemble average or the weighted average of these random variables. At each trial, we have simulated two average values and calculated the ratio. This procedure has been repeated 10,000 times in each case(Table 4).

Figs.3 and 4 show the expectation and variances, respectively. In these figures, squares and circles indicate the weighted and ensemble averages, respectively.

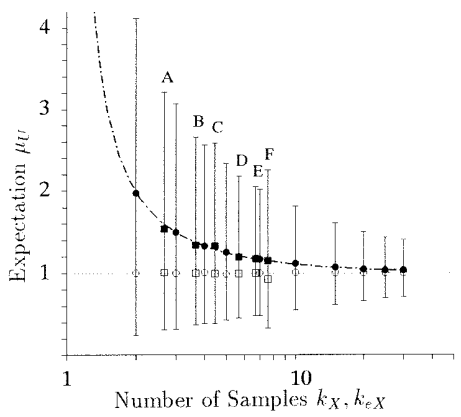

Fig.3 Expectation and 10,50,90 percentails of the spectral ratios estimated by the Monte Carlo technique.

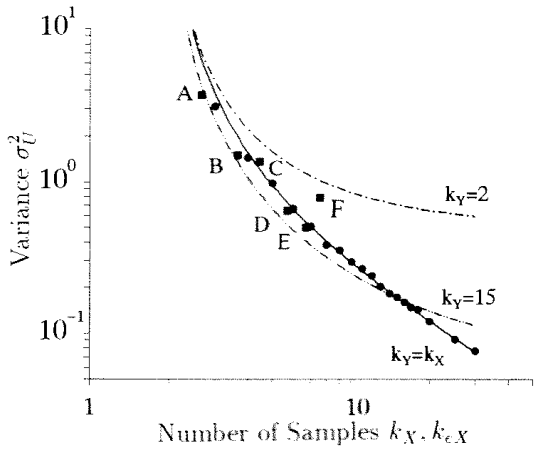

Fig.4 Variance of the spectral ratios estimated by the Monte Carlo technique.

In Fig.3, the one-dotted broken line shows the theoretical expectation(Eq.18), and thin vertical lines show the range of 10-90 percentiles estimated by the Monte Carlo simulations. In this figure, the mediums are also shown by the open circles and squares.

In Fig.4, solid, one-dotted and two-dotted broken lines show the theoretical variance in the case of k_Y=k_X, k_Y=2, k_Y=15, respectively. In Figs.3 and 4, the simulated expectations of the weighted average (Cases A to E) are plotted near/on the theoretical line. From this, we can confirm that the definition of the equivalent number of samples k_e (Eq.16) is also valid for the spectral ratios.

4. MICROTREMOR MEASUREMENTS

4.1 *Instruments and soil structure*

The microtremor measurements were made in the athletic field of Saitama University from 3:30 to 8:45 a.m. on February 16th, 1995. Instruments for measurements are velocity meters. The natural period is 5 sec, and the characteristic of gain is in flat shape from 0.3 to 50Hz. Sampling rate is 100Hz and the total precision is about 4.6μkine/bit.

The soil structure is composed of a very soft (N<10) surface layer of thickness about 35m and

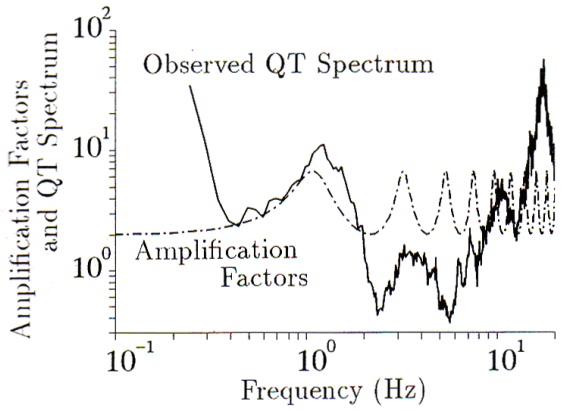

Fig.5 Theoretical amplification factors and the observed QT spectrum.

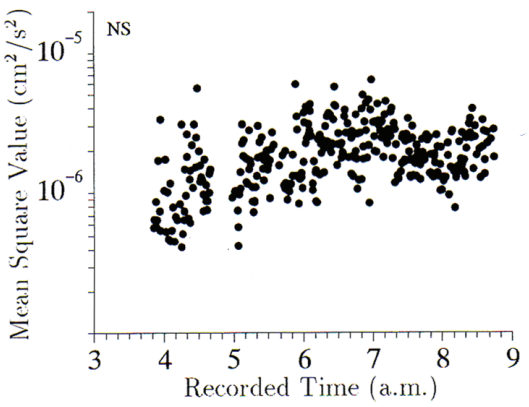

Fig.6 Distributions of the mean square values of the microtremor records.

a hard (N>50) diluvial deposit; their contrast is very clear.

Assuming that the S wave velocity of the surface layer is 150m/sec and that of the basement is 500m/sec, the amplification factors are computed as shown by the broken line in Fig.5. This figure also shows the spectral ratio (NS/UD) whose numbers of samples (k_X and k_Y) are 236. In this figure, we can confirm that the natural frequency of SH waves is 1.2Hz and the Rayleigh waves seem to predominate at 3.0Hz.

4.2 *Measurements and analyses*

Three components(UD,NS,EW) of microtremors were measured at one station for 5.25 hours. In this period, we recorded 332 samples each of whose length is 21sec.

After Fourier transform using FFT, we have calculated the raw power spectra of each record.

Fig.6 shows the mean square values of the time history against recorded time. As shown in Fig.6, the total energies of the microtremors have a tendency to increase with time. For this reason, we used the samples recorded from 5:30 to 8:45 a.m. The total number of samples in this period is 236.

In sections 3 and 4, we have made two assumptions: 1) the raw power spectrum is expressed by

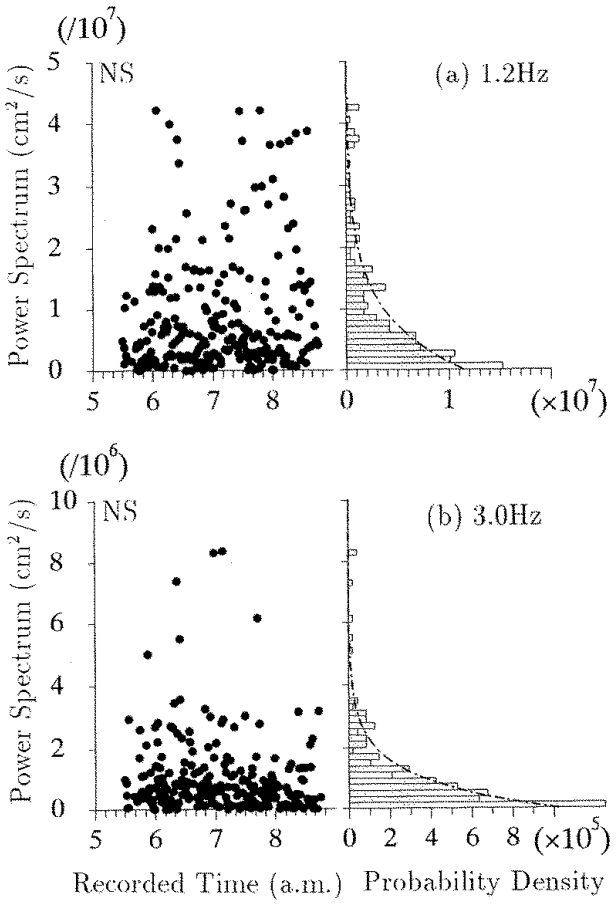

Fig.7 Distributions and the probability density function of the observed power spectra.

the χ^2 distribution; 2) the components are independent from each other.

Figs.7(a)(b) show the distributions of the raw power spectra of the (a)1.2Hz and (b)3.0Hz components. The theoretical probability density functions have been drawn by the broken lines by substituting the observed average into μ in Eq.4. In these figures, each record is found to be distributed randomly, and the probability density function seems to be approximately equal to the χ^2 distribution. However, the variance estimated from the observed records is larger than the theoretical variance $\sigma^2=\mu^2$ by two or three times. This causes a problem in the following discussion on the spectral ratios.

Since the above assumption 2) is acceptable for the soil/rock ratios, we had examined the relation between the horizontal and vertical components observed at the same station and time. Figs.8(a)(b) show the relation between the UD and NS components of the raw power spectra at (a)1.2Hz and (b)3.0Hz. At these frequencies, the correlation coefficients are 0.162 and 0.150, therefore, we can confirm independence between UD and NS components. This means that the formulations in section 3 are also valid for the QT spectrum.

Fig.9(a) shows the ensemble average of the 236 records of the NS component and 10 and 90 percentiles in the case of k =1,10,236. Fig.9(b) shows

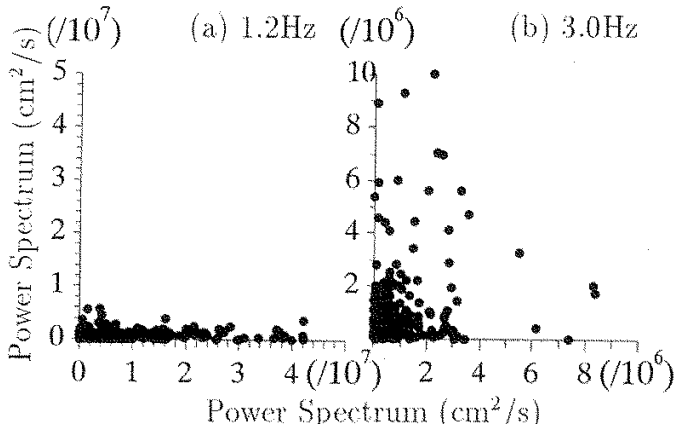

Fig.8 Relation between NS and UD components of the observed power spectra.

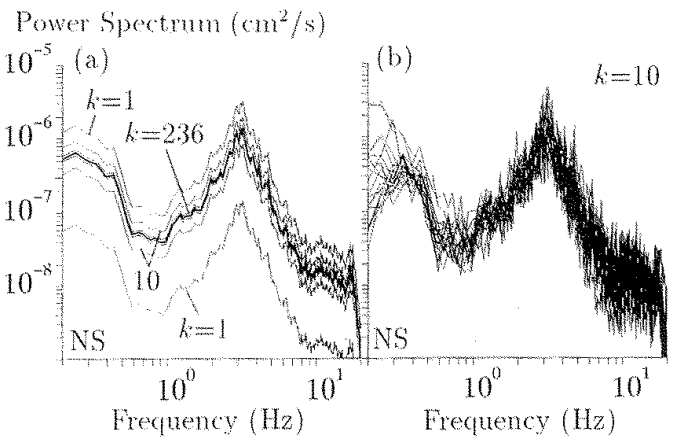

Fig.9 Observed power spectra and 10-90% ranges.

20 samples of the power spectra(k=10), and their variance is in accordance with the theoretical range shown in Fig.9(a).

Figs.10(a)(b) show the comparison of variance between (a)the theoretical range estimated from the observed power spectral ratio(k=236) and (b)observed power spectral ratios(k=10). From this comparison, it can be noticed that the observed ratios have larger variance than the theoretical ones. This tendency is slightly recognized also in Fig.7; the variances of the observed power spectra are larger than those of the theoretical ones. This tendency seems to be magnified by taking the ratio.

5. SUMMARY

We have analytically investigated the probability density functions of the power spectral ratios and their statistical features, and have shown the followings:

1) The expectation, variance and percentiles of the power spectral ratio are formulated.

2) The expectation of the power spectral ratio is larger than the ratio of the expected values.

3) The coefficient of variation of the power spectral ratio is larger than that of the power spectrum.

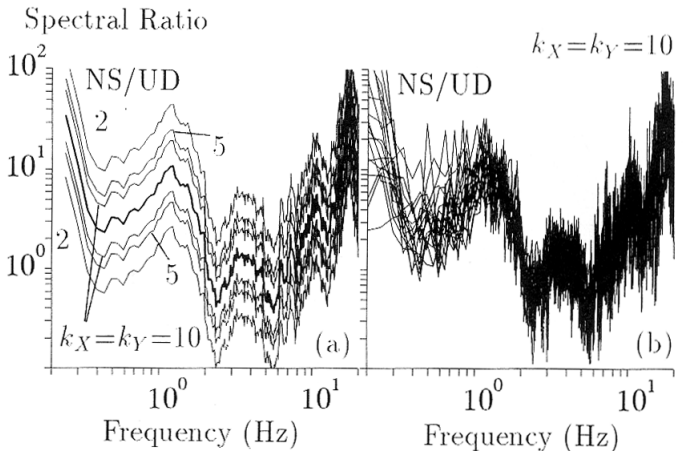

Fig.10 Observed power spectral ratios and 10-90% ranges.

4) These statistical characteristics are remarkable for small numbers of samples.

Next, we have examined the assumptions that we have made in the theoretical investigations using the observed microtremors, and have shown the followings.

5) The observed power spectra and spectral ratios have similar but a little larger variance than the theoretical ones.

6) The vertical and the horizontal components are independent.

In recent years, the ground characteristics have been frequently investigated based on the microtremor measurements. Whenever we refer to these investigations, we have to pay attention to the conditions of measurements and analyses like the number of samples, because this directly affects the spectral ratio.

ACKNOWLEDGMENTS

The authors thank to Professor T.Ohmachi and Associate Professor T.Toshinawa of Tokyo Institute of Technology for their cooperation in microtremor measurements.

REFERENCES

Hino,M.1988. *Spectral Analysis*. Tokyo: Asakura-shoten (in Japanese).

Hoshiya,M. & H.W.Tieleman 1971. Two stochastic models for simulation of correlated random processes. *VPI-E*, 71, 9.

Morikawa,H. & H.Kameda 1991. Simulation of conditional random fields involving given time functions, *UEHR Report*, E8.

Nakamura,Y.1989. A Method for dynamic characteristics estimation of subsurface using micro tremor on the ground surface, *QR of RTRI*, 30, 1: 25-33.

Yoda,H.1966 *Statistics for engineers*. Tokyo: Hobunkan-shuppan (in Japanese).