Offprinted from Trans. Comm. II & IV, Int. Soc. Soil Sci., Aberdeen 1966

PRINCIPLES OF EXPERIMENTATION IN SOIL-CROP STUDIES

PRINCIPLES OF EXPERIMENTATION IN THE STUDY OF SOIL-CROP RELATIONSHIPS

Th. J. Ferrari

Institute for Soil Fertility, Groningen, The Netherlands

INTRODUCTION

To prevent misunderstandings we start with the definition of what will be meant by the study of soil-crop relationships: all intellectual and other activities leading to a better understanding of the interdependence between soil factors and plant growth under farm conditions and also of the control of growth by manipulating those factors. Biological, chemical and physical research necessary to study these ecological factors will be left out of account.

In soil-fertility research the field experiment has always been an important instrument. The first one was laid out in 1834 and was the beginning of a long series of pot and field experiments by which investigators tried to get an insight into soil-crop relationships. Their results have been rather disappointing. The common experience, indeed, has been that the results of their investigations into soil-fertility problems may diverge widely, so that different researchers reach different conclusions about the same phenomenon and the generalization of the results of research is a difficult thing. Bradfield (1961) is the exponent of many complaints by writing: 'It seems to me that we cannot escape the conclusion that the reason we continue to investigate and re-investigate certain aspects of soil fertility is that we are not yet able to predict with confidence what will happen in the case we are interested in, because there are too many parameters involved which we are still unable to identify, measure or interpret.'

It is remembered that the practitioners of agricultural research in the early period were mostly people with a biological, chemical or physical training. Their specialized knowledge was accompanied by the methods which were accepted as the correct ones of their sciences; an agricultural science aware of its own problems did not exist yet. In the natural sciences it has been found particularly advantageous to test a hypothesis against data obtained artificially by changing one factor and keeping the other ones constant (manipulative control and assumption of *ceteris paribus*). In this beginning of agricultural research there was a firm conviction that it would be possible to carry out in soil-fertility research, too, the ideal experiment in a completely isolated system. Soil-fertility problems were attacked as if they were physical or chemical ones. The experimental field is based on this idea.

In connection with the disappointing experiences the question is justified whether the method used has been the best and only one to study and solve soil-fertility problems. In the following we shall discuss the concepts that can be made about the relationships between soil factors and plant growth on the basis of our experiences, the methods to study these relationships and the difficulties connected with this investigation.

PRINCIPLES OF THE STUDY OF MULTI-DIMENSIONAL SYSTEMS

The rationale of research is founded on the assumption that a change of a so-called causal variable is always followed by a certain change of one or more other variables. Hereupon again the idea of the ideal experiment rests: if the variable x is a cause of y and if it is possible to keep constant all other causes of y (isolation), an artificial change of x will be followed by a certain change of y. I take this rule as the starting point of my discussion although I am aware of the fact that the concept of causality is treated incompletely here.

For a good understanding of the value of this method it is necessary to distinguish clearly the several elements of this rule. Firstly the researcher has to deal with the hypothesis of the causal relationship between the variables; this hypothesis can be developed more or less. Next he has deductively to infer conclusions the correctness of which must be tested. Such conclusions often concern the change of a dependent variable by the change of the cause. Finally the correctness of the conclusion is tested by a confrontation with reality.

The cardinal point is the correspondence of prediction and reality. The lack of this correspondence rejects the correctness of the hypothesis, while an agreement makes the hypothesis more acceptable. A proof in the mathematical sense, however, can never be reached as one never knows if other variables have not had an influence. From this it also follows that the characteristic of the experiment is not the technical manipulation, but the empirical verification. This test can also be carried out against data obtained without interference. However, this interference has some advantages with a view to the isolation of the system and the conclusions to the asymmetry of the relationships.

The rule just mentioned also says that the correctness of the method holds only if all other possible causes of y are kept constant and could not have caused the changes of y. The researcher tries to eliminate these influences by studying the hypothetical relationship isolated from all other variables. This isolation can be obtained not only by manipulation and experimental design, but also by treatment of the observations. The first method is applied especially in the natural sciences, while economics and sociology use the second one more.

Relatively little attention has been given to the problem that the interference itself may be the cause of changes of variables assumed to remain constant, which changes in their turn can produce changes in the dependent variable. The necessary isolation is abolished in this case, and the difficulty is that the whole of the changes cannot be analysed any more (identification problems). It may be assumed that one of the most important causes of the differences mentioned between the several investigations is connected with this problem; the interference violates the important assumption of the constancy of the other variables.

In soil-fertility research the problem exists that the correctness of the assumed isolation and constancy not only must often be doubted, but also cannot be proved. One has to do with many variables connected with each other in different ways. Therefore the question always remains if variables left out of consideration have produced the result after all. This means that every causal conclusion always has a certain element of risk. We shall see later that an aspect analysis offers a possibility to reduce this risk. In this respect a physicist often has an easier task; in classical physics an isolated system can be reached more easily at which the influence of disturbing variables can be assumed as small. This is the reason that the experiment with interference has given so many good results in this science.

What kind of variables has the researcher to take into account? To get a good survey the variables producing a change of the dependent variable will be grouped into four categories (Blalock 1961). To the first group belong the variables which the researcher has to test on their causal influence. It may be one factor, but more variables can also be investigated at the same time. The changes can be produced by the researcher, but he may also use the already existing variation (cf. astronomy).

In the second category are classed the variables that might have influenced the dependent variables. They did not produce any effect under the circumstances of the experiment, as the experimenter worked under conditions controlled for these variables, or as these variables just happened not to vary during the course of the experiment. The feature of this group is thus that its variables have caused no variation in the dependent variables. However, we must not forget that these variables may influence the dependent variable under other circumstances. The researcher is inclined to underestimate the importance of these variables in his generalization. Besides, the rules for experimental design always try to change the variables of the next categories into variables of this group. This leads unconsciously to an underestimate.

The third category includes variables which the investigator is not able to control and the changes of which produce changes of the dependent variables, too. These changes are unrelated to those of the independent variables being investigated. There is therefore no danger to ascribe the changes of the dependent variables to others than the assumed causal variables. The variations caused by the variables of this category are used to calculate the statistical significance of the influence of the causal variables.

The fourth category consists of variables that also have produced changes in the dependent variables. The effects are in some way systematically related to those of the independent variables of the first category. The influences of the variables of the first and fourth category are confounded, and can be separated only with difficulty at which the risk exists to make inference on non-existing causality.

The researcher must always realize which variables will be dealt with in the experiment, how large the influence of the several disturbing variables may be, etc. Statisticians, on the other hand, always try to find new methods to eliminate the influences of the second-, third- and fourth-category variables and to estimate these influences. Mathematical statistics has played an important role in this respect, but has led also to an one-sidedness in the use of scientific methods.

It is clear that there is the least risk of a wrong inference if one is confronted only with variables of the first category and/or if the effects of variables of the other categories are small. We have already seen that the researcher of soil fertility meets great difficulties in this respect. It is true, he can try to reach with the help of controlled pot and field experiments, that a large number of variables will belong to the second category. However, the question arises whether there are great possibilities of a generalization of the results in this case. I shall return to this point later.

The great attention given in agricultural research to the experiment with interference and to the elimination of the influences of the second and third category variables by means of statistical methods and experimental design is presumably the cause that the importance of the fourth-category variables has been underestimated. It is undoubtedly true that randomization in the design can change some variables of the fourth into those of the third category. However, it is forgotten that the interference itself changes a number of second- and third-category variables into fourth-category variables. As we have seen, the interference not only produces the intended changes of the dependent variables, but also changes of variables assumed to be kept constant. The importance of an interference is always overestimated, the drawback of it is however, underestimated! The researcher does not realize usually this effect, and accepts the between-trial variation originated by this as inevitable and unexplainable. An awareness of its indirect effect could probably prevent these variations.

The variables of the fourth category play an important role in experiments without interference at which changes are not produced artificially. It is difficult to get an isolated system in this case. We often have to use this kind of research in the study of soil-fertility relationships. Many variables can not or hardly be changed: structure, humus and silt content, profile, groundwater level, etc. An experiment with interference is impossible in this case.

The economist is often in similar circumstances, and has given from the beginning much attention to the solution of the problem connected with this kind of non-experimental research. In spite of this we are confronted with the fact that in the greater part of research attention has been given to subjects belonging to the method with interference, significance tests and experimental design. Wold (1955) once wrote: 'A fourth broad area remains—explanation on the basis of observational data. In this field, which embraces amongst other things a large proportion of social research, progress has been less systematic and spectacular. Current text-books reveal the stepchild treatment which this sort of problem has received from professional statisticians.' The same holds for soil-fertility research. The extensive application of the analysis of variance together with the manipulative experiment has notably delayed the practical importance of soil-fertility research (Ferrari 1966).

MODELS IN SOIL-FERTILITY RESEARCH

Confrontation with the difficulties just mentioned is quite possible, using models representing the reality to be investigated. A model can be considered as an abstracted simplification of this reality in which only elements significant for the problem and already familiar by other investigations are taken. This abstraction is expressed in a language: in words or in diagrams, mathematically or materially. The researcher tries to make this abstraction as simple as possible without, however, leaving out important elements. The main problem remains to select models that are at the same time simple enough to permit the researcher to carry out experiments, but also sufficiently realistic not to lead to conclusions that are highly inaccurate. This is a very important point as a hypothetic model can only be tested negatively on its value.

The significance of the use of models should on no account be underestimated in this complex area of soil fertility. The researcher who does not make a picture of the reality in advance has no notion of the variables he will meet; he does not realize which variables of the three categories are important. Neither can he take the best adapted measures in the experimental design nor realize what consequences an interference will have. Necessary observations will be omitted.

In the following we shall discuss some models that are, or can be, used for the description of soil-fertility problems. This will be done on the basis of diagrams and mathematical equations. The use of mathematical functions to describe complex phenomena has great advantages. The elements themselves are not considered, that being the task of chemical, physical and biological research. The relationships can be represented by one or by a system of equations. The problems met in testing the various models will be discussed.

(a) Models with one equation, with one or more variables

The simplest model is of course the assumption that the dependent variable is only affected by one variable without the use of a function. We put the question: influence or not? No hypothesis of the way in which these variables are related is made, so the experiment cannot give an idea of the function. The absence of a functional hypothesis has the drawback that the results cannot be interpolated or extrapolated. We have no idea how the causal variable is working and on which level the variable is significant. It is remarkable that a great part of soil-fertility research in different countries is or was based on this assumption. Application of the analysis of variance has strongly stimulated this direction. Use of two or more levels does not remove this objection.

The model in which a linear function between both variables is taken has more possibilities. A one-unit increase of the independent variable increases the effect by a constant amount, no matter what value the first variable has. We know that this assumption is not valid in many cases. The linearity can be useful in a limited region of the variable range, but according to experiences in soil-fertility research it should be more useful to utilize non-linear functions reaching a maximum and followed by a depression. There are advantages in choosing the simplest function in this case. Many equations have been proposed in the literature. The best known is the Mitscherlich equation. Other functions are the parabola and the Cobb-Douglas equation. We shall not discuss the properties of these functions. They have in common that they have been developed mainly heuristically and that their theoretical base is very small. The great difficulty therefore is to give the parameters of these functions a fundamental meaning.

Our experiences have made it clear that models with only one causal variable are not very useful. We are therefore obliged to use models with two or more independent variables. The most simple assumption again is that the yield is affected by two or more variables without any indication of of the functions. In spite of the use of more variables the practical meaning is again not large; e.g. advice about the most economical dressing cannot be given. This also applies to the model in which linear functions between variables and yield are assumed. In view of this objection non-linear functions with more variables have been developed. Figure I is an example of a model with many variables of which the function is assumed to be a *linear* one. The model explains the variation of the MgO content of grass with a number of independent variables. The model can be represented by the following equation:

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}y_2 + a_{16}y_3 \tag{1}$$

where y_2 and y_3 are crop variables.

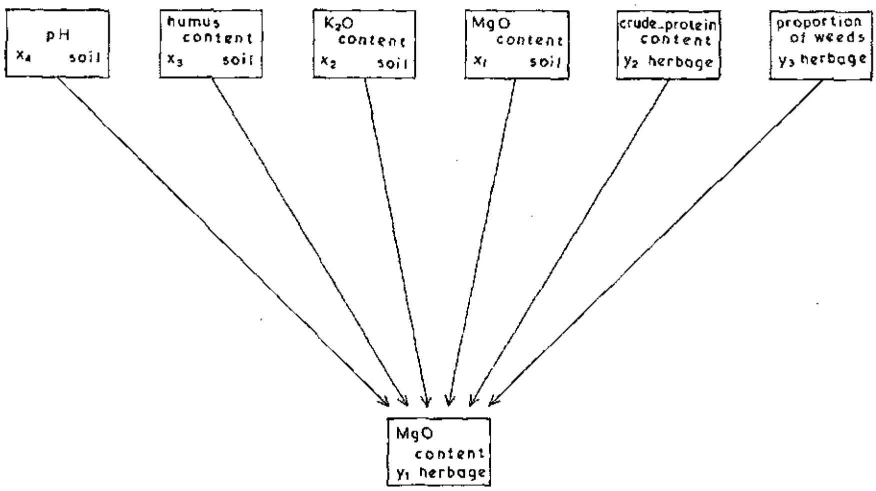

FIG. 1. Diagrammatic model of the influences of independent variables only on the MgO content of grass, represented by one equation.

Our experience is that we are confronted in such factorial experiments with so-called interactions. The size of the influence of a variable is affected by the level of the others. We speak of a nitrogen-phosphate interaction if the effect of the nitrogen is affected by the level of the phosphate dressing. The interaction is represented in the one-equation model by the product term of both variables. In most cases the researcher has no idea of the causal background of this interaction. We shall later show that the interactions can often be explained by chain processes. In such cases one equation to describe *the model is insufficient.* As a matter of fact an interaction is a measure of our lack of knowledge, and in most experiments we try to assess this lack statistically highly significant!

The use of models with many variables is stimulated strongly by the need to investigate variables that cannot or hardly be changed. It is our experience that we have to use such variables in soil-fertility research. As an example we take the explanation of differences in dressing response between the experimental fields, or of differences in yield between the fields. From this it follows that we have to test this many-variable model against data obtained partly without interference. The utmost consequence is an experiment with only single plots. In this way we succeeded in explaining 88% of the potato-yield variation in a certain area.

Something means that a many-variable model is tested by which the risk of false inferences in view of the great number of disturbing variables is great. The researcher therefore will always try to get a great variation range of many variables, whereby he can reach a better elimination of the risk. Finally, the remaining variance gives an answer to the question how the researcher succeeded in obtaining an explanation of the differences found.

This multi-dimensional concept contrasts with the method in which the variation between experimental fields is made as small as possible. In this case the experimenter is more interested in the question how significant the average dressing response has been. A high significance demands a small number of disturbing variables. The researcher tries to reach conditions controlled for as many variables as possible, the experimental fields are laid out under the same conditions of soil, climate, management, etc. This way of working cannot give many practical results because the reality is considered too simple. Sandison (1959) wrote about this: 'A lower between-trial error, leading to significance from fewer trials, is not necessary a matter of congratulation, but suggests that the trial centre or seasons may not have been sufficiently representative.' It is indeed much more realistic to ask which variables have produced the differences between the experimental fields. The answer to this question takes us towards a multi-dimensional concept of soil fertility (Ferrari 1966).

(*b*) *Models with more equations; chain processes*

The models discussed so far are characterized by the assumption that a change of an independent variable affects the dependent variable only, and does not affect the other independent variables. The same assumption is also made in experiments with artificial change according to the *ceteris-paribus* principle, in which case variables of the fourth category are excluded. However, we know that this model is in many cases not in accordance with reality, nor in experiments with interference. The model is false, so the causal inference is questionable.

The reality is often that an interference or a change of the independent variable affects not only the dependent variable, but also other independent variables. The dependent variable is affected by this interference in direct and indirect ways, and we have to do with chain processes. An example connected with the MgO research of Fig. 1 is given in Fig. 2. In this model the independent variables y_2 and y_3 of equation (1) are considered both as cause and as effect. The model of Fig. 1 could be represented by one equation. For the description of this model a system of the following equations is used:

$$y_1 = b_{12}y_2+b_{13}y_3+a_{12}x_2+a_{13}x_3 \qquad (2)$$

$$y_2 = b_{23}y_3+a_{21}x_1+a_{22}x_2+a_{23}x_3 \qquad (3)$$

$$y_3 = a_{31}x_1+a_{32}x_2+a_{33}x_3+a_{34}x_4 \qquad (4)$$

The method of path coefficients comparable with the method of simultaneous equations in econometrics offers the possibility to test the correctness of such models with chain processes and feedback systems and to quantify the parameters of the equations. The necessary data can be obtained by interference or without interference. A comparison between models with one or more equations shows that they must give different inferences. It is clear again that models with a system of equations are often necessary in soil-fertility research. The advantage of the use of models in an experiment with interference is that it shows which determinations have to be carried out.

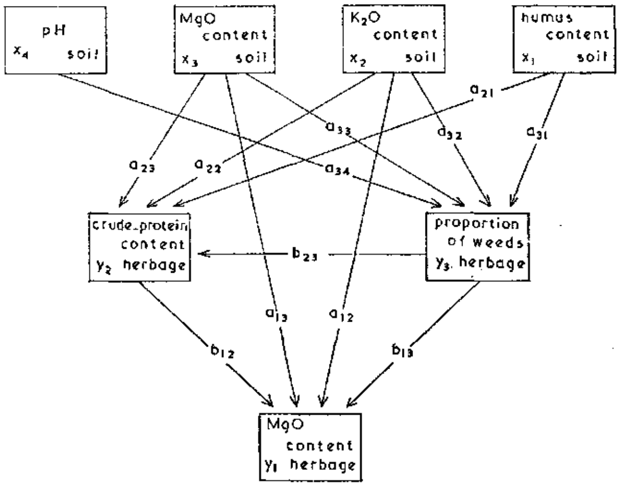

FIG. 2. Diagrammatic model of the influences of variables on the MgO content of grass, represented by a system of three equations.

A model more or less in between that with one equation and that with more equations is the one on which aspect analysis is based. The characteristic of the model of this analysis is that relationships between many variables are assumed. Her no distinction between independent and dependent variables is made at the beginning. Further, we are concerned more with correlations than with regression coefficients. Chain processes are lacking in the model. The relations between the independent variables and their effects take place via hypothetic causal aspects that cannot be determined. The model of an aspect analysis of the MgO investigation already mentioned is given in Fig. 3. The difficulty of the analysis is that such models are not identifiable, and that many solutions are possible. The *caumax* rotation—introduced by us provides the possibility to obtain a causal identification (Ferrari and Mol 1966). The method of aspect analysis has its limits, as appears from a comparison between the models in Figs. 2 and 3, but it is very useful to test and to quantify a less specified hypothesis with a large number of variables. Further, it offers the possibility to give information about variables and number of equations to be taken in more specified models.

This idea of identification forms in research an important element mostly not recognized as such, and needs some explanation. Here the under-identification is especially important, indicating that a certain relationship in the model cannot be resolved; the parameters are indeterminable. We

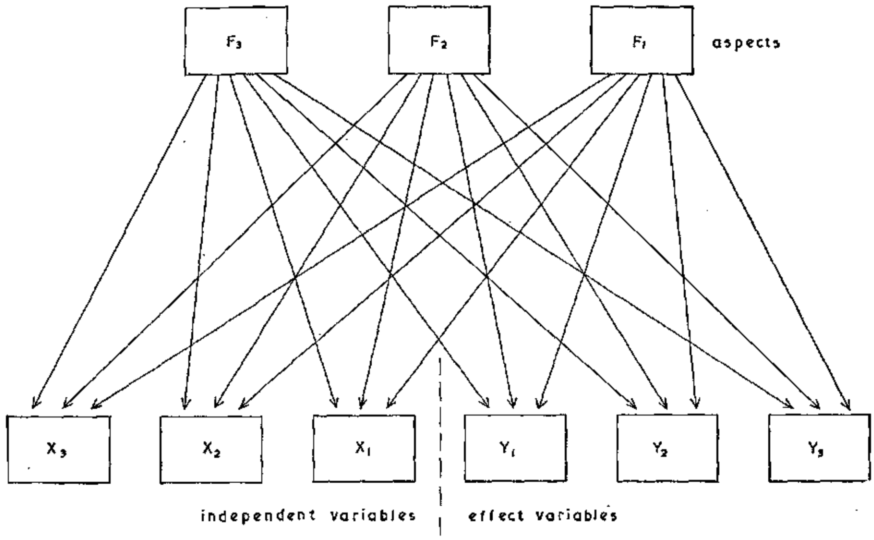

FIG. 3. Diagrammatic model of the aspect analysis with some variables of Figs. 1 and 2.

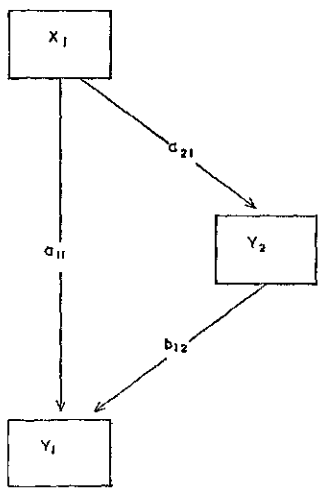

FIG. 4. Model with an unidentified equation.

take as example the model of Fig. 4, in which the influence of x_1 on y_1 is assumed to be direct and indirect via the variable y_2. We suppose that this model must be tested and quantified. For this purpose we change the variable x_1 into more values and determine the obtained changes in y_1 and y_2. In this case, however, it is impossible to separate the direct and indirect influence of changes of x_1 on y_1, as the equation of y_1 as a function of x_1 and y_2 is not identified. This under-identification is related with the ratio of the difference in the number of independent variables in model and equation to the number of dependent variables in the same equation. A solution cannot be reached by replicating the determinations. It can only be obtained by the introduction of a second independent variable in the model. An investigation of such models cannot be designed and performed well if the researcher has not realized the identification conditions. As a matter of fact, we meet this under-identification in many manipulative experiments in soil-fertility research.

DISCUSSION

Every researcher tries to make an ideal experiment, viz. a manipulation of the independent causal variable in an isolated system. Most researchers, however, do not realize that the possibility for carrying out such an experiment and also the value of the conclusions are determined by the problem investigated. It is the experience of soil-fertility research that the relationships are multi-dimensional. The system can only be described by many variables and by more than one equation. This means that it is difficult to get isolated systems. This problem applies not only to an experiment with interference, but also to one without interference. In general, the rule applies that the greater the departure from the ideal experiment, the larger the number of variables and the number assumptions as to how they fit together. Most researchers on soil-fertility believe that they are making an ideal experiment. In reality they are not.

The researcher has the choice between various possibilities to obtain this isolation—on one side the pot experiment under strongly controlled conditions, on the other side the field experiment. It follows from the more-dimensional character of soil fertility that the practical importance of pot experiments must be doubted strongly. Even a basis for comparison is lacking: volume of pot, area or number of plants? The possibilities to use the results of such isolated systems in practice are very small. These objections also apply to field experiments, but not to the same extent. The possibility also exists to obtain an isolation by means of mathematical treatment.

The researcher can also make a choice between interference and no interference. In some cases it is even impossible to carry out an experiment with interference. The extreme consequence of this possibility is an experiment with only single plots without interference. Interference has great advantages with regard to the determination of the asymmetry of the relationship. It is possible, however, to prove a certain asymmetry in another way. In the choice for interference the researcher has to realize which changes of the variables to be held constant may be produced by the interference itself. It is our experience that the assumption of constancy is not real in many cases. The researcher then has the possibility to investigate these changes simultaneously and to realize the identification conditions.

The background of many disappointing results of soil-fertility research is the fact that the concept of the ideal experiment has in many cases up till now defined the subject-matter of the hypothesis. This is an incorrect idea as the conditions for isolation are mostly not fulfilled. The system of soil fertility itself must guide the choice of the number of variables taken up, the amount of manipulation employed and the possibilities for causal inferences. A model offers the possibilities to find the correct method. At the same time we get an idea which variables can disturb the causal inferences.

It is a fundamental difficulty that we still lack a method that ensures that all relevant variables have been taken in the model. Nor do we have any foolproof procedure for deciding which variables to use. The importance of aspect analysis depends on the possibility to give information about the number of equations to be taken and about the variables not to be taken. The choice must also be determined by our present knowledge and by the reasonableness of the assumptions. We shall then find that a lot of specialized research still has to be done. Further we must realize that in research we are approaching more and more to a mathematical description of soil fertility.

REFERENCES

Blalock, H. M. 1964. Causal Inferences in Non-experimental Research. Chapel Hill.

Bradfield, R. 1961. A quarter century in soil fertility research and a glimpse into the future. Soil Sci. Soc. Am. Proc. **25**, 439–442.

Ferrari, Th. J. 1963. Causal soil-plant relationships and path coefficients. Pl. Soil **19**, 81–96.

Ferrari, Th. J. 1964. Auswertung biologischer Kettenprozesse mit Hilfe von Pfadkoeffizienten. Biom. Z. **6**, 89–102.

Ferrari, Th. J. 1965. Models and their testing in agricultural research. Neth. J. agric. Sci. **13**, 366–377.

Ferrari, Th. J. 1966. Towards a soil fertility in dimensions. Neth. J. agric. Sci. **14**, 225–238.

Ferrari, Th. J. and Mol, J. 1966. Aspect-analyse van causale modellen. Rapport 9, Inst. Soil Fertility, Groningen.

Sandison, H. 1959. Influence of site and season on agricultural variety trials. Nature **184**, 834.

Wold, H. 1956. Causal inference from observational data. J. roy. stat. Soc. Ser. A, **119**, 28–60.