Merlewood Research and Development Paper

Number 7

Multivariate Statistical Analysis in the Search for Basic Factors in Forest Site Classification

J. N. R. Jeffers, F.L.S.

R. & D.69/7

Merlewood Research and Development Papers are produced for the dissemination of information within the Nature Conservancy. They should be regarded as internal documents only, and must not be quoted or referred to in formal publications. All opinions expressed in Merlewood Research and Development Papers are those of the authors and must not be taken as the official opinion of the Nature Conservancy.

1. Introduction

The devel•pment of the electr•nic digital computer, and the increased availability of such computers to research workers, have stimulated the interest in mathematical models in many fields of applied research. Where, in the past, scientists have usually sought physical and mechanical analogues as the basis of their conceptual models, they are now willing to abstract the essential elements from the practical problem and to express these elements in the symbolic forms of mathematics. Even in fields of research which are traditionally non-mathematical in orientation, e.g. plant and animal taxonomy, ecology, psychology, sociology, etc., the use of mathematical models as the basic conceptual models has become almost commonplace.

Despite this development of the use of mathematical models, relatively little thought has so far been given to the optimum ch•ice of the strategy to be used in a particular investigation.

Most attention has so far been given, in practical applications of mathematical models, to direct simulation of biological, physical, and chemical processes. Such direct models are built up from consideration of the rates of change taking place in the component parts of the process, and are frequently represented by sets of differential equations or transfer functions. The main aim of the model is to describe the process as accurately as possible, although it is expected that the description will lead to reasonably accurate predictions - essential if the model is to be tested against reality. Mathematical models of this kind may also be used in the making of decisions, either through the prediction of the outcome of various alternative courses of action, or, directly, from the choice of pre-determined parameters.

The direct simulation of processes in this way virtually requires a new model for every application, and the modelling process may itself, therefore, be laborious and time-consuming. There is relatively little "spin-off" from one model to another, except in the experience of the individual research worker, who is frequently able to avoid unproductive approaches to the modelling of a new process. A further disadvantage of direct simulation is that the approach does not necessarily lead to parsimony in the description of the process. The complexity of the process may be such that the modeller is tempted to include large numbers of parameters, all of which have to be estimated. These parameters may reduce the value of the model in practical work, because of the difficulty of obtaining estimates in practical situations, and hence may also reduce the range of applicability of the model.

An alternative approach is that of choosing mathematical models which have desirable properties, even if these models cannot be regarded as direct simulations of the process. One advantage of this approach is that it may be possible to choose models which can be applied to a wide range of problems, so that, even if the efficiency of the solution t• •any one problem is not high, the overall efficiency is nevertheless considerably enhanced. Other advantages are that the mathematics involved in the modelling may be simplified, parsimony of the parameters involved in the description may be achieved, emphasis can be placed on the parameters which are abstracted, and probabilistic elements may be readily incorporated in the model. The approach may not, however, appeal to those research workers who like direct analogues and who are prepared to express physical and chemical ideas in mathematical terms, but are not prepared to accept a conceptual model which is entirely a mathematical artefact. Furthermore, essential elements may be missed more easily where the direct analogue is not available, and the mathematical formulation may introduce assumptions which themselves limit the application of the model.

Clearly, both approaches to the problem of finding mathematical models for practical problems are of value, and the choice between the two approaches will depend partly on the knowledge and training of the investigator, and partly on the computing equipment and mathematical advice which is available. As Williams and Lance, 1968, have pointed out, it may be desirable, for important fields of research, to use both approaches, as it will seldom be possible to tell in advance which of the two methods is the most likely to produce the best working approximation to the real process.

As illustrations of the second approach, this paper considers, first, some simple examples of useful mathematical models, and then gives a brief description of the application of a particular type of multivariate model to a complex problem of forest site classification.

2. Models with Convenient Properties

One of the most frequently used models, although it is not always recognised as such, is that invoked by the analysis of variance of observations taken in a designed experiment. In the formal analysis of this model, an equation is set up for every observation, expressing the observations as the sum of several components. For example, in a randomised block experiment, the equations have the general form

$$y_{ijk} = \mu + \tau_i + \beta_j + e_{ijk}$$

where μ is a general average about which the observations fluctuate,

τ is the effect of the ith treatment applied,

β is the effect of the "block" - an environmental effect,

e is a measure of the experimental error.

Subject to some formal constraints, such that the sum of the treatment effects is zero, and that the sum of the block effects is zero, it is possible to solve a set of such equations so as to be able to test hypotheses about the treatment effects, or to estimate the confidence limits which can be placed on these effects. Three major assumptions have to be made, however. First, it is assumed that the treatment and block effects are additive. Second, residual effects are assumed to be independent from observation to observation, and to be distributed with zero mean and the same variance σ^2. Third, for valid tests of significance and the estimation of confidence limits, the residuals are assumed to follow the normal distribution, (Cochran and Cox, 1950). While these assumptions may not be exactly true, there are many situations where the model is a useful approximation, and the analysis of variance is, therefore, widely applied. Two different versions of this model are actually used in practice (Federer, 1956), although very few research workers distinguish between them:-

Model 1 in which it is assumed that the treatment effects are fixed, and that the e_{ij} are random variates distributed around zero, with a common variance.

Model 2 in which it is assumed that the treatment effects, r_i, and the e_{ij} are random independent variates distributed around zero. The variance of the e_i is estimated as s^2 and the variance of the e_{ij} as s^2_e.

A second example of a mathematical model with desirable properties is that used in multiple regression analysis, in which the values of one variable, the dependent variable, are estimated from the values of two or more variables which are assumed to be independent. In practice, it is seldom the case that the independent variables are uncorrelated, and, as it is confusing to talk about correlated, independent variables, it is preferable to adopt the terminology of Kendall and Stuart, 1961, and to refer to these variables as the regressor variables. Thus, in multiple regression analysis, a predicting equation of the form -

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots\dots\dots\dots + b_p x_p + e,$$

where the b_0, b_1, b_2, ... b_p are coefficients estimated from the data, is derived from sets of observations of the p regressor variables and the dependent variable. In this model, certain assumptions are made. First, it is assumed that the individuals are taken from a population in which the variance is homogeneous, or, in other words, that the variance of the values of the dependent variable about the regression surface is the same for all combinations of the regressor variables. If the variance is not homogeneous, it will usually be necessary to use some weighting procedure. The second assumption is that the deviations of the dependent variable from the regression surface are independent of each other, and that the size and direction of the error for one individual has no relationship to the size and direction of the error for any other individual of the sample, except that all the individuals are from the same population. A final requirement is that the values of the regressor variables are to be measured with, essentially, no error. It should be noted that the fitting of the regression equation by the method of least squares does not require that the values of the dependent variables are normally distributed about the regression surface, but that the commonly used procedures for computing confidence limits and making tests of significance (t and F tests) do formally assume normality.

The multiple regression model is widely used, often with little or no regard to the basic assumptions. When, however, it is used in situations in which the assumptions are at least approximately true, the technique is valuable as a method of predicting the value of the dependent variable which corresponds to given values of the regressor variables.

It is also widely used as a method of exploring the relationships between a dependent variable and several other variables, few, if any, of which will have been measured with any less degree of error than the dependent variable, and which usually display a high degree of intercorrelation. The selection of the regressor variables to be included may be guided by previously nominated hypotheses, and the selection may include linear, non-linear, and interaction terms. More frequently, however, the computations are approached without any nominated hypothesis about the form of the relationships, and it is confidently expected that the regression analysis will define the "best" relationship.

The multiple regression model is multivariable rather than multivariate, in that there is only one variate, the dependent variable. In the multivariate model, several variables are measured on each individual included in the sample. Within the total set of variables, no one variable can be regarded as dependent on the other variables, and, in general, none of the variables will have been measured at predetermined levels. In this situation, we may be more concerned to discover the essential "dimensionality" of the set of variables, and to identify the several dimensions which have been isolated. To do this, it is customary to carry out a principal component analysis of the matrix of coefficients of linear correlation between every pair of variables. Descriptions of the analysis are given by Kendall, 1957, and Seal, 1964.

In this analysis, the original data are expressed as a series of linear transformations such that each transformation is orthogonal to all the other transformations, and that each transformation accounts for the maximum amount of the residual variance. Such a model requires a minimum number of assumptions. However, it is expected that -

(a) all but a small part of the total variability contained by the set of variables can be expressed in some smaller number of linear combinations or components; and

(b) these components can be interpreted as meaningful expressions of biological, chemical, or physical entities.

No assumptions are made, in this analysis, about the normality of the distribution of the population from which the observations are drawn, but homogeneity of the variances and covariances is assumed. It is also assumed that the basic variables can be meaningfully included in linear combinations, and if the assessed variables refer to effects which are multiplicative rather than additive, it will usually be necessary to transform the original data to their logarithms. Experience in the use of multivariate models of this kind suggests that these assumptions are frequently justified. Apart from those assumptions, however, the model does have the very desirable property of expressing the essential variability of the process in the smallest possible number of dimensions, and hence achieves a marked degree of parsimony of description.

In more complex multivariate situations, there may be several sets of variables, such that one or more of the sets may be regarded as dependent on one or more of the other sets of variables. In this situation, a useful strategy can be derived from the application of the principal component analysis to each set of variables, followed by the correlation of the components that have been extracted. By definition, the components from any one set will be independent, but the correlations between the components in the separate sets will indicate the dimensionality of the total problem, and the degree of intercorrelation between the basic dimensions.

3. A Problem of Forest Site Classification

As an illustration of the kind of problem which can be explored by means of the complex multivariate model described above, we may take the classification of sites for the growth of commercial timber. For example, Corsican pine has been widely planted over the last fifty years on sites intended to produce large volumes of timber for pitwood, fencing, fibreboard, chipboard, pulpwood, and selected material for joinery. The Forestry Commission alone has about 50,000 acres of this species, planted during the lasttwenty years. The sites on which this species has been grown are usually at low elevations, particularly on sandy areas near the sea, on light sandy soils and also on heavy clays in south and east England, but relatively little is known about the relationships between the rate of growth of the species, as a crop, and the variables of climate, physiography, soil physics, soil chemistry, nutrient uptake, etc., at individual sites. The problem is complicated by the very large numbers of variables which may possibly have an effect on the growth of the trees.

A strategy which might be adopted for an investigation of this kind is that of a complex multivariate model, constructed to meet the following objectives:-

(a) understanding of the relationships between the various groups of variables that can be assessed on individual sites;

(b) elimination of variables which contribute little or no extra information from future investigations;

(c) selection of the useful variables for prediction of growth rates;

(d) classification of forest sites in relation to their characteristic properties.

A small project, which was carried out by the staff of the Soils Section of the Forestry Commission Research Division, in collaboration with Dr. D. G. M. Donald of Stellenbosch University, South Africa, provided an opportunity to test such a strategy, and data were collected from thirty sites, representing a wide range of the localities and site types on which Corsican pine has been planted. Seventeen of the sites were permanent sample plots, nine were experimental plots, and only four were special plots created in the forest for the purposes of the study. The fifty site and crop variables determined at each site were considered to fall into four groups:-

(i)	climate and physiography variables	(15)
(ii)	soil physics variables	(13)
(iii)	soil and foliage chemical variables	(16)
(iv)	tree growth variables	(6)

There is some degree of logical inconsistency in the inclusion of the chemical variables for the soil and the foliage in the same group, in that the foliage variables represent, at least in part, a response to the soil chemical variables, but the two sets of variables were included in the same group with the idea of expressing soil fertility in both total and available quantities.

A multivariate analysis was carried out on the data from this investigation, including principal component analyses of the four groups of variables, orthogonalised regression of the tree growth variables on the other three groups, and cluster analyses of the sites within the multivariate space described by the principal components of the various groups. A detailed account of the results of these analyses will be published separately. In this paper, only sufficient information about the results will be given to indicate the succe ses of the strategy in achieving the objectives defined.

4. Summary of Results

The components extracted from the analysis of the data are summarised in Table 1. The entries under the heading "interpretation" are merely convenient names for the components, and their exact interpretation should be verified by interested scientists from the more detailed report.

(a) Climate and Physiography

The first six components derived from the analysis of the variables of this group accounted for 97.2 per cent of the total variability described by the fifteen variables of climate and physiography, and, of these, only the first four, accounting for 90.4 per cent of the total variability, are of practical significance.

The first component, accounting for 50.0 per cent of the variability described by the variables, has been interpreted as a component of excess energy, giving greatest weight to variables of evapo-transpiration and growing season temperatures. The second component, accounting for a further 28.9 per cent of the variability is a component of mean winter monthly temperature. The third component, accounting for a further 6.1 per cent, is almost entirely an expression of the aspect of the site, and the fourth component, accounting for a further 5.4 per cent of the variability, is a measure of the Penman coefficient for the month of June.

The variables of climate and physiography, therefore, are dominated by two major components, together accounting for 78.9 per cent of the variability, and which describe the excess energy and winter warmth of the site. Aspect and the Penman coefficient for June describe a further 11,5 per cent of the variability.

(b) Soil Physics

The first six components of the variables of this group accounted for 91.8 per cent of the variability described by the soil physics variables. Of these, the first five components, accounting for 86.6 per cent of the variability, are of practical significance.

The first component, accounting for 31.0 per cent of the variability, has been interpreted as available water capacity. The second component, accounting for a further 23.1 per cent of the variability is a measure of the stoniness and clayiness of the soil, i.e. a stone/clay factor. The third component, accounting for 12.8 per cent, is a mull/mor humus factor. The fourth and fifth components, accounting for 11.3 and 8.5 per cent, are measures of the accumulation of organic matter and the thickness of the surface humus layers, respectively.

(c) Soil and Foliage Chemistry

The first six components derived from the variables of this group accounted for 82.7 per cent of the total variability described by the original sixteen variables, and all of these components are of practical significance.

The first component, accounting for 29.7 per cent of the variability, is an index of the phosphate status of the soil. The second component, accounting for a further 19.0 per cent of the variability, is a measure of the nutrient status of the sample trees, i.e. a measure of nutrient uptake. The third component, accounting for 15.3 per cent, is a measure of the base status of the soil. The fourth component, accounting for 7.4 per cent, has been interpreted as "interference", a contrast between the levels of phosphate and potash in the foliage and the percentage of phosphate in the soil as a derived factor. The fifth and sixth components, accounting for 6.2 and 5.1 per cent, are measures of total phosphate and foliage calcium, respectively.

(d) Tree Growth

Only two of the tree growth components are of any practical significance, together accounting for 90.4 per cent of the total variability contained by the six original variables. The first component, accounting for 64.2 per cent, is a general growth index. The second component, accounting for a further 26.2 per cent, is an index of recent height growth.

The principal component analyses of the four groups of variables were successful in providing a description of the independent sources of variability of the data collected in this study. The parsimony of the description obtained is perhaps most striking for the climate and physiography variables and for the tree growth variables, where the numbers of independent components are only one third of the numbers of the original variables. Nevertheless, the relatively straight-forward interpretations of the components are encouraging, not least because they confirm some of the relationships which were expected.

As an extension of the principal component analysis, the values of each component can be calculated for each site, and the correlations between these transformed values calculated. If it can be assumed that the linear transformations of the original variables implied by the components will result in approximately normal distributions, the significance of the correlation coefficients can be tested in the usual way to discover the inter-relationships between the components of the four groups. For this example, these relationships are summarised in Figure 1. In this figure, the lengths of the arrows are of no significance, but the presence of an arrow indicates a correlation of the sign shown beside the arrow. The components are shown in relation to the general growth index and the index of recent height growth, in an attempt to illustrate the response of Corsican pine to the environment as well as the inter-relationships between the factors of the environment.

There are a number of interesting features of Figure 1. First, it is slightly surprising that the pedology of the site assumes so important a role in the relationships. Of the six components directly correlated with the general growth index, four of the components are concerned with the physical and chemical properties of the soil. Earlier research workers have usually tended to minimise the relationships between tree growth and soil properties, possibly because of the complexity of the relationships between the soil factors themselves. The diagram shows something of this complexity, although the explanation which it gives of the major groups of factors seems reasonable. The lack of any direct correlation between the growth index and water capacity is of rather special interest, and, for other species, e.g. Sitka spruce, quite different relationships might be expected. A detailed interpretation of Figure 1, and of the analyses which can be derived from this figure, will be given in the more detailed report, but the intention here is to show the value of the mathematical model in leading to the understanding of an extremely complex biological, physical and chemical problem.

It is perhaps appropriate to mention at this point that there is an alternative model, that of canonical correlations, which is specifically intended for the situation in which one group of variables is assumed to be dependent on another group of variables, and this model is also described by Kendall, 1957. It would, therefore, have been possible to have correlated the variables of tree growth directly with those of climate and physiography, soil physics and soil and foliage chemistry. Experience with this technique, however, suggests that the results are seldom readily interpreted (Jeffers, 1967).

The second objective of the study of the Corsican pine sites was the elimination of variables which contribute little or no extra information from future investigations. The parsimony of the descriptions achieved in this example certainly suggests that some at least of the original variables could be dropped from future investigations. In the climate and physiography variables, for example, more than 90 per cent of the total variability contained by the fifteen original variables could be described by only four linear combinations, so that eleven of the original variables are nearly expressible in terms of the other four. There is little point, in future investigations, of measuring so many closely correlated variables, and examination of the weights given to the original variables by the principal component analysis provides suggestions for the dropping of variables in future studies. Similar selections can be made from the other groups of variables, and future investigations of the relationships between tree growth and environmental factors can be made on only twenty variables or less, unless the opportunity is taken to introduce variables not so far tested in the current investigations. The complex multivariate model used in this investigation, therefore, leads to the formulation of a conceptual model which is dynamic, in the sense that sets of variables can be screened to assess the total number of dimensions that they represent, and new variables can be introduced in later investigations to assess whether or not these variables increase the number of essential dimensions that are described.

Where the objective of the investigation is to predict the growth of Corsican pine, rather than to seek an understanding of the relationships between the groups of variables, some interesting extensions of the model are available. Inspection of Figure 1 suggests that any efficient predictor of general growth of Corsican pine will require variables of winter warmth, depth of raw humus, mull/mor humus, stone/clay factor, soil phosphate status and sample tree nutrient status, but the intercorrelations between these groups indicate that a large number of possible regression equations could be written which would have approximately similar predicting power.

Of the six variables of tree growth, two may be regarded as of particular importance, i.e. the general yield class and the local yield class. These two variables are widely used as indices of the productive capacity of sites, and as a means of entering the Forestry Commission Management Tables, which enable predictions to be made of the likely future growth of the stands and their intermediate yields. It is possible that the general growth index, derived from the component analysis of the tree growth variables would provide a better measure of tree growth, if only because this index includes the density of the stand and the current top height, but the special importance of yield class in forest management makes it desirable that the relationship of the general and local yield class to the environmental factors should be given special consideration.

In order to clarify relationships between individual dependent variables and the variables of groups of regressor variables, the regressions of the dependent variables can be calculated on each of the components extracted from the other groups of variables, using a technique of orthogonalised regression suggested by Kendall, 1957. Table 2 shows the components of the three groups of regressor variables which make a significant contribution to the variability of the general and local yield class. These orthogonal components of the regression of general and local yield class give a fair indication of the relative importance of the various components in predicting yield class, and also of the likely candidates among the original variables as predictors of yield class. Because the components are orthogonal within any one group of variables, it is possible to estimate the total contribution of any one group, and the estimated contributions of climate and physiography, soil physics, and soil and foliage chemistry are 18.81 per cent, 59.28 per cent and 71.98 per cent respectively. It is, of course, not possible to add together the contributions of the components of the three groups, as these are not independent. The diagram of Figure 1 makes this clear, but some combination of the variables making a significant contribution to the variability of general and local yield class will certainly provide efficient predicting equations, although in general it will not be possible to nominate any one equation as "best" in any realistic sense. Nevertheless, it is felt that this analysis is revealing in that it expresses, in quantitative terms, the predictive ability of the basic components in respect of yield class.

The final objective set for the complex multivariate model used in this example was the classification of forest sites in relation to their characteristic properties. This objective may be achieved by a simple extension of the principal component analysis, in which cluster analyses are carried out on the computed values of the components calculated for each site. As an example, a slightly modified version of the minimum spanning tree method described by Gower and Ross (1969) has been applied to the values of the four components of the variables of climate and physiography, and the resulting classification is given in Figure 2. For further ease of interpretation, the classification can be mapped against the projection of any two components.

The interpretation of the classification is aided by the fact that, by definition, the components are independent, so that, in mathematical terms, the classification has been constructed in Euclidean space. The classification of the sites by climate and physiography divides the sites into three broad groups, each of which is further sub-divided into sub-groups of two or more sites, except in the last group where one site is an outlier to the group.

Conclusions

The example of this paper gives a fair demonstration of the value of what may seem, on first sight, to be a relatively unlikely mathematical model. The model which underlies principal component analysis, whether applied to simple or complex problems, has little theoretical justification. The model does, however, have many useful properties, and experience in its application to a wide range of practical problems suggests that, provided that the investigator remembers that he is employing the model as a mathematical tool, valuable results can be obtained from its use. The ability to handle large numbers of variables, in complicated arrangements of sets of regressor and dependent variables opens up research strategies which have not been available in past research, particularly in the search for understanding of complex relationships, the elimination of variables which contribute little or nothing from future investigations, the selection of useful variables for prediction, and the classification of individuals in relation to characteristic properties.

References

Bradley, R.T., Christie, J.M., and Johnston, D.R., 1966. Forest Management Tables, Forestry Commission Booklet No. 16. H.M.S.O.

Cochran, W.G., and Cox, G.M., 1950. Experimental Design. John Wiley, Ney York, 1950.

Federer, W.T., 1956. Experimental Design. Macmillan, New York, 1956.

Gower, J.C., and Ross, G.J.S., 1968. Minimum spanning tree and single linkage cluster analysis. In press.

Jeffers, J.N.R., 1967. The study of variation in taxonomic research. The Statistician, 17 (1), 1967, pp.29-44.

Kendall, M.G., 1957. A course in multivariate analysis. Griffin, London, 1957.

Kendall, M.G., and Stuart, A., 1961. The advanced theory of statistics. (Vol. 2) Inference and relationship. Griffin, London, 1961.

Levins, R., 1966. A strategy of model building in population ecology. American Scientist, 54 (4), 1966, pp.421-481.

Seal, H., 1964. Multivariate statistical analysis for biologists. Methuen and Co. Ltd., London, 1964.

Williams, W.T., and Lance, G.N., 1968. The choice of strategy in the analysis of complex data. The Statistician, 18 (1), 1968. pp.31-43.

Table 1. Principal Components of the Four Groups of Variables

Component	Percent	Interpretation
Climate and Physiography		
1	50.0	Excess energy
2	28.9	Winter warmth
3	6.1	Aspect
4	5.4	Penman coefficient for June
	90.4	
Soil Physics		
5	31.0	Available water capacity
6	23.1	Stone/clay factor
7	12.8	Mull/mor humus
8	11.3	Accumulation of organic matter
9	8.5	Thickness of surface humus layers
	86.7	
Soil and Foliage Chemistry		
10	29.7	Soil phosphate status
11	19.0	Sample tree nutrient status
12	15.3	Base status of soil
13	7.4	Interference
14	6.2	Total phosphate
15	5.1	Foliage calcium
	82.7	
Tree Growth		
16	64.2	General growth index
17	26.2	Recent height growth
	90.4	

Table 2. Significant Regression of General and Local Yield Class on Components

Variable group	Component	Percentage of Variability General Yield Class	 Local Yield Class	'Best' Variable
Climate and physiography	2	18.81	19.38	Mean winter monthly minimum temperature
Soil Physics	6	21.92	19.32	Percentage clay in metre sample.
	7	37.36	42.52	Humus type
Soil and foliage chemistry	10	42.41	44.27	Percent P in surface sample
	11	20.71	18.18	Mean needle weight of foliage samples.
	12	8.86	5.70	pH of surface sample

Figure 1.

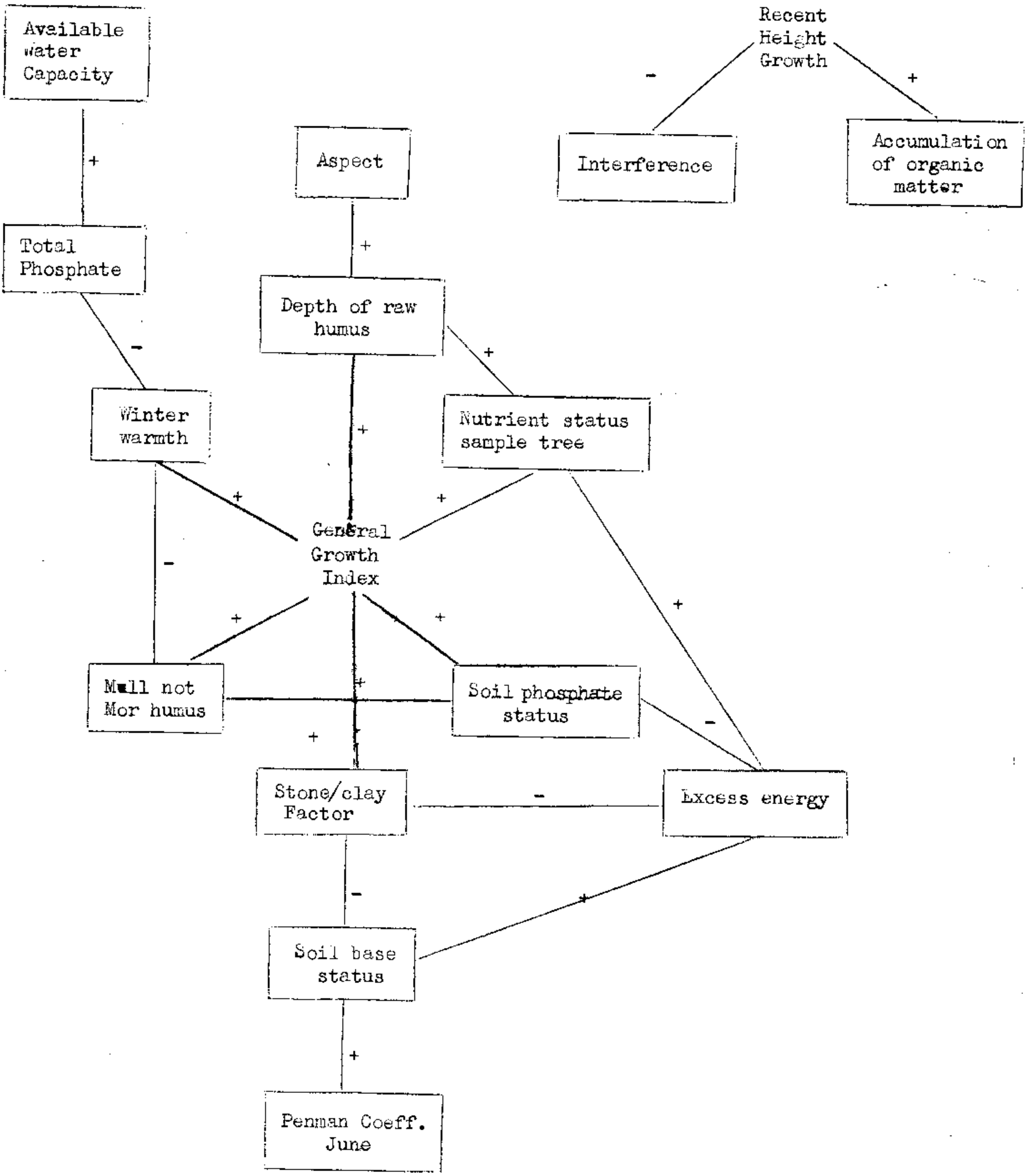
Available water Capacity
+
Total Phosphate
-
Winter warmth
Aspect
+
Depth of raw humus
+
+
Nutrient status sample tree
Recent Height Growth
-
+
Interference
Accumulation of organic matter
+
+
General Growth Index
-
+
+
+
Mull not Mor humus
Soil phosphate status
+
-
+
Stone/clay Factor
-
Excess energy
-
+
Soil base status
+
Penman Coeff. June

Figure 2. Classification based on Climate and Physiography

Symbol			
	Tunstall		236
O	Rendlesham		150
	Rendlesham		371
O	Swaffham		137
	Olleys		181
x	Brendon Top		427
	Brendon Bottom		427
v	Dean Highmeadow		186
	Dean North		187
n	Sherwood VI		601
	Sherwood VI		616
	Alice Holt		430
+	Bramshill		C73
	Bramshill		C68
	Bedgebury		519
	Mogshade		288
_	Mogshade		290
	Wareham	25	IVd
	Ringwood		C77
	Morden Drax		406
	Wareham	25	IVa
	Wareham	25	IIIb
	Wareham	25	IIIc
*	Wareham	26	Control
	Wareham	26	1½
	Wareham	26	3
	Wareham		116
	Wareham		129
	Ringwood		C70b
	Penbrey		153