ON THE PROBABLE ERROR OF SAMPLING IN SOIL SURVEYS.

By G. W. ROBINSON, B.A.,
Adviser in Agricultural Chemistry,

AND W. E. LLOYD, B.Sc.,
Research Scholar, University College of North Wales, Bangor.

THE following paper records experiments made to determine the magnitude of the error involved in the sampling of a soil for survey purposes. While it is not held that the results obtained are true for all localities, they serve to indicate the order of magnitude of the various errors.

The ordinary method of sampling a soil for survey purposes is, first to select a field uniform in itself and representative of the soil type which is being examined. Borings are then taken in various parts of the field. The borings from the top soil are united to form the soil sample. Similarly those from the subsoil are united to form the subsoil sample. The number of borings recommended in text-books is from five up to ten or twelve according to the size of the field sampled.

Now the probable error affecting the analysis of a single boring is a function of two probable errors, namely, (1) the *laboratory error*, that is the error of analytical determination (which itself includes an error of sampling from the laboratory sample), and (2) the *field error*, which is the error due to the normal variation in the composition of the soil from point to point in the field.

Now if the probable error of a determination on one boring be P, the probable laboratory error be p_1, and the probable field error be p_2,

$$P = f(p_1 p_2).$$

P, p_1 and p_2 will have different values according to the soil constituent which is determined. There will thus be a different set of P, p_1 and p_2

for nitrogen, phosphoric acid, coarse sand and so on. By a well-known formula

$$P=\sqrt{p_2^2+p_1^2},$$

provided that, as may be assumed to be the case, the errors are independent.

The object of this investigation is to obtain values of p_2, the probable field error, for various soil constituents.

There is practically no literature on the subject. The only account of previous work accessible to the writers is in a paper by J. W. Leather[1]. Leather took duplicate samples from several plots in different localities. Each sample consisted of a mixture of a dozen borings. He found that the differences between determinations on the duplicates of nitrogen, available phosphoric acid and potash respectively, varied from nothing up to more than 20 per cent. Unfortunately the data do not give any basis for determination of probable errors due to field variation[2].

There are various papers which treat of the probable error of experimental plot yields, but since the composition of the soil is only one factor in determining crop yields, it cannot be expected that these probable errors are any measure of the probable error of sampling.

Two fields were therefore investigated, one on a drift and the other on a sedentary soil. The following notes describe the two fields.

Field A, at the College Farm, Aber, near Bangor. This field has been for many years in grass. The soil is glacial drift of local origin. It is not particularly uniform in texture and appearance and for ordinary survey purposes would be reckoned too variable.

Field B, at Cellar, Aberffraw, Anglesey. This field is now in arable cultivation. The soil is derived from the Pre-Cambrian schists. There is a certain admixture of wind blown sand since the field is not far from the sea. The soil is quite uniform in texture and appearance.

In Field A, 25 samples were taken. The diagram, Fig. 1, shows the order in which the borings were made.

Each sample was kept separately in a bag. On arrival at the laboratory all the samples were analysed according to the usual methods. The following analyses were performed: Mechanical analysis, determination of hygroscopic moisture and organic matter, and the

[1] *Trans. Chem. Soc.* 1902, p. 883.

[2] Since sending in this paper, the authors have seen a paper by Pfeiffer and Blanck, *Landw. Versuchs.-Stat.* LXXVIII. Working on the nitrogen content of an experimental field, they obtained a field error of 2·5 per cent. This was, however, over a small area and the result is scarcely of service for survey purposes.

determination of the total P_2O_5. In the case of the P_2O_5 determinations it was felt that the error of analysis would be smaller if absolute determinations were made. Accordingly, the total P_2O_5 was determined by treatment with sodium peroxide.

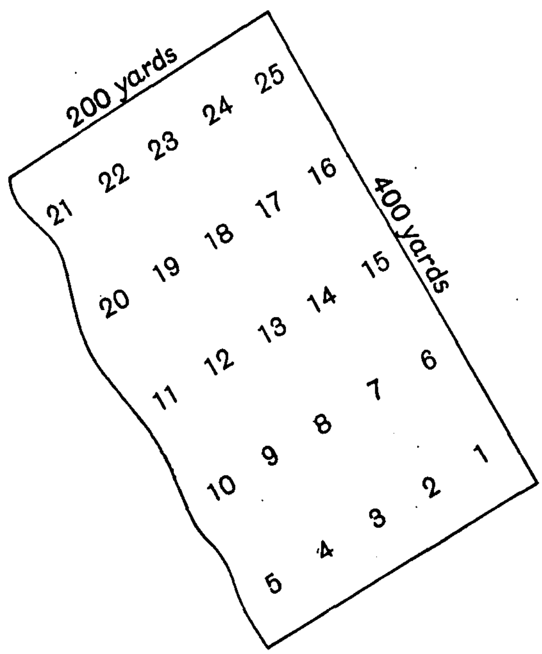

Fig. 1. Field A.

In order to get some idea of the reliability of the figures obtained *i.e.* to determine p_1, six mechanical analyses and P_2O_5 determinations were made of a well mixed composite sample of a number of borings. The results are shown in the following table.

Composite Sample.

	Fine gravel	Coarse sand	Fine sand	Silt	Fine silt	Clay	P_2O_5
1	7·93	16·31	15·42	15·30	24·21	4·80	·343
2	7·79	18·24	15·75	14·84	23·81	5·06	·310
3	7·90	17·56	15·25	14·73	24·65	4·20	·331
4	9·67	17·78	15·53	13·65	22·80	4·88	·315
5	8·03	17·75	15·71	14·45	24·44	5·09	·323
6	8·02	18·17	14·99	14·42	24·96	4·87	·315

The mean values and probable errors calculated from these figures are as follows:

Constituent	Mean value	Probable error
Fine gravel	8·22	± ·52 = ± 6·3 per cent.
Coarse sand	17·63	± ·54 = ± 3·0 „
Fine sand	15·43	± ·20 = ± 1·3 „
Silt	14·56	± ·38 = ± 2·6 „
Fine silt	24·14	± ·52 = ± 2·1 „
Clay	4·81	± ·11 = ± 2·3 „
P_2O_5	·323	± ·008 = ± 2·5 „

Similar determinations were not made for moisture and organic matter. The determinations are made with much less trouble than the mechanical analyses and P_2O_5 estimations, and therefore the moisture and organic matter were determined four times on each boring and averaged. The agreement was so close in all cases that p_1 for these determinations would be exceedingly small both absolutely and in comparison with the corresponding values of P determined.

Fine gravel	Coarse sand	Fine sand	Silt	Fine silt	Clay	Moisture	Organic matter	Phosphorus pentoxide (P_2O_5)
2·43	14·56	18·33	13·78	27·63	4·18	3·65	12·63	·304
7·33	14·80	19·37	11·79	26·47	4·11	4·03	9·94	·261
7·84	16·02	19·20	10·37	27·99	4·06	3·24	8·55	·279
7·34	17·27	21·35	8·59	26·73	4·07	3·77	8·59	·256
7·60	15·95	16·89	10·10	29·38	5·28	3·21	8·89	·261
11·21	20·32	18·09	8·60	22·65	3·46	2·99	10·91	·369
5·32	18·71	16·60	14·95	24·30	3·66	3·62	11·33	·312
17·25	20·36	14·46	11·31	20·10	3·26	2·41	9·71	·329
12·14	19·57	14·91	12·02	21·93	3·00	2·84	12·25	·318
8·37	17·07	19·59	11·45	24·90	3·64	3·13	10·06	·281
7·42	14·59	20·87	13·56	25·12	4·47	2·46	9·61	·314
2·63	11·33	23·37	14·12	28·03	5·54	2·32	9·59	·286
12·49	22·20	14·35	11·51	21·45	3·42	2·07	9·22	·295
11·20	16·16	14·72	13·84	24·92	3·54	2·26	11·26	·298
7·72	16·20	15·67	13·41	27·34	4·35	1·92	10·30	·214
2·83	11·78	15·31	16·13	33·52	5·40	2·10	10·30	·229
8·20	17·90	21·38	12·40	24·77	4·30	1·63	10·75	·250
15·35	19·92	16·09	10·40	21·44	3·79	2·30	9·10	·198
10·56	22·52	15·52	11·39	21·25	4·01	2·63	9·65	·254
7·38	21·48	15·13	12·93	20·00	3·40	2·79	12·72	·440
11·49	20·79	15·74	10·57	22·65	4·90	2·07	9·81	·317
4·73	22·25	15·96	13·43	21·12	3·75	2·96	13·14	·435
4·00	19·25	16·76	15·28	24·25	3·58	2·68	11·69	·303
2·22	11·92	19·83	15·96	31·53	4·01	2·53	10·16	·267
3·10	16·96	19·66	13·21	26·50	6·25	2·40	8·89	·198

It will be seen from the values obtained above that with the exception of the fine gravel p_1 is about 2·5 per cent. as a maximum[1]. These figures are only obtained from six sets of determinations, but they are sufficiently accurate for the purpose of determining p_2, as will be seen later.

The table on p. 147 shows the results of the analyses of the 25 samples from Field A.

The mean values and probable errors calculated from these results are as follows:

	Per cent.	Probable error per cent. of result, P
Fine gravel	7·92 ±2·78	=±35·0 per cent.
Coarse sand	17·59 ±2·24	=±12·7 "
Fine sand	17·61 ±1·64	=± 9·3 "
Silt	12·44 ±1·24	=±10·0 "
Fine silt	26·50 ±2·71	=±10·2 "
Clay	4·10 ± ·52	=±12·7 "
Moisture	2·7 ± ·40	=±14·8 "
Organic matter	10·4 ± ·80	=± 7·7 "
Phosphorus pentoxide (P_2O_5)	·290± ·039	=±13·4 "

In Field B a set of 15 borings was taken as shown on the diagram Fig. 2.

The following results were obtained on analysis:

Aberffraw.

Fine Gravel	Coarse sand	Fine sand	Silt	Fine silt	Clay	Moisture	Organic matter	P_2O_5
4·45	43·39	18·04	8·52	12·91	3·45	2·16	5·54	·220
5·73	44·12	16·57	8·42	12·83	3·20	2·15	5·32	·209
6·13	40·79	16·23	8·81	13·95	3·15	2·42	6·10	·249
4·15	40·41	16·80	8·55	14·55	3·50	1·33	6·78	·214
7·57	42·30	16·44	8·78	14·15	3·20	1·46	5·53	·220
5·58	42·05	16·59	9·20	14·15	3·01	2·14	5·21	·204
6·05	38·97	18·37	10·06	14·86	3·64	2·09	5·22	·188
5·24	47·50	14·39	8·70	13·60	3·14	2·02	5·30	·199
3·37	46·74	16·23	9·53	12·85	3·30	2·00	5·26	·223
4·50	47·36	16·02	8·58	12·34	3·25	1·97	4·74	·264
4·95	43·56	17·11	8·98	12·86	3·82	1·94	5·23	·217
6·64	40·85	17·00	9·65	14·16	3·36	2·00	4·76	·179
4·77	43·43	16·98	8·12	13·01	3·36	2·13	5·17	·269
4·02	46·25	16·42	7·53	12·82	3·02	1·75	5·25	·229
5·40	47·46	16·56	7·35	11·49	3·17	2·06	4·99	·213

[1] In mechanical analyses this error will be much greater if the temperature of sedimentation is not constant, owing to viscosity changes.

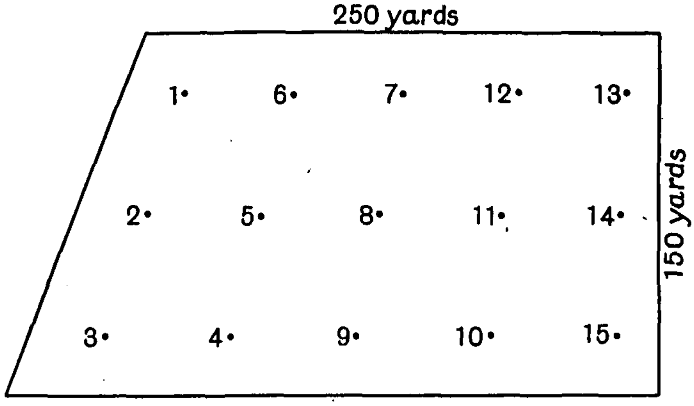

Fig. 2. Field B.

The mean values and probable errors calculated from these results are as follows:

	Per cent.	Probable error per cent. of result, P
Fine gravel	5·23 ± ·73	±13·9 per cent.
Coarse sand	43·68 ±1·76	± 4·0 ,,
Fine sand	16·65 ± ·45	± 2·7 ,,
Silt	8·73 ± ·48	± 5·5 ,,
Fine silt	13·37 ± ·60	± 4·5 ,,
Clay	3·30 ± ·13	± 3·9 ,,
Moisture	1·97 ± ·18	± 9·2 ,,
Organic matter	5·36 ± ·235	± 4·4 ,,
Phosphorus pentoxide (P_2O_5)	·220± ·017	± 7·7 ,,

The results obtained may now be stated:

Field A. Drift Soil.

Determination made	P (per cent.)	p_1 (per cent.)	p_2 $(=\sqrt{P^2 - p_1^2})$ (per cent.)
Fine gravel	35·0	6·3	34·5
Coarse sand	12·7	3·0	12·3
Fine sand	9·3	1·3	9·2
Silt	10·0	2·6	9·7
Fine silt	10·2	2·1	10·0
Clay	12·7	2·3	12·4
Moisture	14·8	v. small since mean of 4 was taken	14·8
Organic matter	7·7		7·7
P_2O_5	13·4	2·5	13·2

Field B. Sedentary Soil.

Determination made	P (per cent.)	p_1 (per cent.)	$p_2 (= \sqrt{P^2 - p_1^2})$ (per cent.)
Fine gravel	13·9	6·3	12·4
Coarse sand	4·0	3·0	2·6
Fine sand	2·7	1·3	2·4
Silt	5·5	2·6	4·8
Fine silt	4·5	2·1	4·0
Clay	3·9	2·3	3·1
Moisture	9·2	v. small since mean of 4 was taken	9·2
Organic matter	4·4		4·4
P_2O_5	7·7	2·5	7·3

Practically the same values for p_2 might have been obtained if analyses had been repeated several times for each boring and averaged. The values of p_1 would then have been greatly reduced and consequently, as in the case of the moisture and organic matter, p_2 could be taken as approximately equal to P.

In view of the great time taken in making soil analyses it was felt that the increase in the accuracy of the values of p_2 which would thus be obtained would scarcely be great enough to justify spending four or five times as long over the laboratory work. Also, no time was spent in doing chemical analyses other than the determination of P_2O_5.

It is seen from the above tables that in the case of Field A, which is too variable for ordinary survey purposes, p_2 is as great as 34·5 per cent. for fine gravel. Now this constituent is comparatively unimportant in characterising a soil unless it predominates over other fractions. An accurate knowledge of the hygroscopic moisture is also, for survey purposes, of no great importance. Of the other determinations the probable field errors of the fractions in the mechanical analysis are about 10 per cent., while that of the P_2O_5 is 13·4 per cent.

If therefore we count on a probable error of 13·4 per cent. for this soil this can be taken as a maximum value. Similarly 7·3 per cent. can be taken as the maximum value of the field error for Field B which is a sedentary soil. Field A is as stated above of too variable a character for ordinary survey purposes. If therefore operations are conducted on the basis of a maximum field error of 10 per cent. it will probably serve for survey work.

The value of the total probable error of the final result can now be reduced to any desired value by increasing the number of borings and the number of determinations.

Thus if n borings are made and one determination made on each boring the probable error of the average is

$$\pm \frac{\sqrt{p_1^2 + p_2^2}}{\sqrt{n}} \text{ or } \frac{P}{\sqrt{n}}.$$

For example, in the case of P_2O_5 in Field A, 9 borings separately analysed and averaged would give a probable error of

$$\pm \frac{13 \cdot 4}{\sqrt{9}} = \pm 4 \cdot 46 \text{ per cent.}$$

If n borings are made and a composite sample obtained by mixing, the field error will be reduced to $\frac{p_2}{\sqrt{n}}$, but the laboratory error will still be present, so that if 9 borings from Field B were mixed, the error of one analysis would be

$$\pm \sqrt{2 \cdot 5^2 + \left(\frac{13 \cdot 2}{\sqrt{9}}\right)^2}, \text{ or } \pm 5 \cdot 05 \text{ per cent.}$$

Where the field error is large in comparison with the laboratory error not much additional accuracy is obtained by analysing the borings separately. Also, however many borings are taken and mixed, it is not possible to reduce the final error to less than the laboratory error; and it is also somewhat difficult to ensure the satisfactory mixing of a very large number of borings.

Performing more than one analysis on a composite sample will serve to reduce the laboratory error, and consequently the final error of result.

If the field error for chemical analyses be taken as ± 10 per cent. and that for the mechanical analyses as ± 5 per cent., we shall probably not err on the side of minimising errors, since no field is sampled for survey purposes unless it appears fairly uniform. As is mentioned above, Field A would not be considered uniform enough for survey purposes.

The following table shows the probable error of final result for different numbers of borings and for single and duplicate analyses of the mixed sample. The probable errors are calculated on the assumption that the final probable error is equal to the square root of the sum of the squares of the separate probable errors. Thus the probable error of an average of two determinations on a composite sample of 5 borings is in the case of mechanical analyses equal to

$$\pm \sqrt{\left(\frac{2 \cdot 5}{\sqrt{2}}\right)^2 + \left(\frac{5}{\sqrt{5}}\right)^2} = \pm 2 \cdot 8 \text{ per cent.,}$$

and in the case of a chemical analysis,

$$\pm\sqrt{\left(\frac{2\cdot5}{\sqrt{2}}\right)^2+\left(\frac{10}{\sqrt{5}}\right)^2}=\pm\,4\cdot8 \text{ per cent.}$$

Number of borings taken and mixed to form composite sample	Final probable error, per cent.			
	Mechanical analysis ($p_2 = \pm$ 5 per cent.)		Chemical analysis ($p_2 = \pm$ 10 per cent.)	
	Single analysis	Average of two analyses	Single analysis	Average of two analyses
1	±5·6	±5·3	±10·3	±10·1 %
2	±4·3	±3·95	± 7·5	± 7·3
3	±3·8	±3·5	± 6·3	± 6·0
4	±3·5	±3·05	± 5·6	± 5·3
5	±3·3	±2·8	± 5·1	± 4·8
6	±3·2	±2·7	± 4·8	± 4·4
7	±3·1	±2·6	± 4·5	± 4·2
8	±3·05	±2·5	± 4·3	± 3·95
9	±3·0	±2·4	± 4·15	± 3·75
10	±2·95	±2·35	± 4·0	± 3·6

The table shows that even with 10 borings there is not much diminution in the probable error by taking an average of two analyses. Since however some check is advisable it is better to make two analyses.

When two mechanical analyses are performed it is seen that not much additional accuracy is obtained by increasing the number of borings beyond six, either in the mechanical or chemical analyses. Taking six as the number of borings the probable error is ± 2·7 per cent. This means that it is an even chance that a result obtained is within 2·7 per cent. of the true result. The chances of a result having lower degrees of accuracy are shown in the following table, calculated from Wood and Stratton's paper on the Interpretation of Experimental results[1].

Percentage deviation from true value	Chances of result being within limits of such deviation
2·7 per cent.	Evens
3·4 ,,	3 to 2
3·9 ,,	2 to 1
4·6 ,,	3 to 1
5·1 ,,	4 to 1
5·4 ,,	9 to 2

[1] This *Journal*, Vol. III. p. 429.

Similarly in the case of a chemical analysis, we have for six borings:

Percentage deviation from true value	Chances of result being within limits of such deviation
4·4 per cent.	Evens
5·5 ,,	3 to 2
6·3 ,,	2 to 1
7·5 ,,	3 to 1
8·3 ,,	4 to 1
8·8 ,,	9 to 2

From these results it is seen that an accuracy of 5 per cent. in mechanical analysis is ensured a probability of 4 to 1 by doing a duplicate analysis on six borings. For survey purposes this is probably sufficient, since it is not conceivable that variation in the amount of any fraction corresponding to 5 per cent. (relative to the amount of the fraction) could have any effect on the properties of a soil.

In the case of chemical analysis it does not seem that the same accuracy can be expected. For survey purposes however the accuracy is probably sufficient. In the case of a critical study of one soil however it would be necessary to reduce the errors much more by repeating analyses and increasing the number of borings.

(*Received March 3rd,* 1915.)