Bull. Sugadaira Montane Res. Cen. No. 8, 109-118 pp. (1987)
(Sugadaira Montane Research Center, University of Tsukuba)

A simulated Experiment for Sampling Soil Microarthropods to Reduce Sample Size

HIROSHI TAMURA

Department of Biology, Faculty of Science, Ibaraki University, Mito 310, Japan

(Accepted Dec. 26, 1986)

Synopsis

An experiment was conducted to examine a possibility of reducing the necessary sample size in a quantitative survey on soil microarthropods, using soybeans instead of animals. An artificially provided, intensely aggregated distribution pattern of soybeans was easily transformed to the random pattern by stirring the substrate, which is soil in a large cardboard box. This enabled the necessary sample size to be greatly reduced without sacrificing the statistical reliability. A new practical method of sampling soil microarthropods in fields was proposed to reliably estimate the abundance of animals using a smaller sample size.

Key words: Microarthropods, soybean, sample size, aggregated distribution, random distribution, stirring

I. Introduction

Soil microarthrods are known to have a patchy distribution in soil, showing various degrees of aggregation and micro-distribution, the degree of aggregation and pattern of distribution varying with the species (TAKEDA, 1973, '79). In making a quantitative survey of soil microarthropods, a number of sample units of soil are usually taken to make a sample, the sample units being taken by using tin samplers of a given size, e. g. 5x4 cm^2 in area and 5 cm deep (KIKUZAWA et al., 1976). Then, as a result of the strong aggregation of animals, a large number of sample units (large sample) are required to avoid the risk of sample units being accidentally concentrated in either dense or sparse concentrations of animals (ELLIOTT, 1977). Using the equation proposed by IWAO and KUNO (1971), the number of sample units necessary for making a statistically reliable estimate (necessary sample size) for collembolans was calculated to be 30 to 50 when using a 5x4x5 cm^3 sampler in a red-pine forest (TAKEDA, 1973) and 40 to 60 when using a 5x5x5 cm^3 sampler in a beech forest (TAMURA, 1978).

It is, however, usually impossible to deal adequately with such large numbers of sample units, especially when the samples are taken frequently to study seasonal fluctuations of the

---

This paper was presented to the VIth Sugadaira Seminar on Entomology held at the Sugadaira Montane Research Center, University of Tsukuba, on December 14-15, 1985.

animals synchronously in different sites. Therefore a compromise must be usually made by arbitrarily taking 5 or 10 units at the sacrifice of statistical reliability (e. g. TAMURA & MIHARA, 1977).

If we could reduce the clumped nature of the distribution of the animals, we could reduce the number of sample units necessary to achieve statistical reliability, and thus greatly reduce the amount of labor needed to achieve usable results. The present study reports a laboratory experiment, using soybeans instead of animals, aimed at transforming a clumped distribution into a random distribution.

II. Method

A topless 95x95 cm^2 and 10 cm high box was prepared of thick cardboard. Soil was sieved evenly into the box to a height of 5 cm to make a substrate. A thousand small soybeans were spread on the substrate in an intensely aggregated distribution pattern.

Before any further treatment, 361 tin cans (5x5 cm^2 in area and 5 cm high) without both top and bottom were reqularly arranged in a square pattern (19 reqularly spaced cans in both rows and columns) on the soil and pushed to the bottom of the box. Soil within each can was scooped into a beaker in order to count soybeans in it, and the number of beans was recorded. After that the soil and soybeans were returned together into the can. In this

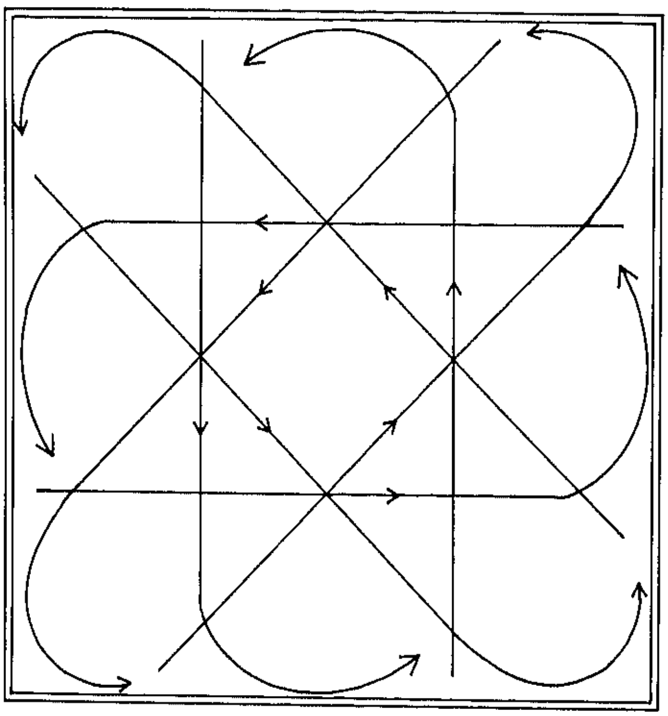

Fig. 1 Schematic representation of the way to stir substrate. One stirring consists of eight strokes of hand.

manner data were obtained for all the 361 sample units of soil. Then all the cans were removed.

Next, the substrate containing soybeans was stirred by hand in the pattern shown in Fig. 1. After the soil surfece was leveled, the 361 tin cans were again put in the substrate and the soybeans in it were counted as described above.

The process described above was repeated with 2, 5, 10, 20 and 50 stirrings. All the data obtained were entered in a computer file and used to calculate mean, variance, necessary sample size and mean crowding (LLOYD, 1967).

The symbols to be used in text and their explanations were listed in Table 1.

Table 1. Symbols to be used in text and their explanations

Symbol	Explanation
N	Total number of soil blocks (sample units, here 361)
x	Number of soybeans found in a soil bloch, which is taken by using a cubic sampler of $5\times5\times5$ cm^3 in volume
m	Population mean (total number of beans used/N)
σ^2	Population variance
$\bar{x}$	Sample mean in terms of beans per soil block sampled ($\Sigma x/n$, n being number of unit soils sampled)
s	Standard deviation in a sample
ns	Necessary sample size (number of sample units necessary for keeping a certain level of statistical reliability)
e	Value of tolerable error, which is here expressed as $0.1\bar{x}$, $0.2\bar{x}$ or $0.3\bar{x}$
ns (e)	For example, as ns ($0.1\bar{x}$) meaning necessary sample size at tolerable error of 10 % of $\bar{x}$
t	Student's t, which defines the statistical reliability in the estimation of ns
d	$e/\bar{x}$
$\overset{*}{m}$	Mean crowding (LLOYD, 1967)
α, β	Indices in m-$\overset{*}{m}$ relationship (IWAO, 1968)
St-T	Tth stirring, e. g. St-10 indicating the time of 10 stirrings

III. Results

Before stirring, at first, population mean (m) and population variance (σ^2) were calculated, and then a patchiness index in terms of σ^2/m was obtained to use as a measure of aggregation. I next randomly selected 10 of the 361 date and used these to compute sample mean ($\bar{x}$), standard deviation (s), mean crowding ($\overset{*}{m}$) (LLOYD, 1967), and necessary sample size (ns) (COCHRAN, 1950 ; IWAO and KUNO, 1971) at tolerable errors (e) of 10, 20 and 30 % of $\bar{x}$ ($0.1\bar{x}$, $0.2\bar{x}$, and $0.3\bar{x}$, respectively) using the equation

(1) $ns=(t\cdot s)^2/e^2$ (COCHRAN, 1950)

,where t is Student's t, here with 9 degrees of freedom (df) at a 0.05 probability level. This process was repeated 100 times and the mean and standard deviation of 100 nss were

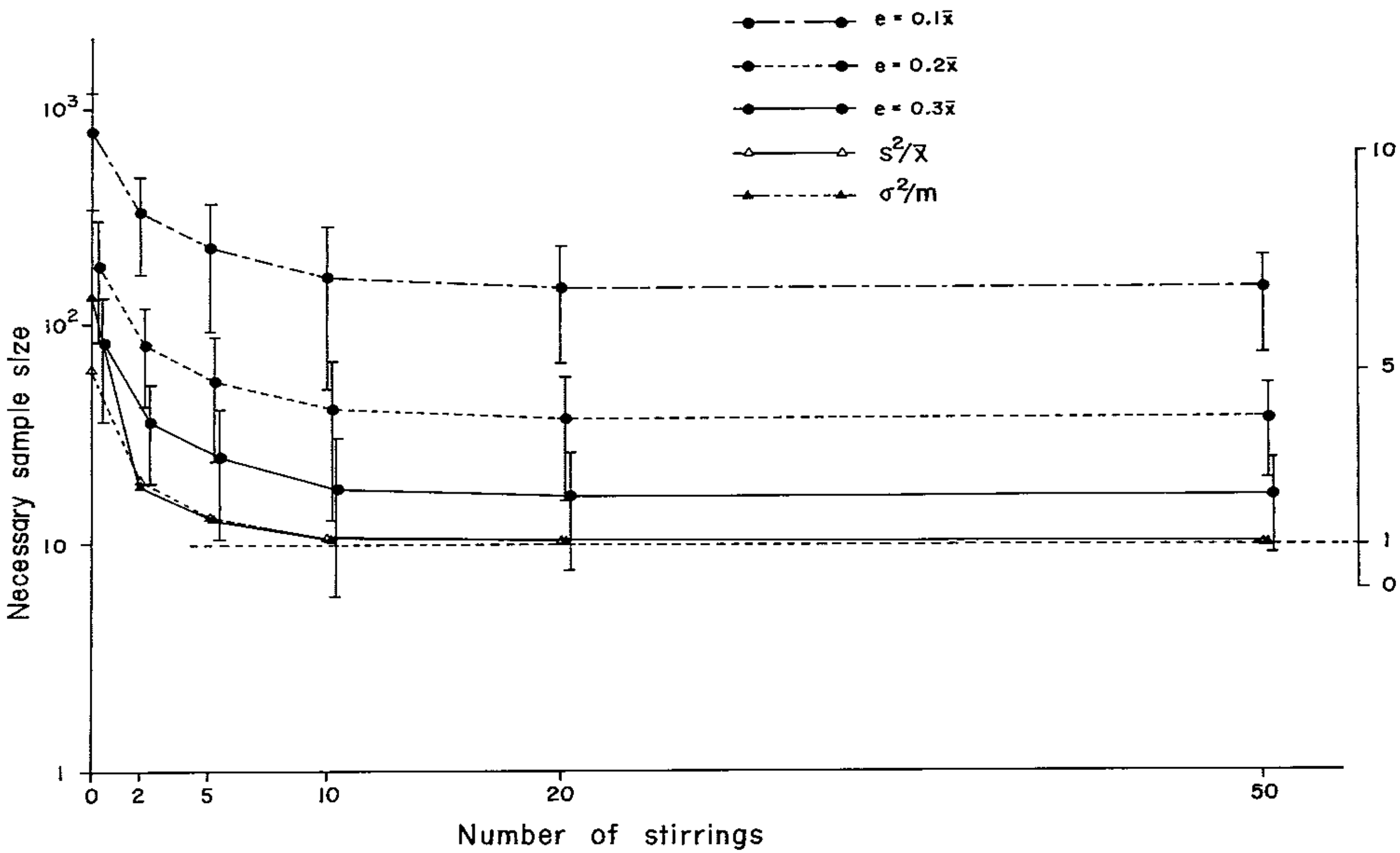

Fig. 2 Changes of necessary sample sizes (statistically estimated number of sample units to be taken) at three different allowable error levels based on COCHRAN (1950) and patchness index in terms of variance to mean ratio (σ^2/m, $S^2/\bar{x}$) on number of stirrings. Symbols and bars indicate mean values and 95 % conficence intervals, respectively.

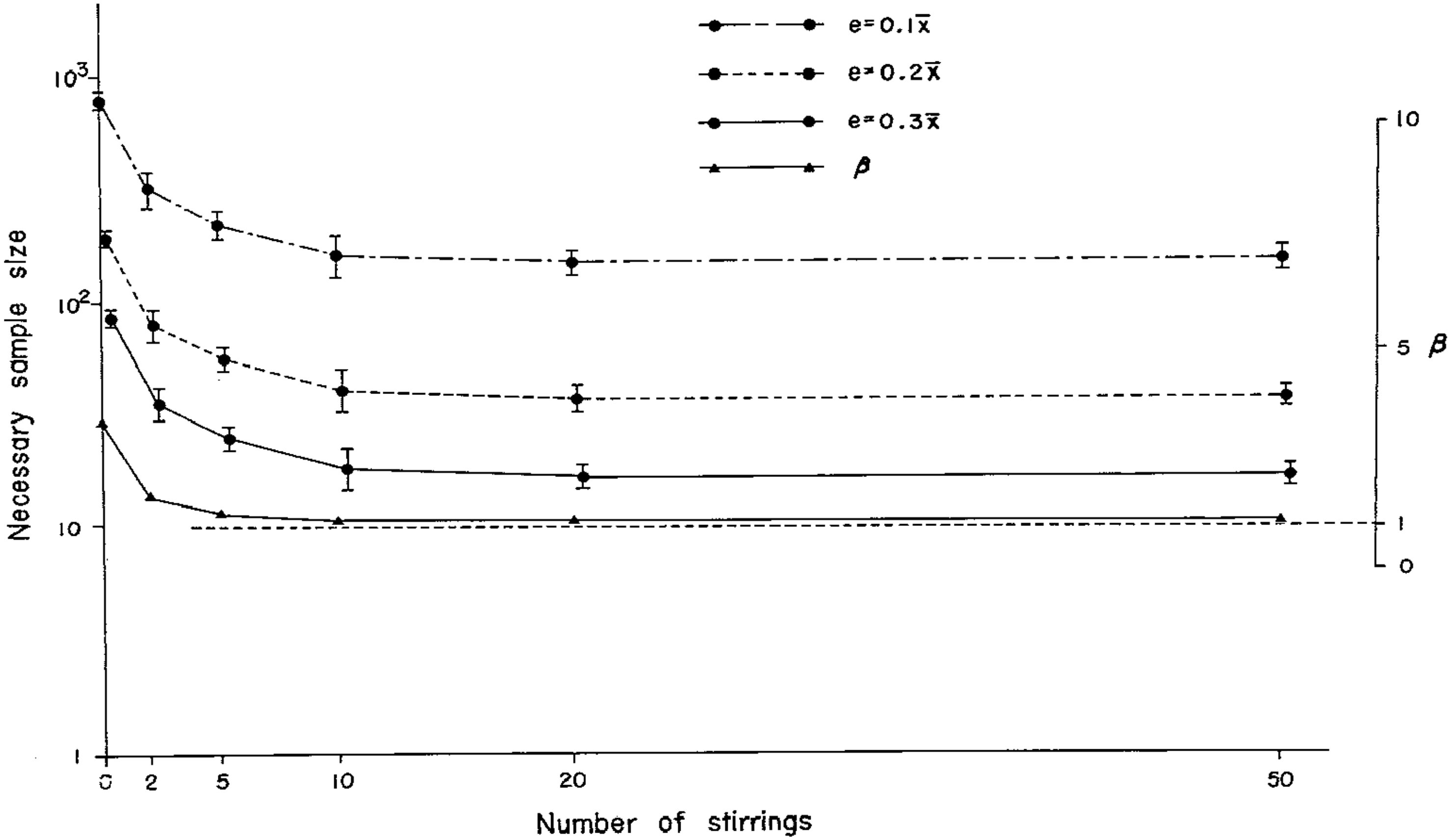

Fig. 3 Changes of necessary sample sizes based on IWAO and KUNO (1971) and patchness index in terms of β in m-$\overset{*}{m}$ relationship (IWAO, 1968) on number of stirrings. Symbols and bars indicate mean value and 95 % confidence intervals, respectively.

calculated. Another measure for the necessary sample size was also computed using the equation

(2) $ns=(t/d)^2 \cdot [(\alpha+1)/\bar{x}+(\beta-1)]$ (IWAO and KUNO, 1971)

,where d is $e/\bar{x}$, and α and β are indices in the m-$\overset{*}{m}$ relationship in IWAO (1968).

These computations were performed for every different number of stirrings and the results were shown in Figs. 2 and 3. Figures 2 and 3 represent families of curves showing the change in sample size (in terms of numbers of soil sample units used) necessary to achieve the different tolerable levels of error under different numbers of stirrings starting with 0 (St-0) and ending with 50 stirrings (St-50).

Examination of figure 2 shows that the number of soil sample units needed to acheive the tolerable level of error drops off sharply with the first five stirsings, more slowly until 10 stirrings and remains relatively stable thereafter. This is clear evidence that the number of sample units needed to achieve a 0.05 probability level can be strikingly reduced by stirring the substrate.

The changes in the patchiness indices, σ^2/m and $s^2/\bar{x}$, were different before stirring, but became almost identical at St-2 and were nearly 1 at St-10 and after. A patchiness index of 1 suggests random distribution of soybeans. To ascertain this, χ^2 test for goodness-of -fit was in turn performed between data at each stirring time and Poisson series, which is the mathematical model expected of a random pattern. The tests were not significant at St-10, St-20 and St-50 ($P>0.9$ or $P>0.5$), but were highly significant at St-0, St-2 and St -5 ($P<0.01$ in all). This indicates that the aggregated pattern of the soybeans could be changed into the random distribution pattern by stirring the substrate.

This change was further investigated through the m-$\overset{*}{m}$ relationship (IWAO, 1968 : see formula 2) (Fig. 3). The sample sizes necessary to achieve a 0.05 probability level at tolerable errors of $0.1\bar{x}$, $0.2\bar{x}$ and $0.3\bar{x}$ reduced quickly with increased stirrings and became stable after St-10 as in Fig. 2. A different patchiness index, β (see IWAO, 1968), also neared unitly at St-5 and kept the unity from St-5 to St-50. Unity here also suggests a random distribution pattern (IWAO, 1968).

These trends show that an aggregated distribution can be easily transformed to the random distribution and the number of sample units necessary to achieve a given level of reliability can be remarkably reduced by stirring the substrate.

When the distribution is random, the variance can be equated to the mean, viz. $\bar{x}=s^2$. Therefore, the equation for neccessary sample size can be simplified as

(3) $ns=(t/d)^2 \cdot (1/\bar{x})$.

Since d is determined by $\bar{x}$, the neccessary sample size is simply a function of $\bar{x}$. The result is that the neccessary sample size linearly declines with the increase of the mean of the individuals per sample unit on double logarithmic scales (Fig. 4). In the case illustrated, t is 2.26 with df=9 at 0.05 probability. For example, when the mean number of individuals per sample unit is 10, then necessary sample sizes at the three different tolarable errors are as follows ; $ns(0.1\bar{x})=52$, $ns(0.2\bar{x})=13$ and $ns(0.3\bar{x})=6$.

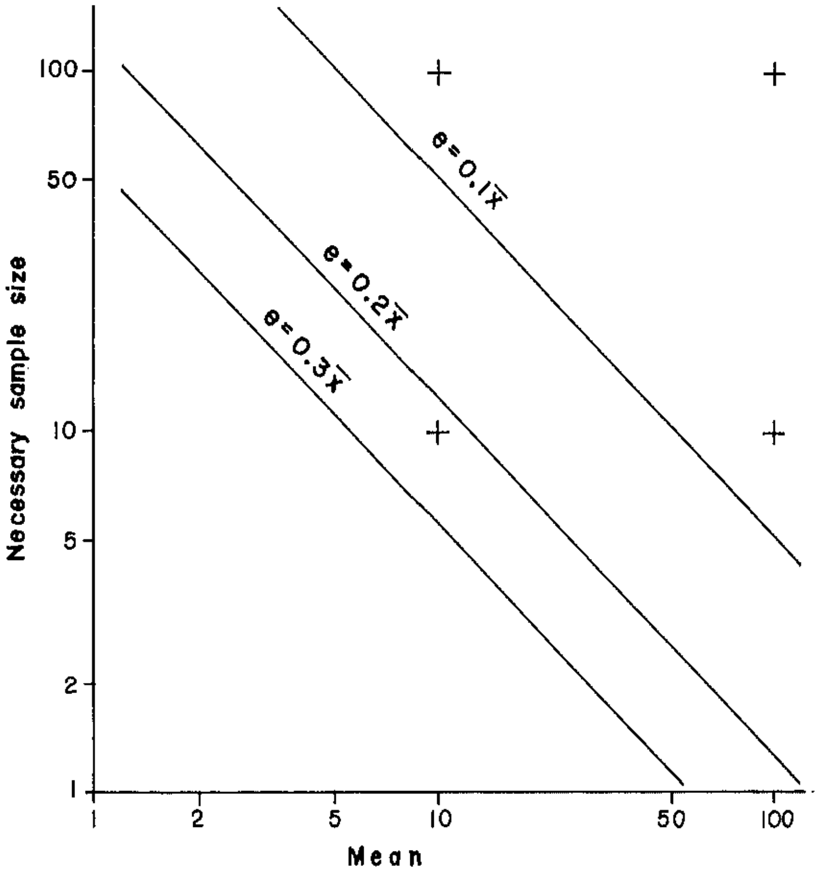

Fig. 4 Relationship between necessary sample sizes at three levels of tolerable errors and mean number of soybeans per sample unit after transformed into the random distribution.

IV. Discussion

The ratio $s^2/\bar{x}$ as a measure of patchiness quickly reaches 1 with stirring, indicating that an aggregated distribution has been changed into a random pattern (Fig. 2). ITOH and KITCHING (1986) recommended the use of β in m-$\overset{*}{m}$ instead of $s^2/\bar{x}$ to show patchiness; however, the value of β also behaved in the same fashion (Fig. 3). It is, threfore, clear that patterns of aggregation are transformed into random patterns by stirring substrates. The randomness makes the determination of the necessary sample size simple and the necessary sample size much smaller without sacrificing statistical reliability.

I try here to ascertain the reliability of the necessary sample size estimated simply in this experiment, where the mean value of soybeans per soil block (sample unit) was 2.77, viz. 1000/361. Using the equation (3), the sample size necessary to achieve a 0.05 probability

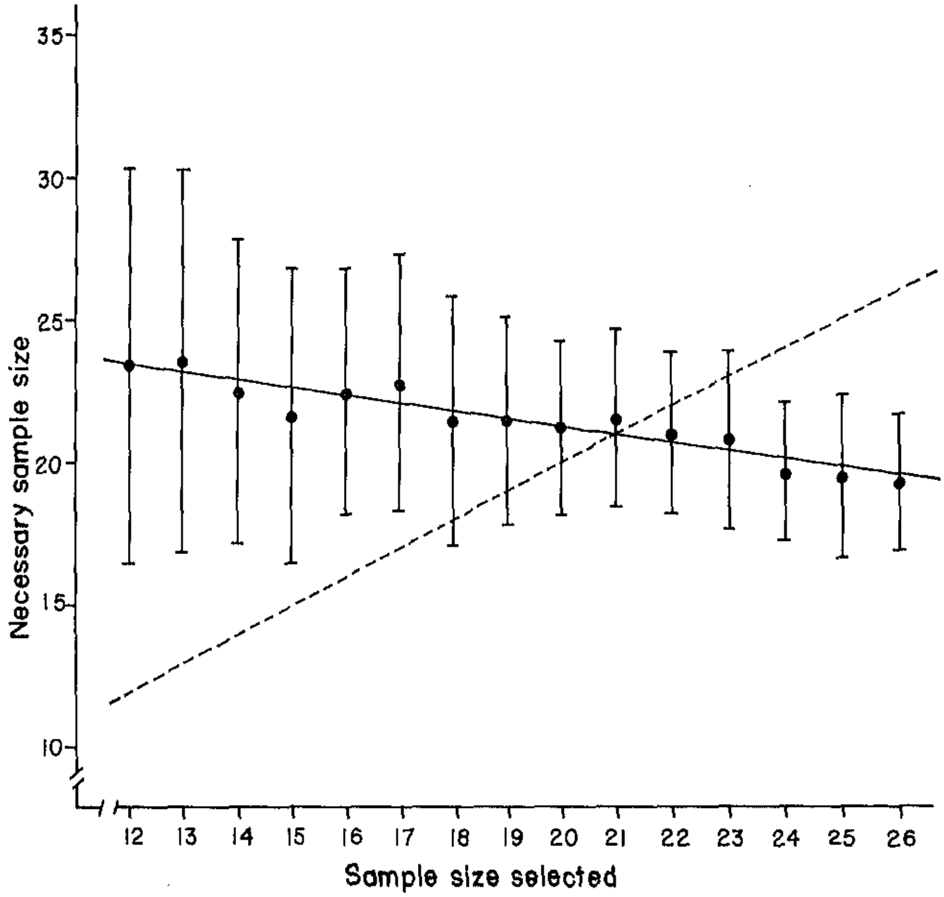

Fig. 5 Regression line through mean necessary sample sizes on arbitrarily selected sample size. Closed circles and bars indicate mean values and 95 % confidence intervals, respectively. Dotted line is of y=x.

level at e=0.3$\bar{x}$ is 21 sample units. For several sample sizes smaller or larger than 21 in the number of sample units, the data of soybeans were randomly selected from the data file at St-10 in the computer 100 times to estimate the necessary sample size for every time, the mean sample size and the 95 % condifence interval, and the regression line through the means was obtained (Fig. 5). The regression line indicates that necessary sample size gradually declines with the number of sample units selected. As shown in the figure, the dotted line, y=x, crossed the regression line on the estimated sample size, viz. x=21. This suggests that the estimated sample size is sufficiently satisfied. On the other hand, in Fig. 5, the line y=x crossed bars of 95 % confidence intervals in the range from x=18 to x=23. Therefore, to secure statistical precision, the actual sample size to be taken should be preferably a little larger than the calculated sample size.

When making quantitative studies of soil microarthropods in fields, it is not difficult to take large numbers of sample units, while it is difficult to prepare collective preparations (TAMURA, 1974) and to count individuals of all the species in each sample unit. Investigators are, therefore, usually forced to sacrifice statistical accuracy by treating an inadequate number of sample units.

It would clearly be desirable to increase the statistical reliability with the examination of a small number of sample uits even if the number of sample units which needed to be taken was increased, since this would diminish the laborious part of the process. This experiment clearly demonstrates that stirring the soil sample clearly accomplishes this end.

A new sampling method for quantitatively studying soil microarthropods is proposed below.

1) In a preliminary survey, 10 or 15 sample units of soil are taken at random using a given sampler from the site to be surveyed later, and the mean and variance among the numbers of animals in the sample units are computed to estimate the necessary sample size to be used later.

2) In the actual survey, a large number of soil sample units sufficiently exceeding the necessary sample size estimated in the preliminary survey are taken over a wide array of the site, 10 of these being randomly selected to measure the respective net weight of soil. Then the mean unit weight is calculated as the standard weight per unit soil.

3) All the sample units of soil taken are together put into a topless box or a large bag with a wide bottom area. Soil blocks are loosened and the sample (total of sample units) is gently stirred so as to avoid damaging any animals contained.

4) A subsample, which is roughly equivalent in weight to the mean unit weight, is quickly taken from the sample to put into a beaker, being weighed its net weight with a portable balance. In the same manner, other subsamples are in turn taken to attain to the number a little more than one fifth of the neccessary number of soil sample units estimated in the preliminary survey.

5) These subsamples will differ in weight, therefore the ratio of the standard weight to the weight of each subsample is calculated to obtain standarizing index to be used later.

6) Each subsamples is put in a separate Tullgren funnel. The animals extracted are mounted in collective preparations. Then, the number of indivduals of every species occurred is counted.

7) The number of individuals counted must be corrected by multiplying by the standarizing index to compensate for the different weithts of soil. Therefore, the numbers of individuals will, in most cases, be expressed in decimals.

A comparison between estimations of the densities of animals obtained from the stirred, homogenized sample and from the non-stirred, intact sample in an actual field survey will be given in another paper.

V. Acknowledgments

I am grateful to students of my laboratory for their assistance in counting soybeans, and to Dr. J. KOJIMA, Tokyo Metropolitan University, for his kind help with computation. Dr. K. CHRISTIANSEN, Grinnell College, Iowa, is thanked for checking the manuscript.

VI. References

COCHRAN, W. G. 1950. Sampling techniques. John Wiley & Sons, New York.

ELLIOTT, J. M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Sci. Publ. No. 25, Freshwater Biological Association, Ferry House, U. K.

ITOH, Y. and R. L. KITCHING 1986. The importance of non-linearity: A comment on the views of Taylor. Res. Popul. Ecol., 28: 39-42.

IWAO, S. 1968. A new regression method for analysing the aggregation pattern in biological populations. Res. Popul. Ecol., 10: 1-20.

—— and E. KUNO1971. Use of the regression of mean crowding on mean density for estimating sample size and transformation of data for the analysis of variance. Res. Popul. Ecol., 10: 210-214.

KIKUZAWA, K., P. SAICHUAE, K. NIIJIMA, M. TANAKA and J. AOKI 1967. On the sampling and extracting technique for soil micro-arthopods. Jap. J. Ecol., 17: 20-28.

LLOYD, M. 1967. "Mean crowding". J. Anim. Ecol., 36: 1-30.

TAKEDA, H. 1973. A preliminary study on collembolan populations in a pine forest. Res. Popul. Ecol., 15: 76-89.

—— 1977. Ecological studies of collembolan populations in a pine forest soil. IV. Comparison of distribution patterns. Res. Popul. Ecol., 21: 120-134.

TAMURA, H. 1974. A method for the measurement of curled body length of Collembola. Rev. Ecol. Biol. Sol, 11: 353-362.

—— 1978. Collembola in a beech forest at Mt. Tanigawa, Nashigahare grassland and Kirigamine grassland. Report on the biomass of wild animals in various ecosystems (ed. Y. Kitazawa), pp. 93-100. (in Japannese with English summary)

—— and Y. MIHARA 1977. Preliminary report on changes in the life-cycle of Collembola with altitude. Rev. Ecol. Biol. Sol, 14: 37-38.