South East Asian Studies, Vol. 13, No. 3, December 1975

Paddy Soils in Tropical Asia

Part 5. Soil Fertility Evaluation

Keizaburo KAWAGUCHI* and Kazutake KYUMA**

According to the Soil Survey Manual,[1] "Soil fertility is the quality that enables the soil to provide the proper compounds in the proper amounts and in the proper balance for the growth of specified plants when other factors such as light, temperature, moisture, and the physical condition of the soil, are favorable." Thus, by definition the chemical nature of a soil determines its fertility. For paddy soils the importance of soil physical properties or tilth is relatively minor because of their specific management condition of' waterlogging. Therefore, soil fertility should deserve a higher attention in attempt to evaluate the capability of paddy soils.

There have been many attempts to evaluate soil fertility in different countries or regions, such as the ones known as land capability classification and the Storie Index. In all these methods so far proposed contributing factors have been only subjectively evaluated and the final result has been presented usually in a qualitative form.

Even the Soil Survey Manual says that "Soil fertility is not directly observable. Thus, soil may be grouped into fertility classes only by inference." Then is it not possible to evaluate soil fertility quantitatively, and to attain a fertility classification in a more objective and reproducible way?

In an attempt to answer this challenging question we have proposed a method of fertility evaluation particularly intended for paddy soils[2,3,4] and a few preliminary applications of it made on tropical Asian paddy soils.[5,6,7] In this paper the results of the most recent study along the same line are presented.

Data and Methods

The data of the plow layer samples of the same 410 tropical Asian paddy soils were used in this study. Descriptions and the results of correlation analysis of these data were given in previous papers.[8,9,10]

For processing these data for the purpose of fertility evaluation we use two

---

* 川口桂三郎, Faculty of Agriculture, Kyoto University

** 久馬一剛, The Center for Southeast Asian Studies, Kyoto University

multivariate statistical methods in this study; principal component analysis and factor analysis. (See, for details, Okuno *et al.*[11]) The aims and the procedures of these two methods are mutually related, as described below.

The method of principal component analysis (PCA) is often used for attaining a "parsimonious summarization of a mass of observation" (Seal[12]). In other words, it is used to extract the hidden essence of a thing or material that is not directly measurable, making use of the correlation between various measurable variables related to it.

Given n samples, each of which is defined by p characters, they can be expressed as n points scattered in a p-dimensional space. PCA aims at reducing the p-axes to orthogonal m-axes, where $m<p$. Mathematically this is attained by an orthogonal transformation of the original p-axes, and only after the transformation, the number of orthogonal axes to be retained, i. e. m, is decided, taking, among others, the expected loss of information into consideration.

Factor analysis (FA) has aims similar to those of PCA. There are, however, certain important differences between the two. In the case of PCA, there are no previous assumptions concerning the number and character of the principal components or factors to be extracted, whereas in the case of FA the following must be assumed:

1) The number of common factors to be considered;

2) The extent of contribution of each variable to the common factors.

The fundamental model of FA is expressed as follows:

$$x_i = a_{i1}f_1 + a_{i2}f_2 + \cdots\cdots + a_{ik}f_k + \cdots\cdots + a_{im}f_m + e_i \qquad (1)$$

where x_i is the standardized ith variable ($i=1, 2, \cdots\cdots, p$), f_k is the score of the kth common factor for each sample ($k=1, 2, \cdots\cdots, m$), a_{ik} is the factor loading for the ith variable and kth common factor, and e_i is error or unique factor of the ith variable that is not explained by the m factors. The unique factor e_i is assumed to be orthogonal to all the common factors and to other unique factors associated with the other variables in the set. Therefore, if there is any correlation between the two variables i and j, it is assumed to be due to the common factors.

In the above model, the communality h_i^2 is defined as follows:

$$h_i^2 = a_{i1}^2 + a_{i2}^2 + \cdots\cdots\cdots + a_{im}^2 \qquad (2)$$

The communality makes up, together with δ_i^2, or variance of the unique factor e_i, the variance of the ith variable x_i or $V\{x_i\}$, thus

$$V\{x_i\} = h_i^2 + \delta_i^2 \qquad (3)$$

Another feature of FA is that the estimated factor can be rotated freely so as to make the interpretation of the factor easier. This is possible because of what is called indeterminacy of factor axes. If the factors, after rotation, are interpretable, computation of factor scores follows, based on the least square estimation.

The actual procedure to be used in this study consists of the following steps:

1) Determination of the number of factors to be considered by means of PCA. A

characteristic equation, in which R is the correlation matrix of the variables used,

$$|R-\kappa I|=0 \tag{4}$$

is solved and the principal components having the eigenvalues (κ) larger than 1 are considered.

2) Preliminary communalities are computed as the squared multiple correlation between a variable and the rest of the variables in the set. In a matrix form it is expressed as

$$H=I-\{\mathrm{diag}\cdot(R^{-1})\}^{-1} \tag{5}$$

where H is a diagonal matrix with p-elements of h_i^2 (i=1, 2, ……, p), and R^{-1} is the inverse of the correlation matrix. Since communality is defined as the proportion of a variable sharing something in common with other variables in the set, the communality of a variable cannot be smaller than the squared multiple correlation.

3) Principal factor analysis; starting from the reduced correlation matrix, the following characteristic equation is solved:

$$|(R-I+H)-\lambda I|=0 \tag{6}$$

The same number of factors as determined in 1) is extracted, and the variances accounted for by these factors become new communality estimates. The diagonal elements are then replaced with these new communalities and the same process is continued until the differences in the two successive communality estimates become negligible.

4) The factor loadings are computed from the final solution of eigenvalues and eigenvectors, as follows:

$$a_{ik}=\lambda_k^{1/2}\cdot l_{ik} \tag{7}$$

where a_{ik} is as before, λ_k is the kth eigenvalue, and l_{ik} is the ith element of the kth eigenvector that corresponds to the kth eigenvalue.

5) The obtained factor axes are orthogonally rotated by Kaiser's varimax method which aims at approaching Thurston's simple structure to facilitate interpretation of the factors. Interpretation is easier when some of the variables make high contribution to a factor, whereas the other variables make a negligible contribution or none at all. In other words, elements of a factor loading vector should approach either of the extremes, unity or zero. Kaiser's varimax method maximizes the variance of the square of the elements of a factor loading vector to attain the above goal.

6) If the results of rotation are reasonably interpretable, then the factor scores for individual soils are computed. They are computed by the following general formula:

$$f_k=b_{1k}x_1+b_{2k}x_2+\cdots+b_{ik}x_i\cdots+b_{pk}x_p \tag{8}$$

where b_{ik} is the factor score coefficient for the ith variable and kth factor. According to the model of FA, the best estimate of x_i in terms of f_k is

$$x_i=a_{i1}f_1+a_{i2}f_2+\cdots+a_{ik}f_k\cdots+a_{im}f_m \tag{9}$$

The ($m\times p$) matrix B (factor score coefficient matrix) can be computed by the formula

$$B=A'R^{-1} \tag{10}$$

where A' is a transposed factor loading matrix and R^{-1} is as before.

Based on the preliminary studies, the characters used in this study are those 11 variables as listed in Table 1. The selection was made taking the following into consideration:

1) The primary data directly obtained from analysis are preferred.

2) One of the items of complementary data is omitted. In other words, one of the characters that are highly negatively correlated is omitted, leaving the character which is easier to interpret relative to fertility. For example, clay and sand are found to be highly negatively correlated (cf., Table 2 in Part 3). High sand content in soils clearly has a negative effect on the fertility, while high clay content does not always have a positive effect. Therefore, the clay data was omitted in this case.

Since multivariate normal distribution of the variables is implicitly assumed in FA, logarithmic transformation was applied to the variables in order to make the positively skewed variables (cf., Part 1 and 2) approach the normal distribution.

All the above computations were done with the programs contained in SPSS (Statistical Package for Social Sciences) of the Data Processing Center of Kyoto

Table 1 List of Characters used for Analysis

Character No.	Name	Brief Description
1	T. C. (Total Carbon)	in % of air-dried soil, Tyurin's wet combustion method.
2	T. N. (Total Nitrogen)	in % of air-dried soil, Kjeldahl digestion and steam distillation.
3	NH_3-N	in mg N/100g of air-dried soil, after incubation for 2 weeks at 40°C.
4	Bray−P	in mg P_2O_5/100g of air-dried soil, Bray-Kurz No. 2 method.
5	Exch. K	in me/100g of air-dried soil, *N* NH_4-acetate extraction, flame photometry.
6	CEC	in me/100g of air-dried soil, buffered neutral *N* $CaCl_2$ medium.
7	Avail. Si	in mg SiO_2/100g of air-dried soil, pH 4 Acetic acid extraction at 40°C.
8	T. P. (Total Phosphorus)	in mg P_2O_5/100g of air-dried soil, either HF−H_2SO_4 or $HNO_3-H_2SO_4$ digestion.
9	HCl−P	in mg P_2O_5/100g of air-dried soil, 0.2*N* HCl extraction at 40°C for 5 hrs.
10	Sand	in % of organic matter-free dried soil, sum of coarse and fine sands.
11	Exch. Ca+Mg	in me/100g of air-dried soil, *N* NaCl extraction, EDTA titration.

University.

Results and Discussions

In a preceding paper[10] the results of correlation analysis of the soil data were presented. The main conclusions were as follows:

1) Characters related to base status, texture, and clay mineralogy are mutually highly correlated.

2) Characters related to organic matter status are mutually highly correlated, but only slightly to insignificantly with other charcters.

3) The same is true for the characters related to phosphorus status.

The correlation matrix for the log-transformed 11 characters used in this study is given in Table 2. The correlations among the variables are generally improved, but the above stated appears still to be valid.

Starting from the correlation matrix of Table 2, PCA was run to estimate the number of factors to be considered. The largest 5 eigenvalues of the correlation matrix and the cumulative percentages of the variances in the total are:

	Eigenvalue	Percent of Total Variance	Cumulative Percent of Total Variance
1.	5.501	50.0	50.0
2.	1.836	16.7	66.7
3.	1.673	15.2	81.9
4.	0.698	6.3	88.3
5.	0.372	3.4	91.6

Since the variances of the first 3 factors are larger than 1, explaining altogether more than 80 % of the total variance and that the fourth accounts for a considerably lower

Table 2 Correlation Matrix of Log-Transformed 11 Characters

	LTC	LTN	LNH3N	LTP	LBRAYP	LHCLP	LCAMG	LEXK	LCEC	LAVSI	LSAND
LTC	1.000	0.944	0.658	0.550	0.244	0.204	0.345	0.465	0.555	0.349	−0.442
LTN	0.944	1.000	0.670	0.558	0.259	0.209	0.245	0.417	0.445	0.264	−0.404
LNH3N	0.658	0.670	1.000	0.391	0.095	0.082	0.142	0.255	0.267	0.241	−0.215
LTP	0.550	0.558	0.391	1.000	0.576	0.563	0.527	0.650	0.603	0.605	−0.274
LBRAYP	0.244	0.259	0.095	0.576	1.000	0.859	0.217	0.405	0.222	0.260	−0.014
LHCLP	0.204	0.209	0.082	0.563	0.859	1.000	0.330	0.398	0.284	0.327	−0.022
LCAMG	0.345	0.245	0.142	0.527	0.217	0.330	1.000	0.775	0.912	0.809	−0.473
LEXK	0.465	0.417	0.255	0.650	0.405	0.398	0.775	1.000	0.807	0.710	−0.509
LCEC	0.555	0.445	0.267	0.603	0.222	0.284	0.912	0.807	1.000	0.812	−0.561
LAVSI	0.349	0.264	0.241	0.605	0.260	0.327	0.809	0.710	0.812	1.000	−0.310
LSAND	−0.442	−0.404	−0.215	−0.274	−0.014	−0.022	−0.473	−0.509	−0.561	−0.310	1.000

variance than the third, we decided to use only the first 3 factors in the next steps in factor analysis.

Using squared multiple correlation coefficients between a variable and the rest in the set as the preliminary communality estimate, principal factoring with iteration was run. After 16 iterations, the final factor loadings and communalities were obtained, as shown in Table 3. The first factor loads moderately to heavily on all the variables, while the second and the third load mainly on the characters related to organic matter status and available phosphorus status, respectively.

The communality figures indicate that NH_4-N and Sand are represented only partially by the 3 factors. In other words, they have larger proportions of the unique factor that is not shared by the other variables.

The varimax rotation gave the results as shown in Table 4. Now the interpretation

Table 3 Factor Loading Matrix and Communality for 3 Principal Factors

	FACTOR 1	FACTOR 2	FACTOR 3	COMMUNALITY
LTC	−0.744	−0.582	0.194	0.930
LTN	−0.699	−0.635	0.304	0.984
LNH3N	−0.455	−0.474	0.172	0.462
LTP	−0.787	0.096	0.220	0.676
LBRAYP	−0.502	0.464	0.664	0.908
LHCLP	−0.514	0.503	0.542	0.811
LCEC	−0.889	0.083	−0.409	0.964
LCAMG	−0.794	0.282	−0.455	0.916
LEXK	−0.833	0.173	−0.179	0.756
LAVSI	−0.746	0.232	−0.292	0.696
LSAND	0.499	0.179	0.259	0.348

Table 4 Terminal Factor Loading Matrix for 3 Factors after Varimax Rotation

	FACTOR 1	FACTOR 2	FACTOR 3
LTC	0.288	0.913	0.118
LTN	0.173	0.965	0.151
LNH3N	0.113	0.668	0.046
LTP	0.478	0.407	0.531
LBRAYP	0.089	0.104	0.943
LHCLP	0.182	0.042	0.881
LCEC	0.936	0.276	0.108
LCAMG	0.944	0.053	0.147
LEXK	0.777	0.246	0.303
LAVSI	0.796	0.118	0.218
LSAND	−0.486	−0.323	0.086

of the factors is much easier. The first factor loads only on the characters related to base status and parent material, such as CEC, exchangeable cations, available silica, total phosphorus and sand, with the last-mentioned being opposite in sign to the rest. Thus, the first factor may be named as inherent potentiality (IP) that is determined primarily by the nature and amount of clay and base status.

The second factor, is related to TC, TN, and NH_4-N. Moderately high loadings on TP and Sand are interpretable in terms of organic phosphorus and textural control on organic matter accumulation (cf., Part 3[10]), respectively. Therefore, the second factor may be named as organic matter and nitrogen status (OM).

The third factor is interpreted clearly as available phosphorus status (AP). Factor loading on TP is much less than those on Bray-P and HCl-P. Contribution of other variables to this factor is very minor.

It is observed that these three mutually orthogonal factors extracted are in accordance with the result of correlation analysis referred to above. This leads to an interesting and important inference that soil fertility of tropical Asian paddy soils is made up of at least three major components and that both organic matter status and available phosphorus status of these soils are independent of what we call inherent potentiality.

The factor scores were computed for individual soil samples, so that the quantitative evaluation of the three components of soil fertility may be made. The coefficient matrix for the score computation is given in Table 5. Since the data were log-transformed before subjecting to the analysis, the same transformation is needed for score computation. Moreover, the transformed data must further be standardized using the mean and the standard deviation vectors as given in Table 6. The scores are the sum of the products between the coefficient and the transformed and standardized datum corresponding to the coefficient, as shown in equation (8).

Table 5 Factor Score Coefficient Matrix for 3 Factors

	FACTOR 1	FACTOR 2	FACTOR 3
LTC	−0.151	0.268	−0.078
LTN	−0.147	0.839	0.010
LNH3N	0.045	−0.012	0.008
LTP	0.051	−0.025	0.084
LBRAYP	−0.091	−0.101	0.701
LHCLP	−0.059	−0.033	0.278
LCEC	0.757	−0.132	−0.214
LCAMG	0.306	−0.144	0.029
LEXK	0.130	0.018	−0.026
LAVSI	−0.058	0.073	0.087
LSAND	0.028	0.012	0.004

The scores thus computed for the samples were standardized with a mean of zero and a variance of unity. Therefore, positive score values indicate above-average status with reference to the overall mean for the 410 sample soils, while negative values below-average status. In Table 7 soils having both the highest and the lowest 10 scores for each factor are listed.

The distribution of the factor scores is given in Table 8. The inherent potentiality scores show somewhat negatively skewed distribution. The highest 10 scores are for the soils of either Grumusols or grumusolic Alluvial soils, 6 from India and 4 from Indonesia. On the negative side, 14 scores are smaller than -2, of which one is even smaller than -3. Twelve of them are from the Northeast Plateau region of Thailand and the other two from Malaysia and India. These are, without exception, sandy soils derived from strongly weathered parent materials.

In the case of the organic matter-nitrogen status, the distribution of the scores is somewhat positively skewed. Here 13 soils have scores greater than 2, of which 6 have scores even greater than 3. These extraordinarily high scores are for swamp soils, 10 from West Malaysia and one each from the wet zone of Sri Lanka, Bangladesh, and Indonesia; the last mentioned is exceptional, being an Ando soil of volcanic ash origin. On the contrary, many of the low-scored soils are from India and Thailand where climatic control seems to be more dominant than textural (cf., Part 3[10)]).

Available phosphorus status shows nearly normal distribution. Three of the four soils hsving scores greater than 2 are from Andhra Pradesh, India, and the fourth is an Alluvial soil from Burma. The poorest soils in terms of available phosphorus are from Cambodia; in fact 5 of the lowest ranked 10 soils are from Cambodia.

Although some reservation has to be allowed regarding the sampling procedure adopted in this study (cf., Part 1[8)]), a rough estimation of fertility status can be made

Table 6 Means and Standard Deviations of the Log-Transformed 11 Charactors for 410 Sample Soils

VARIABLE	MEAN	S. D.
LTC	0.044	0.297
LTN	−0.994	0.282
LNH3N	0.731	0.425
LTP	1.775	0.429
LBRAYP	0.171	0.584
LHCLP	0.608	0.712
LCEC	1.159	0.342
LCAMG	0.993	0.484
LEXK	−0.623	0.449
LAVSI	1.195	0.515
LSAND	1.314	0.544

Table 7 List of Soils having Highest and Lowest 10 Scores for Each of the 3 Factors

IP		OM		AP	
In 33	1.908	M 20	3.484	B 8	3.208
In 29	1.774	Sr 27	3.265	I 60	3.101
I 38	1.738	M 40	3.229	I 58	2.959
I 55	1.728	M 22	3.198	I 56	2.265
In 31	1.709	M 21	3.056	T 62	1.946
I 57	1.645	M 19	3.003	I 8	1.923
In 22	1.633	In 4	2.906	Ph 44	1.919
I 65	1.605	Bd 48	2.417	I 27	1.898
I 74	1.570	M 39	2.404	Bd 15	1.825
I 53	1.564	M 41	2.329	Ph 43	1.781
T 20	−2.179	I 33	−1.802	In 33	−1.867
T 11	−2.182	I 38	−1.816	Sr 5	−1.889
M 14	−2.190	I 27	−1.821	T 72	−1.899
T 25	−2.198	I 11	−1.940	T 3	−1.902
T 15	−2.247	T 27	−2.014	T 15	−1.921
I 27	−2.261	Sr 1	−2.026	Ca 3	−1.945
T 16	−2.290	T 30	−2.113	Ca 11	−1.962
T 13	−2.691	T 18	−2.116	Ca 10	−1.987
T 14	−2.787	I 50	−2.347	Ca 4	−2.004
T 17	−3.181	T 23	−2.699	Ca 2	−2.202

Bd—Bangladesh ; B—Burma ; Ca—Cambodia ; I—India ; In—Indonesia ; M—Malaysia ; Ph—Philippines ; Sr—Sri Lanka ; T—Thailand

Table 8 Distribution of the 3 Factor Scores around the Overall Mean (or Zero) for 410 Sample Soils

	IP	OM	AP
> 3	0	6	2
3 ∼ 2	0	7	2
2 ∼ 1	72	47	67
1 ∼ 0	151	141	131
0 ∼ −1	120	150	148
−1 ∼ −2	53	53	58
−2 ∼ −3	13	6	2
< −3	1	0	0

for each country or region by calculating a mean score for the samples concerned. Table 9 shows the mean values of the 3 factors for each country. The smallest figure of standard deviation is for the organic matter-nitrogen status of Indian paddy soils (0.62) and the largest figure is for inherent potentiality of Thai soils (1.20). Inherent potentiality of Cambodian soils, organic matter-nitrogen status of Sri Lanka soils, and available phosphorus status of Burmese and Philippine soils show as large variance (ca. 1.0) as the corresponding variable for the total of 410 samples.

Inherent potentiality is highest for the soils of Indonesia and the Philippines, followed by those of India. The first two countries are situated in a region influenced by volcanic activity and the parent material of the soil is continuously rejuvenated by fresh volcanic ejecta. India is located in semi-arid to sub-humid climatic regions and the weathering process of the soil material has not been very intensive, especially in the basaltic rock area of the Decan Plateau that constitutes the catchment of the Godavari-Krishna rivers.

On the contrary, soils of Malaysia and Sri Lanka, which are situated in permanently humid to monsoonal climatic regions of the low latitudes, are among the poorest with respect to inherent potentiality. Soils of Bangladesh, Cambodia and Thailand are mostly on the poorer side of the overall mean.

Organic matter-nitrogen status is by far the highest for Malaysian soils, with a mean score as high as 1.40. The second highest is for the Philippines with a mean of 0.34. Conversely, the Indian soils with the lowest mean score −0.78 are the poorest. Soils of the rest of the countries are more or less similar, with mean scores clustering around the overall mean.

The available phosphorus status is high for the soils of India and Bangladesh, while Cambodian soils are the poorest, followed by Thai soils. The regionality observed in this property may be ascribed to the climate and parent material, or in other words,

Table 9 Means and Standard Deviations of the 3 Factor Scores for the Respective Countries

COUNTRY	No. of Samples	IP		OM		AP	
		MEAN	S. D.	MEAN	S. D.	MEAN	S. D.
Bangladesh	53	−0. 438	0. 708	0. 176	0. 704	0. 459	0. 800
Burma	16	0. 118	0. 710	−0. 128	0. 646	0. 308	1. 121
Cambodia	16	−0. 231	0. 990	−0. 155	0. 954	−1. 277	0. 841
India	73	0. 449	0. 837	−0. 780	0. 619	0. 581	0. 906
Indonesia	44	0. 618	0. 766	−0. 014	0. 746	0. 031	0. 734
Malaysia	41	−0. 545	0. 719	1. 398	0. 927	0. 026	0. 726
Philippines	54	0. 618	0. 703	0. 337	0. 673	0. 022	1. 041
Sri Lanka	33	−0. 510	0. 853	0. 150	1. 064	−0. 110	0. 650
Thailand	80	−0. 364	1. 195	−0. 347	0. 939	−0. 641	0. 769

to the degree of weathering of the soil material.

A similar calculation was done for the regions that can be defined more or less discretely with respect to the climate, parent material, and areal extension. The result is given in Table 10. There are some regions in which one or more of the soil fertility components are still highly variable. The Northeast Plateau region of Thailand and the east coast of Malaysia have very low mean scores of inherent potentiality. A great difference between these two regions, however, is seen in the standard deviation figures. In the Northeast Plateau region of Thailand there occur narrow strips of clayey recent Alluvial soils and patches of Grumusols as associated with basalt outcrops, while the greater part of the region is covered by sandy, sevesely weathered and leached soils, scoring mostly −1.5 to −2 in inherent potentiality scale. By contrast,

Table 10 Means and Standard Deviations of the 3 Factor Scores for Selected Regions

Region	No. of Samples	IP		OM		AP	
		Mean	S. D.	Mean	S. D.	Mean	S. D.
Sri Lanka Wet & Interm. Zone	14	−1. 07	0. 64	0. 84	1. 10	−0. 10	0. 64
Sri Lanka Dry Zone	19	−0. 10	0. 76	−0. 36	0. 70	−0. 12	0. 67
Bangladesh Ganges	15	0. 33	0. 35	0. 04	0. 82	0. 90	0. 50
Bangladesh Madhupur-Barind	9	−0. 75	0. 60	0. 35	0. 64	−0. 14	0. 40
Bangladesh Marginal	16	−0. 87	0. 62	0. 36	0. 66	−0. 06	0. 77
Bangladesh Brahmaputra	13	−0. 57	0. 48	−0. 01	0. 64	1. 01	0. 66
W. Malaysia Kedah-Perlis	10	0. 25	0. 44	1. 21	0. 56	0. 02	0. 54
W. Malaysia East Coast	10	−1. 23	0. 20	0. 78	0. 59	−0. 54	0. 57
India Godavari-Krishna	10	1. 38	0. 29	−0. 73	0. 47	1. 11	0. 95
Thailand NE Platesu	32	−1. 18	1. 27	−1. 14	0. 77	−0. 94	0. 70
Thailand Intermountain Basin	4	−0. 27	0. 52	0. 06	0. 69	−0. 72	0. 80
Thailand Upper Central Plain	14	−0. 04	0. 76	−0. 05	0. 49	−0. 31	0. 69
Thailand Bangkok Plain	24	0. 53	0. 64	0. 24	0. 66	−0. 51	0. 86
Thailand South	6	−0. 40	0. 74	0. 58	0. 48	−0. 29	0. 39

the soils of east coast of Malaysia are very homogeneous in their material nature, all having been derived from strongly leached, medium to coarse textured alluvia. A similar comparison may be made between the wet and intermediate zones of Sri Lanka and the east coast of Malaysia, having comparable mean organic matter-nitrogen scores but greatly different standard deviations.

The highest scores for inherent potentiality and available phosphorus status are possessed by the soils of Godavari-Krishna region, which have, however, the second lowest organic matter-nitrogen score. The soils of the Northeast Plateau region of Thailand are characterized by very low scores of all the three fertility components. High organic matter-nitrogen scores are shared by the soils of both the east coast and the Kedah-Perlis regions of Malaysia. Low available phosphorus scores are a common feature of all the regions of Thailand. Of the regions listed, the Kedah-Perlis Plain of Malaysia and the Ganges sediment region of Bangladesh have relatively well balanced soils with respect to the three fertility components.

Now to effect a fertility classification, the whole range of the computed scores was divided into classes at arbitrary class limits. In this study class limits were set at ±0.25 and ±0.84. The assumption underlying the selection of the limits is that if the distribution of the scores is normal, they should give five classes of almost equal size. The potentiality of each fertility component at different class levels could be designated as follous:

Class No.	Class limits	Potentiality
I	> 0.84	very high
II	0.84 - 0.25	high
III	0.25 - -0.25	intermediate
IV	-0.25 - -0.84	low
V	< -0.84	very low

Table 11 Mean Contents of Clay Mineral Species and Selected Total Elemental Oxides for the Samples Falling into Each Inherent Potentiality Class

IP Class	1	2	3	4	5	F-value
No. of Samples	(88)	(92)	(77)	(70)	(83)	
7 Å	27.84	40.27	44.03	52.71	67.83	51.34
10 Å	8.13	15.00	17.60	18.00	11.15	7.95
14 Å	64.03	44.73	38.38	25.50	19.82	78.09
SiO_2	63.86	66.80	71.61	74.94	84.89	73.46
Fe_2O_3	9.24	7.42	5.75	4.52	2.26	74.89
Al_2O_3	20.31	19.31	16.22	15.60	9.78	41.84
CaO	2.13	1.72	1.69	1.09	0.37	11.25
MgO	1.25	1.20	1.09	0.68	0.31	28.14
TiO_2	1.36	1.25	1.13	1.07	0.87	8.91
K_2O	1.50	2.00	2.28	1.90	1.53	5.90

Making use of this 5 grade classification, the distribution of the samples from each country or region is illustrated as in Fig. 1, which is a collection of histograms showing percentage frequencies of the samples falling into each of the five fertilty classes. Each of the regions listed in Table 10 is also shown by a specific pattern within the respectvie countries. By referring to the figure we can understand better the contents of Table 9 and 10 and the contribution of each region to the make-up of the sample population of the respective countries.

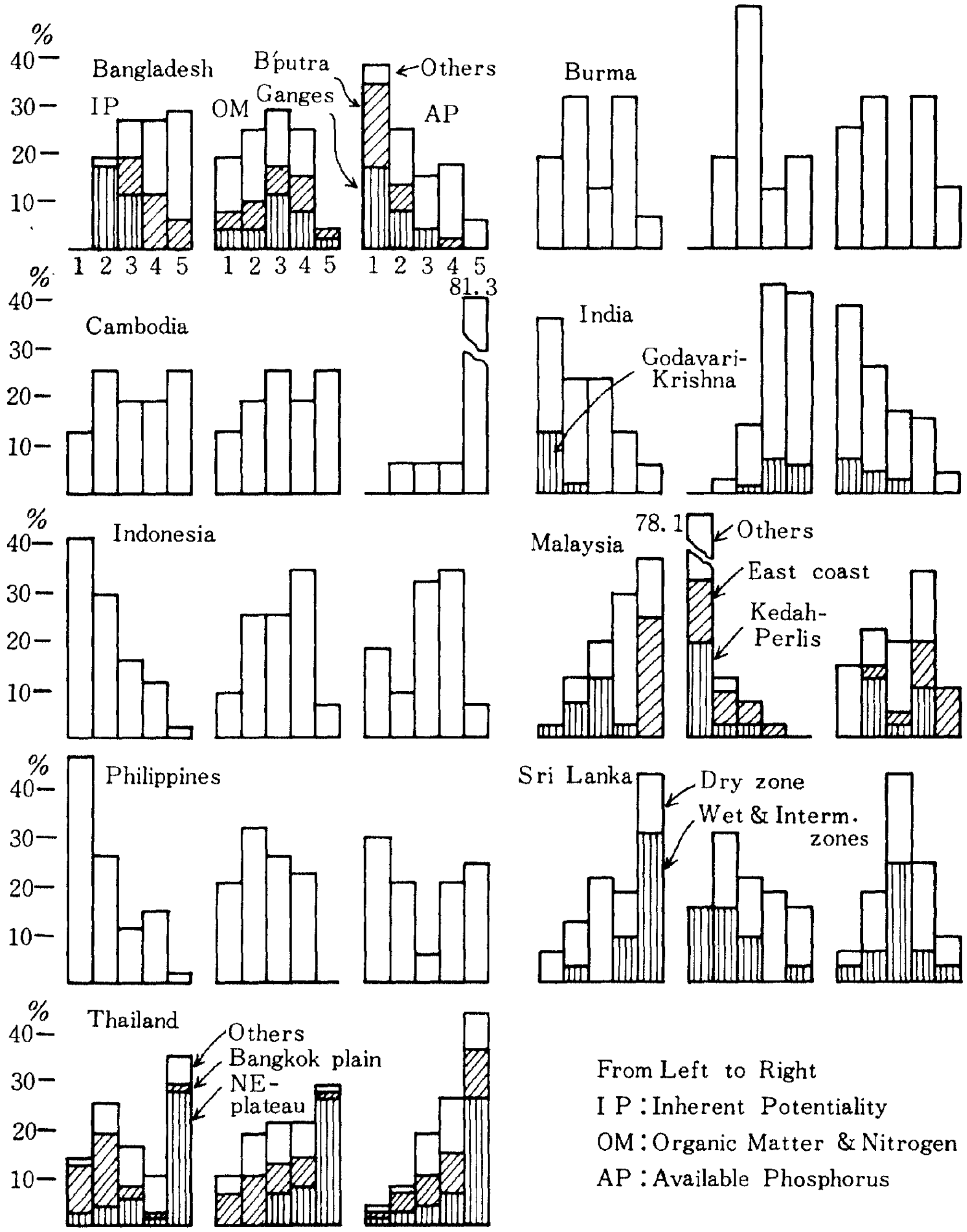

Fig. 1 Percentage Distribution of the Samples from Each Country among Five Classes of Each Fertility Component

Although clay mineral composition and total chemical composition were not directly used in this study, they have been well represented by the inherent potentiality. This is clear from Table 11, which shows the mean contents of clay mineral species and selected elemental oxides for each of the 5 inherent potentialtiy classes. The F-value in the last column is the variance ratio to test against a null hypothesis that there is no difference in the mean contents among the classes. All the F values are much greater than the criterion value of F $(\phi_1=4,\ \phi_2=405;\ \alpha=0.05)=2.4$, indicating a very high significance of the difference in the means. Among the variables listed in the table, 10 Å mineral and total potash contents have a peculiar pattern showing the maximum in the intermediate classes.

Inherent potentiality is thus a compound character closely related to the soil material characteristics. Therefore it naturally follows that it should be correlated with the soil material classes set up in the preceding paper18). To check this the mean and standard deviation of inherent potentiality scores of the samples belonging to each material class were calculated (Table 12). Analysis of variance produced an F-value of 64.0 which is again much higher than F $(\phi_1=9,\ \phi_2=400;\ \alpha=0.05)=1.9$. When the 10 material classes were arranged according to decreasing mean inherent potentiality scores, the following order resulted; VI⩾I>IV⩾III>X>V⩾VII>IX⩾II⩾VIII.
A percentage distribution, among the 5 inherent potentiality classes, of the samples belonging to each material class is shown in Fig. 2. High frequencies of inherent potentiality class 1 and 2 for the material classes VI, I, and IV, and an extremely high frequency of inherent potentiality class 5 for the material class VIII are obvious from it.

In other words the majority of the samples in each material class falls into a narrow range of inherent potentiality classes. This supports the suggestion10) that the material characteristics be used as criteria in soil classification at the "family" and "series" levels, so that these categories will be more homogeneous with respect to soil

Table 12 Mean and Standard Deviation of Inherent Potentiality Scores for the Samples Belonging to Each Soil Material Class

Soil Material Class	Inherent Potentiality	
	Mean	S. D.
Ⅰ	0.955	0.575
Ⅱ	−0.405	0.676
Ⅲ	0.159	0.656
Ⅳ	0.764	0.439
Ⅴ	−0.188	0.766
Ⅵ	0.999	0.613
Ⅶ	−0.193	0.696
Ⅷ	−1.613	0.694
Ⅸ	−0.391	0.636
Ⅹ	−0.046	0.435

capability.

In recent years many attempts have been made to derive a crop yield prediction equation by means of multiple regression analysis, in which soil characters are used as independent variables together with others, such as climatic and management factors that are thought to be relevant to crop yield. Soil characters taken up in such attempts are often humus content, clay content, a certain nutrient content, etc. We believe that the three factor scores estimated in this study are most suitable to the purpose, because

1) the three factors are compound characters derived from many individnal characters and represent the most important fertility components that are relevant to the yield, and

2) these factors are mutually indpendent, so that they best fit the multiple regression model.

Kyuma[4] showed that the three factors alone account for about 60 % of the variance of paddy yield in a study in Malaya on yield data reported by farmers, though such data cannot be considered as representing only soil variability. The coefficient of

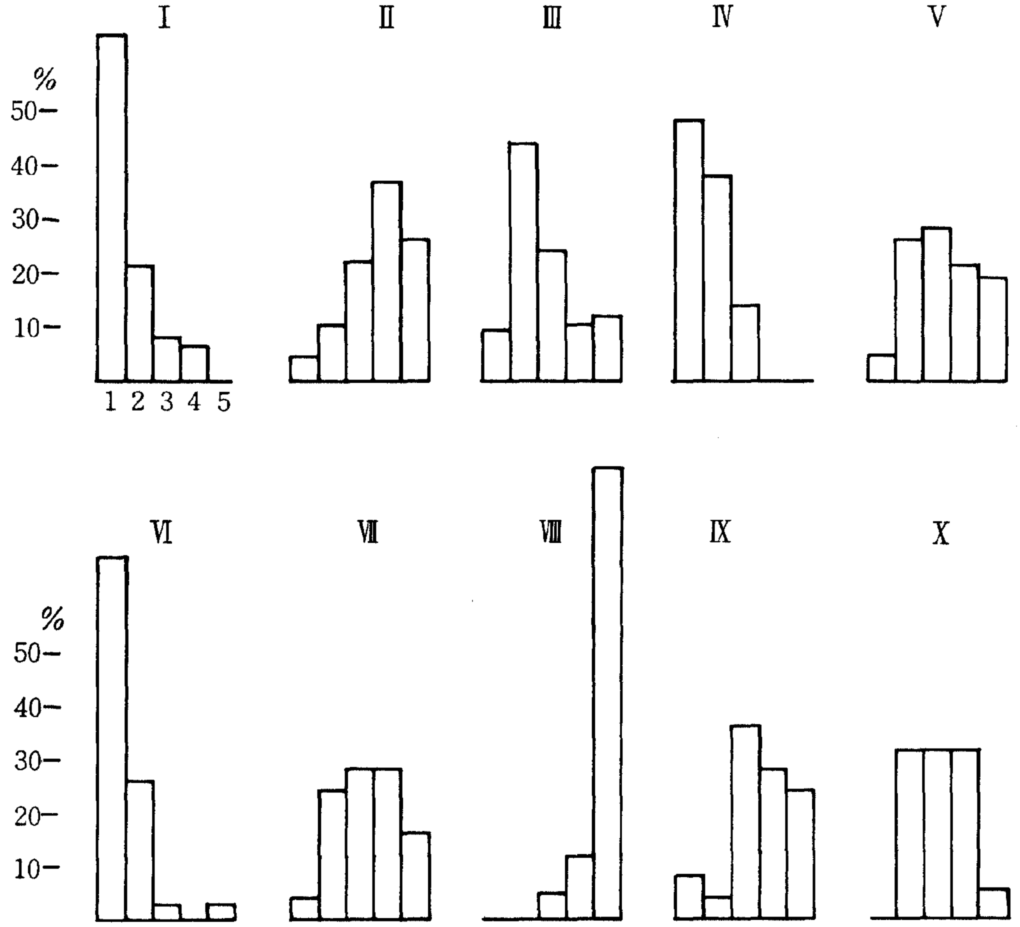

Fig. 2 Percentage Distribution of the Samples belonging to Each of the 10 Soil Material Classes (I～X) among Five Inherent Potentiality Classes

determination is, therefore, regarded as fairly satisfactory. More studies along this line are necessary to check this further.

Summary

Although there have been many studies intended to assess soil fertility, no established method for drawing quantitative conclusions relative to soil fertility is yet available. In this study the method of principal factor analysis was successfully applied to a set of 11 variables (TC, TN, NH_3-N, TP, Bray-P, HCl-P, Ex-(Ca+Mg), Ex-K, CEC, Available Silica, and Sand) routinely analyzed in the soil laboratory.

Three mutually independent and clearly definable fertility component factors were extracted from these data on 410 tropical Asian paddy soils. They were named as inherent potentiality, organic matter-nitrogen status, and available phosphorus status. Inherent potentiality is closely related to the nature and amount of clay and the base status of soil. So it shows close relationship with the soil material classification set up in the previous study.

The samples from each country or region were characterized by the scores of these three factors. Furthermore, based on the factor scores, a five-grade classification (very high, high, intermediate, low, very low) for each of the three fertility components was set up.

The possibility of using these fertility components in a multiple regression analysis for crop yield prediction has also been suggested.

References

1) Soil Survey Staff, *Soil Survey Manual*, USDA Handbook No. 18, Washington, D. C., 1951.

2) Kyuma, K. and Kawaguchi, K., "A Method of Fertility Evaluation for Paddy Soil, I. First Approximation: Chemical Potentiality Grading," *Soil Sci. & Plant Nutr.*, Vol. 19, pp. 1–9 (1973).

3) Kyuma, K., "A Method of Fertility Evaluation for Paddy Soils, II. Second Approximation: Evaluation of Four Independent Constituents of Soil Fertility", *Soil Sci. & Plant Nutr.*, Vol. 19, pp. 11–18 (1973).

4) Kyuma, K., "A Method of Fertility Evaluation for Paddy Soils, III. Third Approximation: Synthesis of Fertility Constituents for Soil Fertility Evaluation", *Soil Sci. & Plant Nutr.*, Vol. 19, pp. 19–27 (1973).

5) Kyuma, K. and Kawaguchi, K., *Fertility Evaluation of Paddy Soils in South and Southeast Asia—First Approximation: Chemical Potentiality Rating—*, Discussion Paper No. 34, CSEAS, Kyoto University, 1971.

6) Kyuma, K. and Kawaguchi, K., *Fertility Evaluation of Paddy Soils in South and Southeast Asia —Second Approximation: Evaluation of Three Independent Constituents of Soil Fertility—*, Discussion Paper No. 40, CSEAS, Kyoto University, 1972.

7) Kyuma, K. and Kawaguchi, K., "An Approach to the Capability Classification of Paddy Soils in

Relation to the Assessment of Their Agricultural Potential", *Proc. Second ASEAN Soil Conference, Jakarta,* Vol. 1, pp. 283–294(1972).

8) Kawaguchi, K. and Kyuma, K., "Paddy Soils in Tropical Asia, Part 1. Description of Fertility Characteristics", *Tonan Ajia Kenkyu,* Vol. 12, No. 1, pp. 3–24 (1974).

9) Kawaguchi, K. and Kyuma, K., "Paddy Soils in Tropical Asia, Part 2. Description of Material Characteristics", *Tonan Ajia Kenkyu,* Vol. 12, No. 2, pp. 177–192 (1974).

10) Kawaguchi, K. and Kyuma, K., "Paddy Soils in Tropical Asia, Part 3. Correlation and Regression Analyses of the Soil Data", *Tonan Ajia Kenkyu,* Vol. 13. No. 1, pp. 45–57 (1975).

11) Okuno, T. *et al.*, *Tahenryo Kaisekiho* (*Multivariate Statistical Methods*), Nikka-Giren, Tokyo, 1971.

12) Seal, H. L., *Multivariate Statistical Analysis for Biologists,* Methuen & Co., London, 1964.

13) Kawaguchi, K. and Kyuma, K., "Paddy Soils in Tropical Asia, Part 4. soil Material Classification", *Tonan Ajia Kenkyu,* Vol. 13, No. 2, pp. 215–227(1975).