The accuracy of analytical data of glasshouse soil testing

C. Sonneveld

Glasshouse Crops Research and Experiment Station, Naaldwijk, the Netherlands

Summary

At intervals over a period of several years, duplicate soil samples were collected from glasshouses which had been sampled by a professional sampler the day before. The samples were analysed in duplicate by means of the 1 : 2 volume extract method.

The analytical data were classified into groups according to the levels found. The mean values and the standard deviations arising from the laboratory work and the soil sampling operations were computed. A close correlation was found between the mean values and the standard deviations. The standard deviation for soil sampling was several times greater than the standard deviation for the laboratory analyses. Total standard deviations of more than 20 % were found for some determinations.

The investigation has proved that systematical errors in the sampling method will affect the soil testing results. Therefore, it is of great importance that very precise instructions, adapted to crops and growing systems, are issued to the sampler.

Introduction

The accuracy of analytical soil testing data depends on the degree of accuracy with which soil sampling and analytical determinations are carried out. The errors made in soil sampling are generally greater than those made in laboratory analyses (Vermeulen, 1960a; Peck & Melsted, 1973; Cline, 1944).

Factors like the number of cores per sample, the uniformity of the soil and the working methods of the sampler all have a great effect on the results. It can be demonstrated statistically that the error in sampling is subject to a decrease which is inversely proportional to the square root of the number of cores used to make up the sample. The relationship between the error in sampling and the number of cores is expressed by the following equation $s_p = s \cdot p^{-1/2}$ in which s_p is the standard deviation resulting from sampling with p cores per sample, p is the number of cores and s is the standard deviation resulting from sampling with one core per sample.

Because of the risk of error, the sampler in the Netherlands is instructed to collect at least 40 randomized cores per sample (Vermeulen, 1960b). More cores per sample involve more labour and produce only a relatively slight decrease in the sampling error.

Knowledge of the variability of the soil is a very important factor in sampling. Systematic components in this area have to be looked for and they should be included in the instructions to the sampler. Precise instructions to the sampler must be drawn up and these should be observed in order to standardise sampling.

The results of investigations into systematic deviations in glasshouse soil sampling will be discussed in this paper. Also, information will be given about the degree of errors in analytical data obtained in soil testing on a practical scale.

Methods

In order to advise growers on fertilizer requirements, about 50 000 glasshouse soil samples are analysed in the laboratory for soil testing at the Glasshouse Crops Research Station te Naaldwijk every year. Most of the samples are collected by professional samplers.

At intervals, an inspector collected samples from glasshouses which had been sampled by a professional sampler the day before. The soil samples were analysed in duplicate by means of the 1 : 2 volume extract method (Sonneveld & van den Ende, 1971). The electrical conductivity (EC), chloride (Cl), nitrogen (N), phosphate (P), potassium (K) and magnesium (Mg) were determined in the extract. The EC was expressed as mS/cm at 25 °C, phosphate as mg/litre and the other elements as meq./litre. At the time of sampling, information about the soil type and the crop grown in the glasshouse was also recorded.

In this way, the analytical data of 400 glasshouses were collected and divided into two groups. The data of 200 glasshouses sampled during the years 1975 and 1976 were included in one group and the data collected during the years 1977 and 1978, were included in the other group. The dividing into groups was carried out, because generally it is not advisable to make computations with results gathered over long periods.

The four figures of every determination per glasshouse obtained in the laboratory will be marked as follows:

$x_{1.1}$ and $x_{1.2}$ are the values of samples collected by the professional sampler;

$x_{2.1}$ and $x_{2.2}$ are the values of samples collected by the inspector.

The following denotations will be used with the statistical computations:

$x_1 = (x_{1.1} + x_{1.2})/2$ and $x_2 = (x_{2.1} + x_{2.2})/2$

$d_{1.12} = x_{1.1} - x_{1.2}$ and $d_{2.12} = x_{2.1} - x_{2.2}$

$d_{12} = x_1 - x_2$

n = the number of glasshouses in a group

$s_t = [(\Sigma d_{12}^2)/2n]^{1/2}$, which is the total standard deviation made in soil testing

$s_a = [(\Sigma d_{1.12}^2 + d_{2.12}^2)/4n]^{1/2}$, which is the standard deviation made in laboratory analyses

$s_s = (s_t^2 - \frac{1}{2} s_a^2)^{1/2}$, which is the standard deviation made in soil sampling.

298

Standard deviations expressed as percentages will be denoted as C_t, C_a and C_s respectively.

Mean values of a group of samples are denoted as M and mean differences between diplicate results as $\bar{d}$.

Results of statistical calculations

The analytical data of the glasshouses were classified into ten groups on the basis of the mean values for the different determinations. The twenty lowest were included into group 1, the twenty next lowest into group 2 and so on. The mean M and the standard deviations s_a and s_t of every group were calculated. Next, every group was checked for values of $d_{1.12}$ and $d_{2.12}$ greater than $3 \cdot 2^{1/2} \cdot s_a$ and values of d_{12} greater than $3 \cdot 2^{1/2} \cdot s_t$. These values were marked as outliers and deleted from the group. Only some analytical data had to be removed and new values for mean and standard deviations were calculated for the groups concerned.

A close linear correlation was found between the mean values and the standard deviations for all determinations. The regression equations found are given in Tables 1 and 2. An example of the scatter diagrams is shown in Fig. 1. In the 1977-1978 group of equations, the regression coefficient for phosphate is exceptionally high and the intercept has a great negative value. This is impossible and is

Table 1. The relationships between the mean values of the analytical data of groups of samples (M) and the total standard deviations (s_t).

Determi-nation	1975 - 1976		1977 - 1978	
	regression equation	r	regression equation	r
EC	$s_t = 0.120 \, M + 0.01$	0.79	$s_t = 0.100 \, M + 0.06$	0.76
Cl	$s_t = 0.116 \, M + 0.06$	0.90	$s_t = 0.154 \, M + 0.03$	0.93
N	$s_t = 0.227 \, M + 0.02$	0.97	$s_t = 0.174 \, M + 0.08$	0.89
P	$s_t = 0.205 \, M + 0.16$	0.93	$s_t = 0.303 \, M - 0.29$	0.98
K	$s_t = 0.176 \, M - 0.01$	0.94	$s_t = 0.246 \, M - 0.06$	0.98
Mg	$s_t = 0.188 \, M + 0.05$	0.92	$s_t = 0.187 \, M + 0.06$	0.89

Table 2. The relationships between the mean values of the analytical data of groups of samples (M) and the standard deviations arising from laboratory analyses (s_a).

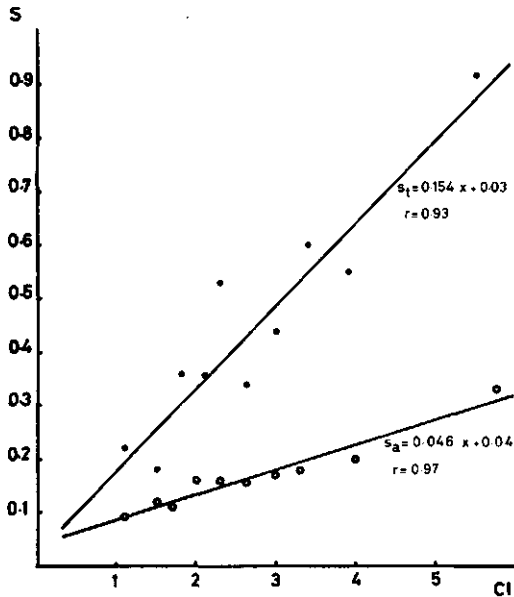| Determi-nation	1975 - 1976		1977 - 1978	
	regression equation	r	regression equation	r
EC	$s_a = 0.044 \, M + 0.01$	0.88	$s_a = 0.052 \, M + 0.01$	0.84
Cl	$s_a = 0.038 \, M + 0.09$	0.89	$s_a = 0.046 \, M + 0.04$	0.97
N	$s_a = 0.044 \, M + 0.11$	0.94	$s_a = 0.062 \, M + 0.05$	0.98
P	$s_a = 0.043 \, M + 0.18$	0.86	$s_a = 0.070 \, M + 0.05$	0.96
K	$s_a = 0.057 \, M + 0.05$	0.69	$s_a = 0.059 \, M + 0.02$	0.98
Mg	$s_a = 0.043 \, M + 0.11$	0.89	$s_a = 0.070 \, M + 0.03$	0.95

Fig. 1. The relationships between the chloride content (meq/litre) of the 1:2 volume extract and the standard deviations s_t (total) and s_a (analyses). Results of the 1977-1978 group.

caused by the fact that a significant value for $\bar{d}_{12}$ is found in all groups, but especially for the highest levels. This increases the value of the regression coefficient and decreases the value of the intercept. After the values of $\bar{d}_{12}$ of the groups of phosphate determinations had been corrected, the following regression equation could be calculated:

$s_t = 0.217 \, M - 0.03; r = 0.98.$

This equation comes up better to expectations and agrees rather well with the equation found for phosphate in the years 1975 and 1976.

The conclusion may be drawn from Table 1 that the total standard deviation for EC and Cl tends to be in the range of 10 to 15 %. The standard deviations for other determinations are higher, reaching values of more than 20 %. The standard deviations resulting from laboratory analyses are lower, i.e. values of between 5 and 10 %. The contribution of s_a to s_t is only a factor $2^{-1/2}$ times the values given in Table 2, because the soil samples were analysed in duplicate and because the mean values of the duplicates were used in the calculation of s_t.

The value of s_s may be calculated from Tables 1 and 2, as well as an estimation of the contribution to s_t by the standard deviation resulting from soil sampling and the standard deviation resulting from laboratory analyses. The ratio between the two contributions may be calculated by using the formula $s_s \cdot s_a^{-1} \cdot 2^{1/2}$. The value of this product roughly varies between 2 and 5. This shows that the errors made in soil sampling are considerably greater than those made in laboratory analyses.

Besides the classification of analytical data into groups according to the levels found, the data were also classified according to soil types, crops and samplers. With respect to soil types and crops, no significant differences were found in the

300

Table 3. The total standard deviations (s_t) and the standard deviations as percentages (c_t) calculated for five soil samplers.

Samplers	EC		Cl		N		P		K		Mg	
	s_t	c_t	s_t	c_t	s_t	c_t	s_t	c_t	s_t	c_t	s_t	c_t
A	0.16	11	0.37	12	0.79	22	1.3	27	0.29	16	0.31	13
B	0.25	18	0.51	19	0.64	20	1.9	32	0.56	29	0.70	28
C	0.15	10	0.39	15	0.53	16	1.6	33	0.38	22	0.40	18
D	0.29	20	0.66	23	1.10	30	2.0	33	0.65	32	0.81	28
E	0.19	15	0.48	20	0.57	18	1.3	21	0.26	16	0.63	24

values of s_t. However, there were significant differences between the soil samplers. The values of s_t calculated from the analytical data of 1977 and 1978 are shown in Table 3. The values of s_t are nearly always highest for samplers B and D. It would appear that these samplers have worked less accurately than the others.

Not only casual errors will affect the value of s_t, but also systematic errors in the sampling method, in which case significant values for $\bar{d}_{12}$ will be found. This was the case especially in the phosphate determinations. The values of $\bar{d}_{12}$ found for the five samplers are given in Table 4. Positive values were found for all samplers. Thus on average, the phosphate content of the soil samples taken by the samplers was higher than that of the samples collected by the inspector. Systematically differences between the samples taken by samplers and those taken by the inspector were also found in the case of other determinations. However, the latter were less evident than the differences in the phosphate contents.

The most striking differences between the samplers and the inspector were found by calculating the deviation of each sampler for the same crop. Table 5 shows the values of $\bar{d}_{12}$ established for the five samplers, together with the mean values of the determinations of the samples collected in tomato houses. The data were obtained from the samples gathered in the years 1977 and 1978 and for each sampler there were 10 to 15 duplicate samples available from tomato houses. As may be seen, sampler C shows the least deviation from the inspector and sampler B shows the

Table 4. The mean differences ($\bar{d}_{12}$) between the analytical data of the phosphate determinations of the soil samples taken by the samplers and the inspector.

Samplers	Data 1975 - 1976		Data 1977 - 1978	
	M	$\bar{d}_{12}$	M	$\bar{d}_{12}$
A	5.0	+ 1.30	5.0	+ 0.86
B	5.6	+ 1.14	6.0	+ 1.52
C	5.3	+ 0.22	5.0	+ 0.82
D	5.9	+ 0.78	6.0	+ 1.24
E	5.1	+ 1.26	5.9	+ 0.67

Table 5. Mean values of the differences between samples collected by the samplers and by the inspector in tomato houses.

Samplers	EC		Cl		N		P		K		Mg	
	M	$\overline{d}_{12}$	M	$\overline{d}_{12}$	M	$\overline{d}_{12}$	M	$\overline{d}_{12}$	M	$\overline{d}_{12}$	M	$\overline{d}_{12}$
A	1.6	+0.11	2.9	+0.25	3.9	+0.75	5.2	+1.4	2.1	+0.33	2.7	+0.24
B	1.5	+0.31	2.5	+0.29	3.6	+0.91	6.8	+2.8	2.4	+0.87	3.0	+0.90
C	1.7	—0.00	2.9	—0.03	4.4	+0.17	5.5	+1.1	2.4	+0.26	2.9	—0.10
D	1.9	+0.15	2.9	+0.19	4.8	+0.43	7.1	+1.8	3.0	+0.46	4.1	+0.58
E	1.5	+0.03	2.3	—0.10	4.7	+0.40	7.8	+1.1	2.4	+0.19	3.3	+0.17

greatest deviation. For the phosphate determinations all samplers displayed great deviation.

Discussion

In the testing of glasshouse soils, relatively great deviations in the analytical data obtained have to be taken into account. These deviations may be ascribed in particular to the soil sampling methods. The total standard deviations found in soil testing easily amounted to 10 to 15 % and could reach values of more than 20 %. The highest values were found for the determinations of nitrogen, phosphate, potassium and magnesium. This is understandable since the contents of these elements in the soil are affected strongly by fertiliser applications which, as much as anything, lead to uneven distribution of the elements in the soil (Peck & Melsted, 1973; Gallager & Herlihy, 1963; Little & Tong Kwong Yuen, 1973; James & Dow, 1972; Leo, 1963).

A standard deviation of about 10 % in analytical soil testing results should be acceptable and should still provide a reasonably sound basis for advice to growers on fertiliser requirements. The most obvious solution to achieve a reduction in a high standard deviation would be to increase the number of sites from which the sample is taken. However, in order to reduce C_t from 20 % with 40 cores per sample to 10 %, an increase in the number of cores per sample would be required to

$$40 \left[0.20^2 - (C_a \cdot 2^{-1/2})^2 \right] / \left[0.10^2 - (C_a \cdot 2^{-1/2})^2 \right] = 186$$

on condition that $C_a = 0.06$ and the samples were analysed in duplicate like in our investigation. This would lead to an unacceptable increase in the labour required for soil sampling.

This investigation has shown that soil sampling errors are caused not only by coincidence, but also by systematic deviations. This is evident from the analytical data of the samples collected by the samplers which are consistently higher than those of the samples taken by the inspector. The explanation is that in collecting the sample, the inspector will take the cores near the centre of the growing bed, as recommended in the instructions. The pressure of work on the profession sampler is such that he will collect the cores closer to the better attainable edge of the grow-

Table 6. Average values of analytical data of soil samples collected from different layers in three glasshouses shortly after fertilization.[1]

Depth (cm)	Analytical data			Relative value		
	N	P	K	N	P	K
0– 8	5.2	13.7	4.2	100	100	100
8–16	4.4	8.4	2.9	85	61	69
16–24	2.9	4.2	1.5	56	31	36
24–40	2.1	1.4	1.0	40	10	24

[1] Unpublished data, Glasshouse Crops Research and Experiment Station, Naaldwijk, the Netherlands.

ing bed where salt concentrations are usually higher than in the middle (van den Ende & Knoppert, 1959). Another explanation for the difference in the analytical data may be found in the sampling depth. If the professional samplers do not push in the auger deeply enough, the result would be an increase in the levels of the analytical data. The salt content of soils generally decreases with sampling depth (James & Dow, 1972; Cameron et al., 1971), particularly in the case of phosphate in glasshouse soils (Hamaker & van den Burg, 1978; unpublished data; Table 6). This offers an explanation in particular for the differences in phosphate contents found.

It is clear that very precise instructions for soil sampling are of great importance. The instructions should be adapted in detail to the different crops and growing systems. In that way, it would be possible to decrease the deviations caused by soil sampling errors.

At present, standard deviations of between 10 and 20 % have to be taken into account in glasshouse soil testing. This means that the analytical data do not give a close estimation of the true average values of the soil sampled. For many determinations, at a level of significance of k %, the following confidence interval is applicable:

$$x - u \cdot 0.2\,x < \mu < x + u \cdot 0.2\,x$$

in which x = the result of a determination, u = the standard normally distributed unit going with a level of significance of k % and μ = the true value of the determination in the soil sampled. In the furnishing of advice on fertilizer requirements based on the results of routine soil testing, insufficient attention is being given to the deviations which should be taken into account. At a 5 % level of significance, the value of u amounts to as much as 2 and for certain determinations a deviation of 40 % belongs to the confidence interval.

References

Cameron, D. R., M. Nyborg, J. A. Toogood & D. H. Laverty, 1971. Accuracy of field sampling for soil tests. Can. J. Soil. Sci. 51: 165-175.
Cline, M. G., 1944. Principles of soil testing. Soil Sci. 58: 275-288.

Ende, J. van den & J. P. C. Knoppert, 1959. Grondonderzoek ten behoeve van het bijmesten. *Groenten en Fruit* 14: 769.

Gallager, P. A. & M. Herlihy, 1963. An evaluation of errors associated with soil testing. Irish. *J. Agric. Res.* 2: 149-167.

Hamaker, Ph. & A. A. M. van der Burg, 1978. De water- en mineralenbalans van een glastuinbouwbedrijf in de periode 1976/77. Nota 1072, Instituut voor Cultuurtechniek en Waterhuishouding, Wageningen.

James, D. W. & A. J. Dow, 1972. Source and degree of soil variation in the field: the problem of sampling for soil tests and estimating soil fertility status. *Bull. Wash. agric. Exp. Stn* No 749.

Leo, M. W. M., 1963. Heterogeneity of soil agriculture land in relation to soil sampling. *Agric. Food. Chem.* 11: 432-434.

Little, R. C. & L. Tong Kwong Yuen, 1973. Variation in nutrient content of glasshouse soils. *Exp. Hort.* 25: 102-111.

Peck, T. R. & S. W. Melsted, 1973. Field sampling for soil testing: Soil testing and plant analysis. Soil Science Society of America, pp. 67-75.

Sonneveld, C. & J. van den Ende, 1971. Soil analysis by means of a 1:2 volume extract. *Pl. Soil* 35: 505-516.

Vermeulen, F. H. B., 1960a. Fehlerquellen bei der Bodenuntersuchung. Stand und Leistung agrikulturchemischer und agrarbiologischen Forschung. VII, 80-85.

Vermeulen, F. H. B., 1960b. Bemonstering van de grond voor het chemisch bodemvruchtbaarheidsonderzoek. *TNO-Nieuws* 167: 20-22.