Numerical method to evaluate land suitability for agriculture using a parametric model

by

Chang-Huei Tsai

A thesis
presented to school of graduate study
in partial fulfillment of the
requirements for the degree of
Master of Arts
in
Department of Geography

OTTAWA, Ontario, 1983

ABSTRACT

A parametric model is one of the most widely used methods in land use evaluation. However, there are two key drawbacks in parametric methods. One is that they fail to account for interactions among factors. The other is the subjectiveness of evaluating values of physical factors. The former has been solved by Lok (1979, Lok & Phipps, 1981) when the LASTAN model --- LAnd SuiTability ANalysis model --- was established. The latter still appears even in this LASTAN model.

The LASTAN model combines two elements: site conditions (Site vector) and land use requirements (Use matrix). Through a series of simple applications of matrix algebra and matrix transformations, the performance indices of site suitability for a specific land use type can be calculated. It incorporates not only the main effects of soil factors but also two-way interactions between coexisting states into performance indices. However, the Use matrix is generated on the basis of evaluations of pedologists and agronomists. It lends itself to the easy criticism of being lacking objectivity.

A computer program COUSMAT (COmputed USe MATrix) is used to calculate the Use matrix defined in the LASTAN model. The COUSMAT program is mathematical and, therefore, it is less subject to inclination. This program takes the operations of the LASTAN model into consideration, and goes through a series of converse operations of the LASTAN model to generate a computed Use matrix. The test of soil conditions in the Peace River area for growing barley and wheat demonstrates that the COUSMAT program is a mathematical program which meets the purposes of developing a mathematical method to generate the Use matrix in order to avoid the criticism of subjectiveness to the LASTAN model.

RESUME

Les modèles paramétriques sont parmi les plus largement utilisés pour l'évaluation de l'utilisation du sol. Cependant, il existe deux inconvenients à ces méthodes. Le premiér est qu'ils ne considèrent pas l'interaction entre les facteurs. L'autre est la subjectivité de l'evaluation en ce qui concerne la valeur de ces facteurs physiques. Le premier fut résolu par Loh (1979, Loh et Phipps, 1981) lorsque le modèle LASTAN --- LAnd SuiTability ANalysis model --- fut établi. Le deuxième apparait toujours, même dans le modèle LASTAN.

Ce modèle combine deux éléments: les conditions du site (vecteur site) et les exigences de l'utilisation du sol (matrice d'utilisation du sol). Par une suite de simples applications d'algèbre matriciel et de transformation de matrices, les indices de la possibilité du site pour le type spécifique d'utilisation du sol peuvent etre calculés. Ceci tient compte non seulement les effets principaux du sol mais aussi des interactions binaires entre les états en coexistence. Cependant, la matrice d'utilisation est produite sur la base d'evaluations faites par des pédologues

et des agronomes. On peut toujours la suspecter de non-objectivité.

Un programme d'ordinateur COUSMAT (COmputed USe MATrix) a été réalisé pour calculer une matrice d'utilisation telle qu'elle a été définie dans le modèle LASTAM. Le programme COUSMAT est un programme mathématique, donc, il évite une trop grande subjectivité. Ce modèle utilise le même fonctionnement que celui du modèle LASTAM mais d'une manière inversée afin de générer une matrice d'utilisation estimée. Une application empirique aux conditions de sol dans la région de la rivière de la Paix pour la culture de l'orge et du blé nous démontre que le programme COUSMAT permet d'atteindre les objectifs de ce travail: le développement d'une méthode mathématique produisant une matrice d'utilisation afin d'éviter la critique de subjectivité opposée aux modèles paramétriques.

ACKNOWLEDGEMENTS

There are many people, to whom I would like to express my appreciation for their contribution to this study.

The warmest thanks must be given to my supervisor, Dr. Michel Phipps (Associate professor, Department of Geography, University of Ottawa) for his cheerful assistance in all aspects of this study.

I am also indebted to Dr. Chang Wang (Land Resource Research Institute) for his advice and encouragement during my study period at the University of Ottawa.

Appreciation is extended to Dr. Julian Dumanski (Head, Land Use and Evaluation Section, LRRI) for offering significant suggestions to this study.

In particular I am grateful to Dr. K. B. MacDonald (Land Resource Research Institute), Mr. J. Dobb (British Columbia Ministry of Agriculture) and Dr. R. B. Stewart

(Crop Production Division, Agriculture Canada) for their kindness in allowing me to use their valuable data.

It is also a pleasure to acknowledge the assistance of my dear friends, Lesley Hoermann and Jose A. Borello for their kind assistance in correcting my English writing.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

INTRODUCTION

1.1 OVERVIEW

Land, mother of all creatures, is a limited resource. It is expected to satisfy an ever-increasing range and volume demand for goods, services and amenities. The increasing population needs more and more land for housing, education, transportation, business and recreation. Rapid business expansion and urbanization in the last two decades have competed with agriculture for the use of land. It will continue to be pressured by competing uses. More and more farmland has been converted to non-agricultural use. In the United States, between 1967 and 1975, approximately 9.1 million hectares were converted from cropland to urban and water uses. One third of these, about 3.2 million hectares, were prime farmland (Schmude 1977). In Canada, each year, Canadian urban areas cumulatively consume a land area equal in size to Hamiltom, Ontario. Not only is the quantity of land consumed of serious concern, but more importantly the irreplaceable quality of that land. Of those lands converted 63 percent were of high agricultural capability (Neimanis, 1980). In many other countries, the situation of land use being converted from prime farmland to non-agricultural use is the same as, or more serious than that of the United States or Canada. This change indicates the great importance of land use planning directed to conservation, proper allocation and use of the limited resource more wisely. Knowledge of the physical characteristics of land is fundamental to rational land use planning (Dumanski, et. al. 1978). In this view, soil suitability evaluation plays a significant role in land use planning.

Soil suitability, as defined by Vink (1975), is the physical capability of soil and climate to produce a specific crop or group of crops. According to this definition, the valuation of socio-economic factors is not explicitly considered within the area examined. The same definition holds for defined uses other than agricultural uses. Agricultural scientists working on soil evaluation around the world try to use many different methods to assess or calculate the performance indices of soil on crop growth. Each of such methods, such as capability classification (this allows the soil survey interpreter to synthesize or integrate field conditions and crop requirements, weighing the positive or negative effects of each soil characteristic in isolation and in combination with other soil factors to arrive at a capability rating) (Goodchild, 1976); or the analogue system (in which, the physical input/output data necessary for evaluation are extrapolated from experimental sites to analogous areas defined by land or soil classification) (Nix, 1968), has its own strengths as well as its weaknesses. Two other methods are: site factor method (this attempts to relate certain measured soil properties to known crop yield in a multiple regression equation, which is used to define the relationship between soil factors and productivity) (Vink, 1975), and the Parametric model (in which, attention is concentrated on the quantification of those land attributes that are considered significant to the particular land use) (Stewart 1968). The parametric model, however, might be one of the most widely used. This does not imply that the parametric model is perfect for land suitability evaluation, but in comparison with other methods its advantages outweigh its disadvantages.

1.2 THE PROBLEM

Parametric methods, defined and reviewed by Riquier (1974), consist of

1. separately evaluating the different properties of soils and giving them separate numerical valuations

according to their importance within and between each other,

2. combining these factors (numerical values) according to a mathematical law taking into consideration the relationships and the interactions between the factors to produce a final index of performance,

3. which in turn is used to rank soils in order of their (agricultural) value.

There are various approaches to combine the influence of factors on crops productivity, such as (1) additive, (2) additive and subtractive (3) multiplicative, (4) more complex equation, etc.. For example:

$$P = fX + f'Y + f''Z \quad .... (1)$$

$$P = fX + f'Y - f''Z \quad .... (2)$$

$$P = (fX) * (f'Y) * (f''Z) \quad .... (3)$$

$$P = a(1-(X-b)*(Y+c)*(Z-d)) \quad .... (4)$$

where P is the performance index;

X,Y,Z are soil factors: such as slope, texture, drainage, soil subgroup and so on;

f,f',f'' are the mathematical functions appropriate to the individual factors; and

a,b,c,d are constants. (Riquier, 1974)

No matter what approach is adopted, each factor has its own influence on the final indices. Some authors have agreed that the interactions between factors play a significant part to any approach of parametric models.

The principal advantages of parametric models are that these systems are quantitative and are particularly adapted for computer applications with large data banks. When the mathematical models are programmed, and input data of selected parameters are stored in a data base, a quantitative evaluation of the potential of a particular crop in the defined area could be immediately derived. However, the most critical element of parametric models is that the values of physical factors and interactions among those factors which have an effect on crop production are difficult to calculate mathematically. Usually these values are estimated by researchers whenever they are needed. In short, parametric methods are somewhat subjective in appraising values of soil factors and their interactions. In addition, the contributions of interactions among factors on crop productivity are frequently neglected. None of the four approaches mentioned above takes interactions between factors into consideration in a satisfactory manner.

The latter has been solved by Lok (1979, Lok & Phipps, 1981), when the LASTAN model, Land Suitability Analysis model, was established. Through a series of simple applications of matrix algebra in the LASTAN model, the performance indices of site suitability for a defined land use can be calculated, which incorporate the main effects of parameters and two-way interactions.

Although a test of the results provides a crude evaluation of the success of the model (Lok, 1979), unfortunately the elaboration of the Use matrix, which allows for the calculation of performance indices in the LASTAN model, remains subjective. The researcher must establish an ordinal rating system which defines the degree of isolated contribution of each land factor state. A set of interaction values must be established in order to evaluate the degree of interacting contribution of each land factor state with the others. Two researchers with either the same or different background and study experience would do different appraisals, and thus produce two different Use matrices. Obviously, this would result in various consequences in the performance indices. It is obvious that one of the key problems in the interpretation of soil suitability by parametric models, i.e., the effects of interactions among land factors on crops productivity, may be solved using the LASTAN model. If the other problem, the way of assessing values of coefficients of the Use matrix could be replaced by a mathematical model, the parametric method would be considerably improved.

The objective of this thesis is to attempt to establish a mathematical model to calculate those values of interactions as defined in the Use matrix in order to reduce the inadequacies of parametric methods. The goal of this study is essentially of a methodological nature. The work consists in finding an original solution to the specific methodological problem mentioned above: how to determine a Use matrix in a way which would avoid the arbitrary setting up of the coefficients, that is, in a way which ~~would~~ improve the objectivity of the parametric methods. A second goal of this work is to demonstrate how the mathematical model established may be put into operation, using a set of actual data.

1.3 METHODOLOGY

The LASTAN model combines two key elements: Site vector, representing soil or climate conditions of a particular area, and a Use matrix, representing land use requirements. Through a series of operations of the LASTAN model, the performance index of site suitability for a defined land use can be calculated. The Use matrix plays an important role in the LASTAN model. However, due to the fact that it is generated by evaluation, it causes this model to be criticized as being 'non-objective'. This study attempts to find a mathematical method to generate the Use matrix on the basis of empirical data, e.g. crop yield, in order to calculate site suitability for a specific land use by adopting the LASTAN model.

1.3.1 Developing a mathematical model

The LASTAN model is a mathematical model which could be represented by mathematical equation as:

$$R = f(S, U)$$

The Site vector and the Use matrix are independent variables and the result of the LASTAN model --- the performance index --- is the dependent variable in this equation. The Use matrix is the variable term which is to be calculated in this study. Mathematically, when the two variable terms, the Site vector and the performance index, are known, the Use matrix can be calculated via the inverse operations of the LASTAN model. The equation becomes:

$$U = f(S, R)$$

The development of the mathematical model illustrated above is the highlight of this study.

1.3.2 Selecting the physical factors

The physical and chemical properties of soil affect crop production to various degrees. Including additional soil factors into the analysis may heighten the precision of results obtained. However, in order to demonstrate the applicability of a mathematical method, it is not necessary to take all soil factors into consideration. In addition, sometimes the data available includes only a few factors. Therefore, selection of the ~~physical~~ factors for this study should be on the basis of the data available.

1.3.3 Selecting sampling sites

Usually the population of a sample space contains a large amount of observations, but not all the observations will be used in an analysis. Some sampling techniques are thus adopted in order to expect that reasonable generalizations can be achieved. In this study, the random sampling technique is used which means that any site within the sample space had an equal and independent chance of being selected. However, those sites in which soil conditions are out of the fixed states of soil factors will be discarded. In addition, mean yield of long run observation is ~~more~~ representative than that of short term observation for a specific soil condition. However, due to the difficulty of obtaining data, the data of short term observation used does not impair the value of this work as a methodological demonstration which was actually the main objective of this study.

1.4 ASSUMPTIONS

The number of factors influencing crop production is tremendous. All probable influential factors are difficult to consider together in an experimental model. Some factors had to be assumed under control at a significant level. It is necessary to make these kinds of assumptions in this study.

According to economists, in agriculture, the four main factors of production could be classified into two parts: physical influence, (land); and human influence, (labour, capital and entrepreneurial skills). Also, the environmental factors such as climate and cultural influences such as social and economic factors, are significant to crop production.

1. Physical influence : Land characteristics vary infinitely and continuously. Some variations are gross and conspicuous and have radical and obvious effects upon crop production. Due to the fact that data being used in this study is obtained from the database in Land Resource Research Institute (LRRI), and only part of important factors are included, the other unemployed factors such as drainage, pH value, stoniness etc., which are absent from the data list obtained from LRRI, must be assumed at an average level and their effects on this study would be constant.

2. Human influence : Any agricultural production needs human inputs in the form of labour, capital and management or entrepreneurial skills. Apparently those inputs have a considerable influence on the level of output. In this study, however, only land characteristics are emphatically considered as analytical factors. Although this human influence (socioeconomic) data could be introduced into the LASTAN model (Lok, Phipps, 1981), it is better to assume an average level of management, which includes timely tillage, fertilization, weed control and other cultivation techniques.

3. Environmental influence : Climatic factors have paramount effects upon crop grwoth, especially in this study area. The climatic data obtained include mean daily temperature, mean precipitation, soil temperature and soil moisture. Data of mean daily temperature and mean precipitation was obtained from Crop Production Division of Agriculture Canada. As data of crop yield and soil conditions, including soil temperature and soil moisture, is recorded in tabular form, this data is difficult to accurately superimpose upon maps of mean daily temperature and mean precipitation. Thus, mean daily temperature and mean precipitation data has to be foregone. Only soil temperature and soil moisture are taken into account as climatic factor in this study.

4. Cultural influence : Social, economic and cultural factors are not considered. These factors include transportation facility, location, intensity of land use, tilling technique, habit of cultivation, etc..

1.5 STUDY AREA

All mathematical models for land evaluation must be tested by field data to ensure that these models are applicable. Hence, this study requires field data to test its applicability. Data of the Peace River area in British Columbia has thus been selected for this purpose.

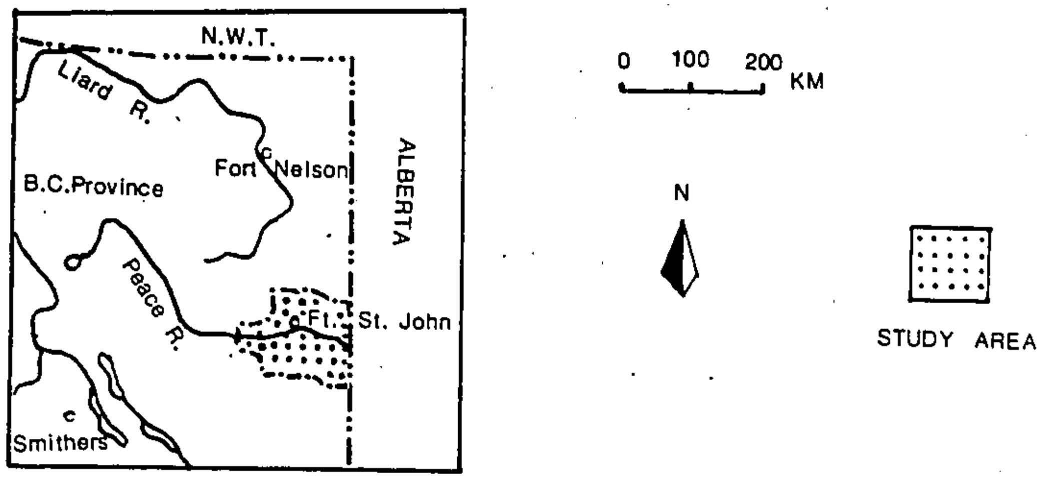
1.5.1 Location

The Peace River area is located in the northeastern part of British Columbia, and comprises part of the Peace River drainage basin lying west of Alberta (Farstad et. al. 1965). It includes the intermountain valleys of the Pine, Peace and Halfway rivers. The total area, which lies partly in the Rocky mountain foothills and partly in the high plains, is about 1,745,242 hectares (see figure 1). Surface and soil drainage varies greatly throughout the area. The most satisfactory drainage conditions, from the viewpoint of agriculture, are found on long, gentle slopes in the Peace River basin. Cereal grains, such as wheat, oats and barley, are grown in the greater part of the cultivated land which centres mainly on the broad platform or basin that borders the Peace river.

1.5.2 Soil

Soil contains the organic material on the surface of the land which is capable of supporting plant growth. Soil types are grouped according to the way in which they are formed, and each group is associated with a particular environment (Farley, 1979). According to the Canadian System of Soil Classification, there are several kinds of soils represented in the Peace River area. Table 1 shows the 5 orders, 9 great groups, and 16 sub-groups of soils in this study area (see table 1).

Figure 1 : Location Map of the Peace River Area



Source: Farstad 1965

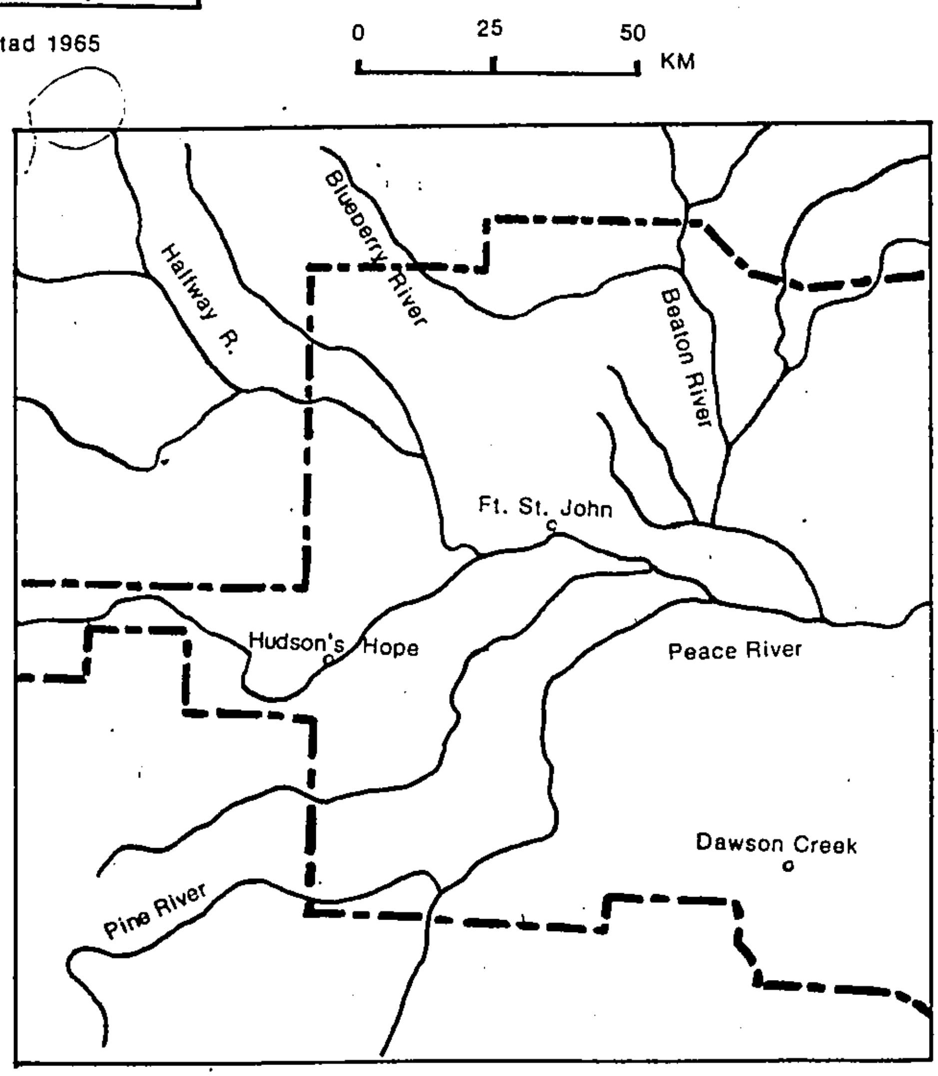


TABLE 1

Soils of Peace River area

Order	Great Group	Sub-Group
Brunisolic	Eutric Brunisol	Orthic Eutric Brunisol
		Eluviated Eutric Brunisol
Chernozemic	Dark Brown	Eluviated Dark Brown
	Black	Rego Black
		Eluviated Black
		Solozetzic Black
Gleysolic	Humic Gleysol	Orthic Humic Gleysol
	Gleysol	Fera Gleysol
	Luvic Gleysol	Orthic Luvic Gleysol
Luvisolic	Gray Luvisol	Orthic Gray Luvisol
		Dark Gray Luvisol
		Brunisolic Gray Luvisol
		Solonetzic Gray Luvisol
		Gleyed Solonetzic Gray Luvisol
Solonetzic	Solonetz	Black Solonetz
	Solod	Black Solod

Source: Canadian System of Soil Classification, 1976 revised.

1.5.3 Climate

The study area is characterized by a moderate, continental climate, moderately warm in summer and relatively cold in winter. The temperature ranges from a mean daily minimum of -16.7 in January to a mean daily maximum of 15.6 degree Celsius in July at Fort St. John. The mean annual temperature appears to be about 1.7 degree Celsius in this area. The annual precipitation varies from 16 to 18 inches. About 40 to 50 percent of this falls during the growing season by the way of local thunderstorms. June and July are the wettest months (Farley, 1979).

In sum, the study area is a predominantly rural zone. Agriculture and related processing industries are important economic activities of this area. Thus, this region lends itself well for the research undertaken in this study.

Chapter II

THE LASTAN MODEL

2.1 OVERVIEW

A major problem of numerical methods of determining land suitability was their inability to clarify interactions among land characteristics (Lok and Phipps, 1981). This interaction is defined as the difference between the effect that two coexisting factors have, and the effect generated by each factor in isolation on the contribution to soil suitability. The aim of improving this inaccuracy of ignoring interactions for land suitability stimulated the development of the computerized LASTAN model. This model is based on the assumption that interactions may be positive (synergistics), neutral (no effect or interaction) or negative (limiting), depending upon their contribution to land use. Through a series of matrix algebra calculations, including a matrix transposition, those effects were incorporated. The final output generated by the LASTAN model is a series of performance indices representing the soil suitability of each site for a particular land use type.

The following sections of this chapter will describe the elements, implementation and characteristics of the LASTAN model. Its advantages and defects will then be discussed.

2.2 STRUCTURE OF THE LASTAN MODEL

2.2.1 Site Vector

The Site vector, 1(S)m, represents a single site or homogeneous land area. It records properties of land factors within this unit, and particularly those factors which are pertinent to the land use, (i.e. stoniness, soil texture, pH value, topography and so on). Each factor section is further subdivided into individual entries of this vector. Each entry represents a state of a factor. A set of binary values (0 or 1) is employed on the Site vector. In other words, when the value of an entry is 1, it indicates that this particular land factor state is present; when this value is 0, it indicates that the particular factor's state is absent.

According to the scaling system, the data representing the physical condition or state of soil, may usually be classified as a nominal measurement level, which is employed when things are distinguished only on the basis of qualitative considerations without any implication of a quantitative relationship (Robinson, et. al. 1978). Although some factors or variables could be quantitatively defined, for example, pH value, slope class, etc., frequently they are substituted for the hierarchy type (qualitative states) of classification in soil suitability evaluation in order to correspond with those nominal variables. Therefore, all variables (factors) used in this study are divided into states. Each state indicates a physical situation within a variable in reality. Slope class (SLOP) for instance, was divided into three states in this study. State 1 of variable SLOP indicates the surface slope ranging from 0 to 2 %, State 2 from 2 to 5 %, and State 3 from 5 to 9 % (see subsection 3.3.1). When a site is identified, only one state is significant per variable. For example, if only four variables are considered, only four binary values may equal 1; one '1' value for each variable. These '1' values represent those states which are present at these sites. The remaining values must equal zero, and these zero values represent those states which are absent at this site. For example: if only four physical factors : Var 1 (slope), Var 2 (soil texture), Var 3 (drainage) and Var 4 (pH value) are considered, and Var 1 is divided into three states whereas the others are divided into four states each, also suppose that the physical situations of a hypothetical Site are that state 'b' of slope, state 'a' of soil texture, state 'd' of drainage and state 'b' of pH value, then, the hypothetical Site vector could be presented as shown in figure 2.

Figure 2: A hypothetical Site vector

Factors	---	Var 1	Var 2	Var 3	Var 4
States	---	a b c	a b c d	a b c d	a b c d
Site X	---	0 1 0	1 0 0 0	0 0 0 1	0 1 0 0

2.2.2 Use matrix

The Use matrix, m(U)m, represents the land requirements of a single land use type. It consists of two parts, the factor ratings, representing the degree of contributions of each possible land factors state, and the interactions of these states to the suitability of the land use type. The factor ratings, indicating the 'm' states, lies on the 'm' entries making up the diagonal of the Use matrix. The interactions are defined here as the difference between the rated effect that two coexisting variable states have, and the effect generated by each variable state in isolation (assuming other variable states are constant), upon a contribution to soil suitability (Lok, 1979). These effects may be positive, neutral or negative. Thus, the interactions between pairwise variables can form a submatrix. The total submatrices, (V*(V-1)/2, where V is the number of variables), build up the second part of the Use matrix positioned below the diagonal. The other entries of the Use matrix, including those entries to the right of the diagonal and those entries lying between the diagonal and the adjacent submatrices below it, are equal to 0. Thus, the Use matrix is a lower triangular matrix. For example, the factors and states being considered here are the same as the situations discussed before in the example of hypothetical Site vector, four factors are Var 1 (slope), Var 2 (soil texture), Var 3 (drainage), Var 4 (pH value), and each factor is divided into four states except Var 1 (slope) which is divided into three states. A hypothetical Use matrix can be seen in figure 3.

Figure 3: A hypothetical Use matrix

Factors ---		Var 1			Var 2				Var 3				Var 4			
States ---		a	b	c	a	b	c	d	a	b	c	d	a	b	c	d
	a	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Var 1	b	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
	c	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	a	2	1	-1	3	0	0	0	0	0	0	0	0	0	0	0
	b	1	-1	-1	0	3	0	0	0	0	0	0	0	0	0	0
Var 2	c	1	2	-2	0	0	1	0	0	0	0	0	0	0	0	0
	d	-2	1	1	0	0	0	2	0	0	0	0	0	0	0	0
	a	4	2	1	-2	-1	1	1	4	0	0	0	0	0	0	0
	b	2	1	2	3	1	2	1	0	2	0	0	0	0	0	0
Var 3	c	-3	-1	-1	-1	-1	-2	-2	0	0	2	0	0	0	0	0
	d	-2	-1	-1	-2	-1	-2	-1	0	0	0	3	0	0	0	0
	a	2	2	1	2	3	1	1	-1	2	1	3	2	0	0	0
	b	1	1	2	1	2	1	-1	3	2	2	1	0	3	0	0
Var 4	c	-1	1	-2	-1	2	1	-1	4	2	1	2	0	0	3	0
	d	-3	-1	-1	2	1	-1	2	3	1	1	1	0	0	0	1

2.2.3 The operations of the LASTAN model

From the operations of the LASTAN model, as described by Lok (1979), the indices of land suitability are calculated in two steps. Firstly, according to the laws of matrix algebra, the dot product of a Site vector, 1(S)m, and the Use matrix, m(U)m, is calculated. Then, a new row vector 1(I)m is generated, which has the same orders as the Site vector.

$$1(S)m * m(U)m = 1(I)m$$

Before the second operation starts, the new generated row vector 1(I)m has to be transposed becoming a column vector m(I.t)1. The scalar product of the multiplication between the Site vector 1(S)m and the transposed new column vector m(I.t)1, is the index of land suitability (R) of the site represented by Site vector for the land use represented by the Use matrix.

$$1(S)m * m(I.t)1 = (R)$$

Therefore, (R) is an arbitrary evaluation of the land site suitability which will then be classified according to a set of suitability classes.

When the mediate row vector 1(I)m is generated, values contained in its 'm' entries are ratings, interactions and combinations of these two. When the Site vector 1(S)m is multiplied with the transposed column vector m(I.t)1, its scalar product (R), incorporates those corresponding ratings and interactions. The LASTAN model is favoured over other approaches because it accounts for interactions between soil characteristics and then uses these interactions to derive the indices. Suppose the same hypothetical Site discussed before is exemplified here again, the physical situations of this site are : state 'b' on Factor 1 (slope), state 'a' on Factor 2 (soil texture), state 'd' on Factor 3 (drainage) and state 'b' on Factor 4 (pH value). Then, the result of the LASTAN model for this hypothetical site can be obtained from the following equation.

$$\begin{aligned} R = \; & E1b + E2a + E3d + E4b \\ & + I1b2a + I1b3d + I1b4b \\ & + I2a3d + I2a4b + I3d4b \quad \ldots\ldots \quad (\text{Eq.1}) \end{aligned}$$

where R is the performance index of LASTAN model for a hypothetical site;

E1b, E2a, E3d, E4b are effects of each state, which are present at this hypothetical site; and

I1b2a, I1b3d, I1b4b, I2a3d, I2a4b, I3d4b are pairwise interactions between coexisting states.

From the above equation (Eq.1), the LASTAN model is an additive parametric model. However, the distinction is that not only are the effects of each factor's states, which contribute to soil suitability of a specific land use type considered in the LASTAN model, but also the 2-way interactions between states are incorporated.

2.3 GENERAL CHARACTERISTICS

2.3.1 Advantages

The conventional additive parametric model postulates that each factor operates without mutual interference. Although the LASTAN model is an additive parametric model, this method overcomes the principal flaw of conventional additive techniques which ignored the interactions among factors. In addition, there are several other advantages which are as follows:

1. It is a computerized model. When the data of evaluating sites are collected and coded, they can be manipulated quickly and precisely by computer. It allows rapid assessment of large geographical areas composed of many land units simultaneously.

2. The technique is simple and inexpensive. Once the input data, including site conditions and land use requirements is established, various land use types can be evaluated for a site.

3. To improve the precision of this model, the Use matrix could be expanded easily to incorporate many additional variables, such as climatic and economic factors (Lok, 1979).

2.3.2 Disadvantages

Although the LASTAN model has many advantages as described above for land suitability evaluation, it obviously has some drawbacks as well.

1. The principal problem perceived in using this model lies in the difficulty of accurately assessing the numerical value of interactions. They rely upon the expertise of pedologists and agronomists to assess the effects of isolated states and pairwise interactions.

2. Even though the effects and interactions are estimated by experts in this field, agronomists and pedologists, this model may easily be criticized as being lacking objectivity owing to the absence of experimental support.

Chapter III

THE GENERATION OF THE COMPUTED USE MATRIX

3.1 OVERVIEW

Mathematically, when an equation is established, the independent variables (factors being considered) must be known, then, the dependent variable (result of this equation) can be calculated. In the LASTAN model for instance, the operations of this model are that first, calculating the dot product of a Site vector 1(S)m (independent variable) and the Use matrix m(U)m (independent variable), and secondly, transposing the dot product from row vector 1(I)m to column vector m(I.t)1, then the scalar product of the multiplication between the Site vector and the transposed column vector is the result of the LASTAN model. By equations, these operations can be expressed as :

$$1(S)m * m(U)m = 1(I)m \;\ldots\ldots\; (M\text{-}1)$$

$$1(S)m * m(I.t)1 = (R) \;\ldots\ldots\; (M\text{-}2)$$

The independent variables (1(S)m and m(U)m) must be known, then the dependent variable (R) may be computed. Unfortunately, the Use matrix m(U)m is unknown, if it has not already been estimated. In this case, if the other two of the three dependent and independent variables ((R) and 1(S)m) are known, through a series of converse mathematical operations, the unknown variable m(U)m could be calculated as well.

The aim of this chapter is to illustrate how the Use matrix, m(U)m, is calculated. The various steps are described in detail in the following sections.

3.2 THE PERFORMANCE INDICES

As shown in the above equations, (M-1) and (M-2), there are three variables: 1(S)m, the Site vector; m(U)m, the Use matrix; and (R), the result of performance index of a specific site. In reality, only the Site vector could be deduced from soil surveys. The others are unknown. The Use matrix, m(U)m, is the term which is to be solved in this study. The result or performance index (R) is obtained through the operations of the LASTAN model. Due to the fact that one of the independent variables, the Use matrix m(U)~~m,~~ is unknown, it seems impossible to produce the scalar result (R) via the LASTAN model. As mentioned in the prior section, if and only if those two of the three variables in the combined equations of the LASTAN model are known, the unknown variable can be solved through a series of mathematical operations. To calculate the result, (R), is the first problem faced in this study.

The result of the LASTAN model is the performance index of land suitability of a site for a particular crop production or land use type. All crops will not grow equally well on any given soil. The highest index specifies the soil most suited to that crop yielding the highest production. In other words, the performance index in the LASTAN model is the integrated contribution of the isolated effects of each land factor and the interactions between coexisting states. The crop yield is the result of these contributions. Performance index and crop yield play the same roles on land evaluation. Owing to the fact that the performance index does not exist before the operations of the LASTAN model, the homogeneous data which may be obtained directly from soil survey and investigations is the crop yield. Therefore, crop production is to be used instead of the performance index (R), in calculating the Use matrix.

However, crop yield results from all possible contributions combined together. In this study, only parts of significant physical characteristics are considered. Therefore, the crop yields could not replace the results or performance indices of the LASTAN model directly. Also, due to the fact that the performance indices are used to calculate the 2-way interactions, the joint contributions of each coupled coexisting factors on crop production must be calculated in advance in order to determine to what degree these pairwise factors affect the crop yield. The magnitudes of those strengths may be calculated by a descriptive statistic, ETA square, in the ANOVA subprogram of SPSS (Nie et. al., 1975). The technique of computing the ETA square will be illustrated in the following section. The scalar product of multiplying the crop yields by the ETA square values is used to replace the result or the performance indices (R) of the LASTAN model. Thus, the result (R) is identified in this study.

3.3 CALCULATION OF FACTOR RATINGS

The Use matrix is composed of two important parts. One is the factor ratings, the other is the interactions. The factor ratings, which occupy the diagonal entries of the Use matrix, represent the degree of isolated contributions of land factor states. In Lok's case study, each variable considered in the LASTAN model was divided into five classes which construct the rating system; each integral point standing for a corresponding hierarchy level. For instance, a state of '5' stands for excellent, '4' for good, '3' for mediocre, '2' for poor and '1' for unsuitability for a particular land use type. Those rating points are assigned to fill the entries of the diagonal to frame the factor ratings according to the experience of specialists in this field.

In this section, a method of computing factor ratings will be described, and its computed factor ratings will be used to replace the rating system in the LASTAN model.

3.3.1 Variable selection

Selection of the independent variables should be based on the data availability as well as suitability. The physical and chemical soil attributes may contribute to crop production, either in isolation or in combination with each other. However, due to the lack of data available, it is difficult to combine all the physical and chemical factors together into a model for analysis. Only some of momentous soil characteristics were selected for this study. The data, which were obtained from Land Resource Research Institute, include information on five significant soil/land factors. These factors are slope class, surface texture, soil order, soil subgroup and climate area. Due to the fact that soil order and soil subgroup are levels of soil classification in Canada, the information contained in soil order levels is comprised in that information contained in soil subgroup levels. Hence, the soil order was dropped, leaving only four factors or variables which are considered in this study.

3.3.1.1 Slope class

According to the Manual for Describing Soil in the Field, the slope gradient is divided into ten classes (see table 2). However, there is no sampling site on which slope is steeper than 9%. There are only three sampling sites belonging to class 1 (0 - 0.5%). For the sake of simplicity and avoiding erroneous generalizations, class 1 (0 - 0.5%) and class 2 (0.5 - 2%) were merged into one class (0 - 2%). Therefore, there are only three states of slope class in this study (see table 3).

TABLE 2

Slope classification

Class	Slope (%)	Approximate degrees	Description
1	0 - 0.5	0	level
2	0.5 - 2	0.3 - 1.5	nearly level
3	2 - 5	1.5 - 3	very gently sloping
4	5 - 9	3.5 - 5	gently sloping
5	9 - 16	6 - 8.5	moderately sloping
6	16 - 30	9 - 16.5	strongly sloping
7	30 - 45	17 - 24	very strongly sloping
8	45 - 70	25 - 35	extremely sloping
9	71 - 100	36 - 45	steeply sloping
10	> 100	> 45	very steeply sloping

Source : Manual for describing soils in the field, 1976 revised.

3.3.1.2 Surface texture

Textural classes are estimated in terms of size and the distribution of primary particles. Although more than twenty five field textures have been defined in Canada, only nine textures appear in the Peace River area. Within these nine, some soil textures such as organic and sand have site samples less than ten frequencies in the population of the sample space (totally 1,246 sites for barley). For the sake of avoiding zero samples and erroneous generalizations, only the top five textures in order of their sampled sites: Loam (622), Sandy loam (257), Clay loam (96), Silt loam (82) and Loamy sand (78), are selected as states of this variable. Hence, this variable is further divided into five states (see table 3).

3.3.1.3 Soil subgroup

Soil subgroup is the third level of soil classification in the Canadian system. Subgroups are formed by subdivisions of each great group. They not only carry the differentiating criteria of the order level and great group level, but also of the subgroup itself, which are differentiated on the basis of the type and arrangement of horizons (Canadian system of soil classification, 1978 revised). Within the sixteen subgroups which exist in the Peace River area as delineated in Table 1, only five of them have a frequency of site samples over one hundred. So, those five subgroups are adopted as class states of this factor (see table 3).

3.3.1.4 Climate area

The climate described here does not include mean precipitation or mean daily temperature for growing crops, but rather the soil temperature and soil moisture. There are five soil temperature classes : 1. Arctic, 2. Subarctic, 3. Cryoboreal, 4. Boreal, 5. Mesic in Canada, and more than ten groups of soil moisture. Their combination forms the criteria of climate area of this study. The soil temperature of the Mesic class is mild to moderately warm. It seems illogical to have the soil temperature of this class in the Peace River area. However, from the soil survey, investigated in conjunction by the Land Resource Research Institute and British Columbia Ministry of Agriculture at the Peace River area, it shows that there are only 5 samples belonging to the Boreal class, and 44 to the Mesic class of the soil temperature in the total sample of 1,246 sites for barley. Therefore, the Boreal class of the soil temperature is discarded and five climate area groups are defined. They are, thus, the states of this variable (see table 3).

TABLE 3

Variables, states and descriptions used in this study

Variable	States	Description	
Slope (SLOP)	0 - 2%	level to nearly level	
	2 - 5%	very gently sloping	
	5 - 9%	gently sloping	
Texture (TXTR)	Loam		
	Sandy loam		
	Clay loam		
	Silt loam		
	Loamy sand		
Subgroup (SBGP)	GL-D	Dark Gray Luvisol	
	LG-O	Orthic Luvic Gleysol	
	BL-E	Eluviated Black	
	GL-O	Orthic Gray Luvisol	
	HG-O	Orthic Humic Gleysol	
Climate (CLIM)	1C**	Arctic	Subaquic
	2GF	Subarctic	GF
	3GF	Cryoboreal	GF
	5GF	Mesic	GF
	5F	Mesic	Subhumid

** 1. arctic : extremely cold;
2. subarctic : very cold;
3. cryoboreal : cold to moderately cold;
5. mesic : mild to moderately warm;
C. subaquic : soil saturated for short periods;
F. subhumid : soil dry in some parts when soil temperature is > 5 degree C in some years, water deficits 6.4 to < 12.7 cm;
G. semiarid : soil dry in some parts when soil temperature is > 5 degree C in most years, water deficits 12.7 to 19.1 cm;
GF. dominantly semiarid to subhumid.

NOTE : As mentioned on page 36, the present of class 5 in the Peace River area is illogical. It has been left for computation because it is included in the data list obtained.

3.3.2 Weights

The conventional additive parametric technique is adopted to calculate the contributions of individual factors to crop production. The LASTAN model, in addition, summed up those individual contributions and 2-way interactions to generate indices of sites for a specific land use. When crop yields are used instead of the performance indices, the factor ratings must yield information concerning these crop yields in order to cope with the results of the LASTAN model. Hence, the extent of effect of each factor is to be measured and used as a weight in computing factor ratings. In the ANOVA subprogram of SPSS (Analysis of Variance), a descriptive statistic, ETA square, is used to measure the strength of effects of independent variables (physical factors) on dependent or criterion variable (crop yield). This technique, therefore; is to be employed here to calculate the weight of each factor contributing to the crop production.

The basis of analysis of variance is the decomposition of variation or sums of squares corrected for the mean (Kim et. al. 1975). Suppose only one factor (drainage) is taken into account and that this factor is further divided into several groups or states. The structual model, sometimes called a linear model, for each observation or site (production) can be written as :

$$Pij = GM + Ei + Rij \quad \ldots\ldots\ldots \quad (1)$$

where Pij is the dependent variable (production) for the jth site in group i;

GM stands for grand or overall mean of all sites;

Ei is the effect of being in group i (effect of the independent variable); and

Rij is the effect on this observation of all other variables (residual);

The deviation of a group mean from the grand mean is taken to represent the effect on each observation from belonging to that particular group Ei (Iversen and Norpoth 1978). Thus, Ei could be present as :

$$Ei = Pi - P \quad \ldots\ldots\ldots \quad (2)$$

where Pi is the mean that appear in group i, (group mean);

P is the grand or overall mean;

And the deviation of an observation from its group mean is taken to represent the effect on that observation of all variables other than the group variable (drainage) (Iversen and Norpoth 1978). Rij can also be presented as :

$$Rij = Pij - Pi \quad \ldots\ldots\ldots \quad (3)$$

By replacing the terms GM, Ei, Rij in (1) by P, (Pi - P), (Pij - Pi) respectively, the structual model can be rewritten as :

$$Pij = P + (Pi - P) + (Pij - Pi) \quad ..... \quad (4)$$

$$\Longrightarrow (Pij - P) = (Pi - P) + (Pij - Pi) \quad ..... \quad (5)$$

The sum of a group of squared deviation scores makes a variation or sum of squares (Schmidt 1979). In ANOVA, three sum of squares are computed by squaring both sides of the equation above.

$$\sum_i \sum_j (Pij - P)^2 = n\sum_i (Pi - P)^2 + \sum_i\sum_j (Pij - Pi)^2 \quad ..... \quad (6)$$

or denoting as :

SStotal = SSbetween + SSwithin

The first term, the total variation of individual scores around the grand mean, is equal to the sum of two components, the 'between groups sum of squares' and the 'within groups sum of squares'. The second term, SSbetween, is called the 'between groups sum of square', because it contains the deviation of each group mean around the grand mean, and represents how extensive the effect of the independent variable (drainage) is on the dependent variable (production). And the third term, SSwithin, is called the 'within groups sum of squares', because it contains the deviation of each score in a group around the group mean and shows the extent of the effect, of all other variables upon the dependent variable.

In the ANOVA subprogram, a descriptive statistic, ETA square, is the ratio of the sum of the squares of the explanatory (independent) variable to the total sum of squares.

$$\text{ETA square} = \frac{\text{SSbetween} \; \text{---} \; n \sum_{i} (P_i - P)^2}{\text{SStotal} \; \text{---} \; \sum_{i} \sum_{j} (P_{ij} - P)^2}$$

It represents how much of the variation in the dependent variable is explained by the explanatory variable. In this study, this value is used as the weight of the independent variable (physical factor) on the dependent variable (production).

3.3.3 Factor ratings

The factor ratings represent the strength of isolated contributions of land factor states. Although the magnitude of ETA square for each factor indicating the proportion of variation in the dependent variable (production) explained by each factor can be calculated, the contribution of state level cannot. The differences between states influencing the crop production could be displayed on the mean yields based on the statistics of states level. Hence, the multiplication of mean yield of each state and the factor weight which the state belongs to frames the factor ratings.

3.4 COMPUTATION OF INTERACTIONS

When all the necessary data, including Site vectors n(S)m, results or crop yields (R), and factor ratings in the Use matrix are completed, then the principal part of this study, computing the interactions to generate the Use matrix m(U)m, will be presented. Following the converse operations of LASTAN model it is not difficult to calculate values of interations in the Use matrix. For example, there are three physical factors being considered as main variables which have significant influences on a particular crop production. They are named Factor A, B and C. Factor A and C are divided into two states respectively, whereas Factor B is divided into three states. Then, a hypothetical Site vector 1(S)7 would be present as :

S1	S2	S3	S4	S5	S6	S7

where the first two entries S1, S2 indicate the site condition about Factor A; the next three entries S3, S4, S5 illustrate Factor B; the last two entries, S6 and S7 represent Factor C. The Use matrix 7(U)7, can be divided into two components, the factor ratings and submatrices of interactions. The factor ratings are denoted as :

A1	A2	B1	B2	B3	C1	C2

which occupied the entries of diagonal of the Use matrix. They are computed by multiplying the mean yield based on the statistics of each state with factor weight which the state belongs to. This method was illustrated in the above section.

Owing to the fact that only three physical factors, Factor A, B, C are considered, there are three submatrices of interactions (V*(V-1)/2, where V, the number of considered factors, is three) in the hypothetical Use matrix. Submatrix X stands for the submatrix of interactions between Factor A and B. There are 2 states in Factor A and 3 states in Factor B, so, six interactions (2 by 3) X1, X2, X3, X4, X5, X6 are in this submatrix X. Submatrix Y, the submatrix of interactions between Factor A and C, contains four interactions (each factor consists of two states) Y1, Y2, Y3 and Y4. And Submatrix Z, the submatrix of interactions between Factor B and C, comprises six interactions (3 by 2) Z1, Z2, Z3, Z4, Z5, Z6. Thus, the hypothetical Use matrix 7(U)7 is denoted as follows :

A1	0	0	0	0	0	0
0	A2	0	0	0	0	0
X1	X2	B1	0	0	0	0
X3	X4	0	B2	0	0	0
X5	X6	0	0	B3	0	0
Y1	Y2	Z1	Z2	Z3	C1	0
Y3	Y4	Z4	Z5	Z6	0	C2

From the operations of the LASTAN model, the steps of computing performance index of this hypothetical site are :

(1) 1(S)7 * 7(U)7 = 1(I)7

The operations of this step, thus, are composed of

S1*A1 + S3*X1 + S4*X3 + S5*X5 + S6*Y1 + S7*Y3 = (I)1

S2*A2 + S3*X2 + S4*X4 + S5*X6 + S6*Y2 + S7*Y4 = (I)2

S3*B1 + S6*Z1 + S7*Z4 = (I)3

S4*B2 + S6*Z2 + S7*Z5 = (I)4

S5*B3 + S6*Z3 + S7*Z6 = (I)5

S6*C1 = (I)6

S7*C2 = (I)7

This step combines the isolated contribution of each state and all of the possible interactions with respect to this state into entries of a new generated row vector separately. Before the second step starts, the new row vector 1(I)7, has to be transposed becoming column vector 7(I.t)1, and then the Site vector 1(S)7 multiplied with this transposed column vector 7(I.t)1 to generate the result index.

(2) 1(S)7 * 7(I.t)1 = (R)

S1 * 1(I.t) + S2 * 2(I.t) + S3 * 3(I.t) + S4 * 4(I.t)
+ S5 * 5(I.t) + S6 * 6(I.t) + S7 * 7(I.t) = (R)

This step incorporates those corresponding ratings and interactions into its scalar product (R), the performance index of the LASTAN model. To extend this equation, an equation for computing performance index of LASTAN model can be obtained.

S1*S1*A1 + S1*S3*X1 + S1*S4*X3 + S1*S5*X5
+ S1*S6*Y1 + S1*S7*Y3 + S2*S2*A2 + S2*S3*X2
+ S2*S4*X4 + S2*S5*X6 + S2*S6*Y2 + S2*S7*Y4
+ S3*S3*B1 + S3*S6*Z1 + S3*S7*Z4 + S4*S4*B2
+ S4*S6*Z2 + S4*S7*Z5 + S5*S5*B3 + S5*S6*Z3
+ S5*S7*Z6 + S6*S6*C1 + S7*S7*C2 = (R) .... (EQ.3-1)

There are no unknown variables within those terms, S1*S1*A1, S2*S2*A2, S3*S3*B1, S4*S4*B2, S5*S5*B3, S6*S6*C1, S7*S7*C2 in equation (EQ.3-1), Let

$$R = (R) - S1*S1*A1 - S2*S2*A2 - S3*S3*B1 - S4*S4*B2 - S5*S5*B3 - S6*S6*C1 - S7*S7*C2$$

Then the above equation (EQ.3-1) can be rewritten as :

$$\begin{aligned} & S1*S3*X1 + S2*S3*X2 + S1*S4*X3 + S2*S4*X4 \\ & + S1*S5*X5 + S2*S5*X6 + S1*S6*Y1 + S2*S6*Y2 \\ & + S1*S7*Y3 + S2*S7*Y4 + S3*S6*Z1 + S4*S6*Z2 \\ & + S5*S6*Z3 + S3*S7*Z4 + S4*S7*Z5 + S5*S7*Z6 \\ & = R \text{ ....(EQ.3-2)} \end{aligned}$$

There are sixteen unknown variables (X1, X2, X3, X4, X5, X6, Y1, Y2, Y3, Y4, Z1, Z2, Z3, Z4) in equation (EQ.3-2). It needs sixteen homogeneous equations to establish a set of linear equations to get the solution of this combined linear equations. When those variables are solved, the Use matrix 7(U)7, is then obtained.

3.5 CASE STUDY

To further clarify how the model of computing a Use matrix in the LASTAN model works, the set of real data gathered in the Peace River area has been tested. It includes the three most important cereal crops, barley, wheat and oat in this study area. For example, 150 sites of barley, or about one eighth of the total investigated sites, have been selected as input data for computing statistical values including overall mean, ETA squares and so on. The steps of generating the Use matrix are illustrated as follow:

(1) Deciding the physical factors and states of each factor is the commencement of this computation. Due to the limitation of data obtained, only part of the acquired soil attributes, as mentioned in subsection 3.3.1, are taken into account. They are slope class (SLOP), surface texture (TXTR), soil subgroup (SBGP) and climate area (CLIM). These factors and their states are described in table 3. However, the sites having been selected showed that none of them appears at state HG-O (Orthic Humic Gleysol) in Factor soil subgroup, and states 1C & 5F in Factor climate area. Therefore, these three states were dropped in computing the Use matrix for barley, but were retained for calculating the Use matrices for other crops.

(2) When the factors and states are decided, the data is transformed into metric (for yield) and non-metric (for states) data in order to be put into an ANOVA subprogram for computing ETA squares which represent the magnitude of strength of each factor contributing to crop production. The formulations and processes of computing ETA squares are illustrated in subsection 3.3.2. The output of this analysis is showed in table 4.

TABLE 4

Weight of each factor for barley

Factors	SLOP	TXTR	SBGP	CLIM	TOTAL(%)
Weights	2.89	1.00	8.41	6.25	15.5

In this study, although four soil factors or characteristics are taken into account, they do not play the equivalent roles on agricultural production. As shown in table 4, the weights of soil subgroup and climate area are higher than that of slope class and surface texture. This is because that information, including soil subgroup and climate area which has significant contribution on crop production, is much more important than that of slope class and surface texture. For instance, the climate area information contains soil temperature and soil moisture; in the soil subgroup, it contains properties of pedons, the effects of dominant, soil-forming processes, the properties that reflect differences in the strengths of dominant processes, the arrangement of horizons, and so on (The Canadian System of Soil Classification). However, information in slope class and surface texture is singular. The variety of information contained in each factor results in their difference on weights.

The statistical mean yield of each state is also calculated and shown in table 5.

TABLE 5

Statistical mean yield of barley

	state 1	state 2	state 3	state 4	state 5
SLOP	44.53	46.62	46.00		
TXTR	46.12	45.33	46.08	47.29	44.90
SBGP	47.80	44.49	45.38	46.31	
CLIM	46.65	43.35	45.67		

*** Unit : Bushel

To multiply these means with their weights respectively, the factor ratings, therefore, are generated as shown in Table 6.

TABLE 6

The factor ratings for barley

	state 1	state 2	state 3	state 4	state 5
SLOP	1.2869	1.3473	1.3294		
TXTR	0.4612	0.4533	0.4608	0.4729	0.4492
SBGP	4.0200	3.7416	3.8165	3.8947	
CLIM	2.9156	2.7094	2.6544		

(3) The Site vectors and result indices have to correspond with each other, because they represent the physical condition of a particular site and its crop yield respectively. In this study, only two-way or two-state interactions are expected to be calculated. In other words, each state is matched into pairs with each state of other factors and the interactions between these pairwise states are to be calculated. The Site vector and its result index have to cope with the pairwise states they represent. In this case, fifteen states are taken into account totally, eighty three pairs of states are thus generated, which represent the pairwise interactions between coexisting states and are divided into six submatrices, 15 (3 by 5) in

submatrix SLOP-TXTR, 12 (3 by 4) in submatrix SLOP-SBGP, 9 (3 by 3) in submatrix SLOP-CLIM, 20 (5 by 4) in submatrix TXTR-SBGP, 15 (5 by 3) in submatrix TXTR-CLIM, and 12 (4 by 3) in submatrix SBGP-CLIM. Therefore, there are eighty-three unknown variables (interactions) in the Use matrix, which require eighty-three Site vectors 83(S)15, and their corresponding result indices (R)83 to solve those unknowns. Parts of the Site vectors are displayed in figure 4.

Although the factor ratings in the Use matrix and the Site vectors are acquired, without the resultant indices the interactions in the Use matrix could still not be generated. The method of computing result indices for this study was described in section 3.2. In this case, it needs eighty-three result indices corresponding with the eighty-three unknown variables in the Use matrix and the Site vectors. The first fifteen indices, in accordance with the fifteen unknown variables (interactions) in submatrix SLOP-TXTR, represent the result of joint contributions between factor SLOP and TXTR. They are computed by multiplying statistical mean yields of each coupled coexisting state with the joint effect (weight) of these two factors. These indices are shown in figure 5. The others can be computed in the same way.

Figure 4: Parts of Site vectors of barley

Factors	SLOP			TXTR					SBGP				CLIM		
States	1	2	3	1	2	3	4	5	1	2	3	4	1	2	3
Site 1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
7	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
8	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
10	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
11	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
12	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
13	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
14	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
15	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
16	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
17	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
19	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
20	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0

Figure 5: Parts of the result indices of barley

		SLOPE		
		0 - 2%	2 - 5%	5 - 9%
	Loam	1.7429	1.8293	1.8022
T	Sandy-Loam	1.7983	1.7679	1.7983
Y	Clay-Loam	1.6380	1.8330	1.7550
T	Silt-Loam	1.6380	1.8599	1.7983
R	Loamy-Sand	1.7811	1.7421	1.7160

Three of these indices (Sandy-Loam:0-2, Sandy-Loam:5-9, Silt-Loam:5-9) in figure 5 are replaced by the mediate value which is calculated by the multiplication of overall mean with the joint effect of these two coexisting factors, due to the fact that there are no such soil conditions of those coexisting states in sampling sites.

The variables Site vectors 83(S)15, factor ratings (D)15 and result indices (R)83 are obtained, through the operation of a computer program, COUSMAT (COmputing USe MATrix) (see appendix A), those unknown interactions are solved, and the Use matrix 15(U)15, is, thus, acquired and shown in figure 6.

Figure 6 : The computed Use matrix for barley by considering four factors

SLOPE CLASS			SURFACE TEXTURE				
0 - 2%	2 - 5%	5 - 9%	Loam	Sandy-L	Clay-L	Silt-L	L-Sand
1.2869	0	0	0	0	0	0	0
0	1.3473	0	0	0	0	0	0
0	0	1.3294	0	0	0	0	0
-0.0052	0.0108	0.0116	0.4612	0	0	0	0
0.0581	-0.0327	0.0156	0	0.4533	0	0	0
-0.1097	0.0249	-0.0352	0	0	0.4608	0	0
-0.1218	0.0397	-0.0040	0	0	0	0.4729	0
0.0452	-0.0542	-0.0624	0	0	0	0	0.4490
-2.6969	-0.9978	-1.1869	0.5618	0.6227	0.2335	0.3026	0.8350
-1.0217	-1.0902	-1.0210	0.4283	0.6005	0.5930	0.5809	0.3854
-0.6934	-1.1435	-0.9960	0.5178	0.2542	0.5182	0.5061	0.8305
-0.9696	-1.0759	-1.0841	0.4073	0.4475	1.3645	0.5506	0.1543
-0.1651	-0.1349	-0.1442	0.0731	-0.0389	0.2089	0.1109	-0.0420
-0.6347	-0.1222	-0.0788	0.1098	0.2495	-0.7282	0.2299	0.2538
-0.3133	0.0663	-0.2238	0.0640	0.1045	0.0970	0.0849	0.1088

Figure 6 : The computed Use matrix for barley by considering four factors (continued)

SOIL SUB-GROUP				CLIMATE AREA		
GL-D	LG-O	BL-E	GL-O	2GF	3GF	5GF
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
4.0200	0	0	0	0	0	0
0	3.7416	0	0	0	0	0
0	0	3.8165	0	0	0	0
0	0	0	3.8947	0	0	0
-0.6533	-0.8922	-0.8845	-0.8337	2.9156	0	0
-1.2933	-0.8550	-0.5776	-0.6700	0	2.7094	0
-0.9262	-0.7046	-0.7226	-0.8009	0	0	2.8544

Chapter IV

DISCUSSION

4.1 OVERVIEW

As mentioned before, the purpose of this study was to attempt to find a mathematical method to calculate the Use matrix defined in the LASTAN model, in order to avoid subjective interpretation. The computer program COUSMAT is a mathematical model which meets this objective. The method used in COUSMAT was illustrated in the first chapter. Its result, the computed Use matrix, was shown in figure 6.

Mathematically, the COUSMAT program can achieve the purpose of computing the Use matrix which was defined in the LASTAN model. However, further discussion is necessary to clarify those representations of the result showed in the computed Use matrix and their significance in reality.

The Use matrix represents the contribution of the variables to the production of a specified crop. Thus, the essential soil and climate requirements for successfully growing barley need to be reviewed briefly. The compiled data also affects the result of the computed Use matrix greatly. In addition, the Use matrix is composed of two key components: factor ratings and submatrices of interactions. The compiled data and components of the Use matrix will be reviewed separately in this section.

Barley has a rather weak root system and since it has a short growth period, it grows best on porous soils, particularly fertile, well drained soil in which the plant nutrients are available in easily assimilated form. Thus, fertile, well drained, loam and clay loam produce the highest grade barley (Royen, 1954). Cold, acidic, heavy clay and sandy soils as well as organic soils are generally unsuitable for barley production.

4.2 DATA REVIEW

Data collection plays an important role in a scientific analysis. Its precision greatly influences the results of the analysis. As the emphasis of this study has been placed on methodological demonstration, incomplete data would not impair the value of this work. Obviously, however, the further demonstration of the validity of results obtained through the application of this model should require the collection of appropriate and accurate data. Some shortcomings of the data used will be disscussed to aid in future applications of this model.

4.2.1 Data adoption

Data adopted for scientific analysis must be data which was obtained under general or fair conditions. In agriculture, crop yield is affected by many uncontrollable physical, ecological and social factors, e.g., a storm may cause the destruction of an entire crop. In general, mean yield of long run observation in agriculture is more representative than that of short term observation for a specific soil condition. Most studies in agriculture have adopted mean yields of long term observations as input data for analysis. Unfortunately, since such long term mean yields were not available, the data of crop yield adopted in this study is from one year's observation. Although nothing showed that the data is grossly incorrect, it is nonetheless dangerous to take the crop yield of one year to be a general representation of average crop yields. Note, this does not impair the value of this work as a methodological demonstration which actually was the main objective of this thesis.

4.2.2 Data sampling

As declared earlier, the data used included approximately twelve hundred sites for barley, within which only one hundred and fifty sites were sampled for this analysis. The random sampling technique was used, meaning that any site within the sample space had an equal and independent chance of being selected. In fact, those sites used were selected one by one from the top of the data list, excluding those sites in which soil conditions were out of the fixed states of soil factors. This caused the result of zero samples in some cells of coexisting states. The grand mean yield of barley was thus used to compute the indices replacing those cells, in order to calculate their interactions. Of course, this influences the correctness of the computed Use matrix.

4.2.3 Data distribution

The input data was selected by random sampling. However, the sampling frequencies of states are not equal. Some states like State Sandy-Loam and Loamy-Sand in Factor surface texture (TXTR), State BL-E in Factor soil subgroup (SBGP), and State 5GF in climate area (CLIM), have frequencies of less than ten sampling sites individually.

Several samples are required to avoid erroneous generalisations which may affect results adversely. One conspicuous example is that there is only one sampling site in the soil conditions of coexisting states between State 0-2% in variable SLOP and State GL-D in variable SBGP. Also, the investigated yield of this site is very low, 29 bushels per acre only. This causes the interaction between these two coexisting states to have the lowest value (-2.6969) in the computed Use matrix. Such low interaction, therefore, results in the comparatively lower indices of those sites which contain the coexisting soil conditions 0-2% in slope class and GL-D in soil subgroup. Table 7 shows soil conditions of those comparatively lower indices of sites and their performance indices which are affected by the exception mentioned above. There are still several similar exceptions appearing in the computed Use matrix.

TABLE 7

Comparatively lower indices of sites : their soil conditions and performance indices

	SLOP			TXTR					SBGP				CLIM			
Site	1	2	3	1	2	3	4	5	1	2	3	4	1	2	3	Index
1	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	5.8017
2	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	4.5190
3	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	5.3067
4	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	5.6060
5	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	4.7750
6	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	5.4635
7	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	5.5043
8	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	3.2478
9	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1	4.9065
10	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	5.4754
11	1	0	0	0	0	0	1	0	1	0	0	0	0	1	0	4.2750
12	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	4.9636
13	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	5.9980
14	1	0	0	0	0	0	0	1	1	0	0	0	0	1	0	4.9744
15	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	5.6629

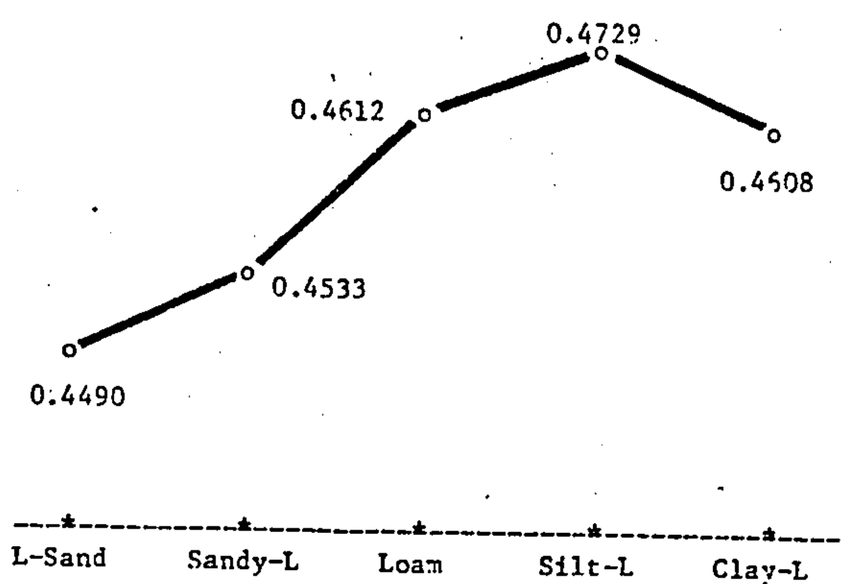
** The grandmean of performance index is 6.8655.

4.3 FACTOR REVIEW

Factor ratings are the degree of contributions of states in isolation. Generally speaking, within a factor variable, the higher rating indicates the more suitable soil condition for barley production. However, the random sampling technique used in this study sometimes causes unequal frequencies for states of factors which could result in unexpected consequences in interpretation. Two factor variables and their ratings are reviewed in the following examples.

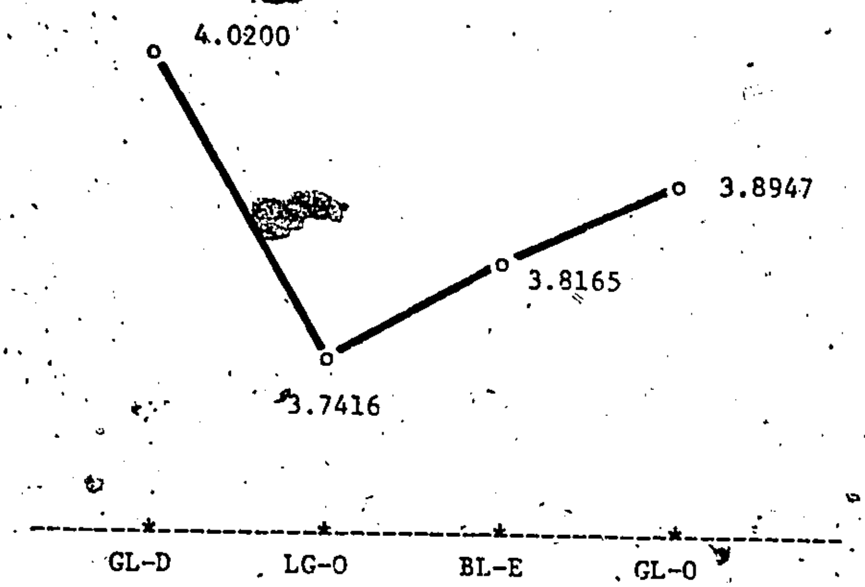
The surface texture was divided into five states --- Loam, Sandy-Loam, Clay-Loam, Silt-Loam and Loamy-Sand --- according to the size and distribution of primary particles. For example, Loam is a soil material that contains 7-27% clay, 28-50% silt, and less than 52% sand. Clay-Loam, then, contains 20-35% clay, less than 28% silt and 45% or more sand. Barley tends to need loam & clay-loam soil, which contains average percentage of clay and sand such as Loam, Clay-Loam and Silt-Loam, for optimum conditions of production. Figure 7 shows the isolated contributions of states within factor surface texture (TXTR), which framed parts of diagonal entries of the Use matrix. It represents that Silt-Loam texture contributes the highest effect on barley production within these five soil textures. The following order is Loam and then, Clay-Loam and Sandy-Loam. In comparison with the other four soil textures stated above, Loamy-Sand contributes the least to the growth of barley. This result demonstrates that the ranking of factor ratings in this study is consistent with the requirements of growing barley.

Figure 7: Contributions of soil textures



The properties of features contained in the lower level of the soil classification system are inclusive of characteristics specified in higher levels from which they are derived. The four states in factor soil subgroup (SBGP) are shown in three order levels, Chernozemic (BL-E), Gleysolic (LG-O) and Luvisolic (GL-D & GL-O). The features of soil of Gleysolic order indicate that soils in this order are comparatively unsuitable for growing barley, due to the fact that producing barley requires well-drained soils. However, Gleysolic soils are usually associated with either a high ground water table at some period of the year or temporary saturation above a relatively impermeable layer. The isolated effect of State LG-O (Orthic Luvic Gleysol) in factor soil subgroup (SBGP) which inherits features from Gleysolic order, is therefore, the lowest in comparison with other states in this factor. In general, the organic material covered in the Black soil is higher than in Gray Luvisol soil. This results in the characteristics of Eluviated Black soil: higher soil temperatures, fertility and better drainage, all of which are conducive to barley growth. Therefore, the expected rating order of states in factor soil subgroup in order of their contribution to barley production should be: BL-E (Eluviated Black) > GL-D (Dark Gray Luvisol) > GL-O (Orthic Gray Luvisol) > LG-O (Orthic Luvic Gleysol). The statistical yield means and frequencies calculated on the basis of the population observations (1,246 sites) for barley have demonstrated this expected order. However, based on the sampled observations (150 sites), the order becomes: GL-D > GL-O > BL-E > LG-O. In other words, state GL-D ranks among others; state GL-O follows; state BL-E becomes the second lowest and state LG-O is the lowest. That is because there are only 4 sampling sites for state BL-E while the other three states have more than 40 sampling sites individually. Less samples may result in erroneous generalizations. The statistical yield means on the basis of sampled observations are adopted to calculate factor ratings as the isolated contributions of states. This is to maintain data consistency and because incomplete data will not affect the value of this study as a methodological demonstration. Figure 8 indicates the differential contributions of states in factor soil subgroup on barley which were calculated on the basis of sampled observations.

Figure. 8: Contributions of soil subgroup

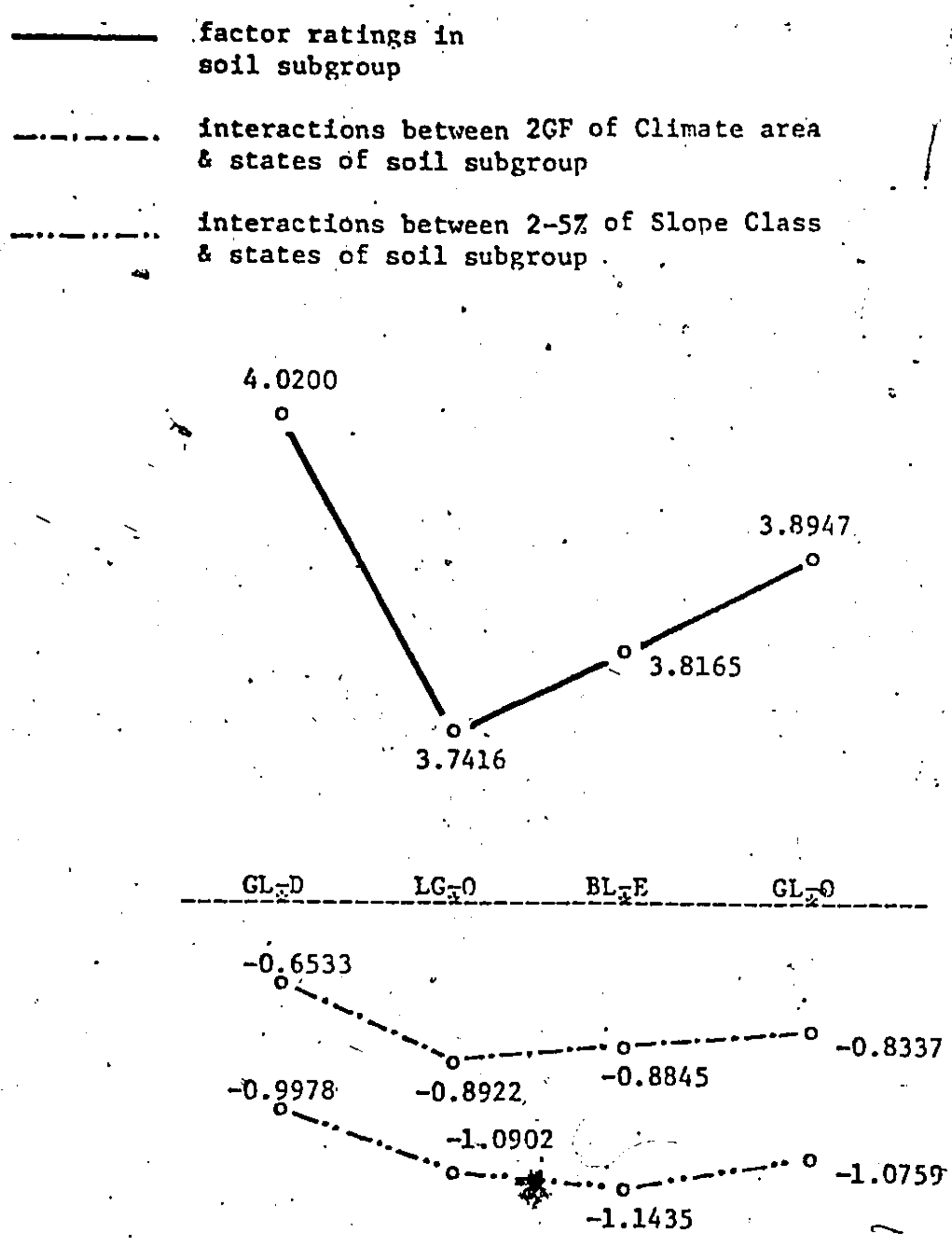


4.4 INTERACTIONS REVIEW

The interactions between coexisting states may be positive, neutral or negative. The computed Use matrix reveals that the interactions are either positive or negative. None of them is exactly neutral. This demonstrates that interactions are significant in the parametric method. However, interactions in some submatrices of the computed Use matrix, such as submatrix between slope class and soil subgroup, were all negative. Although the land factor interactions are recognized, the knowledge of their nature and effects is far from complete to date. The phenomena of all negative interactions in a submatrix may be not expected, this possibly results from that the data used are one year's observation; or the random sampling technique adopted is easy to result in zero samples. Both of these cause the consequence of erroneous generalizations.

Generally speaking, the magnitude of interactions are irregular. However, there have the tendency that interactions between a specific state and states in another factor are consistent with ratings of that factor. The highest interaction corresponds with the highest factor rating. When the contributions of states in a specific factor are lined in ascending order, the interactions between states of this factor and a state of another factor are lined in the same order. Let us take states in soil subgroup for instance: as shown in figure 9, the state LG-O contributes the lowest effect on barley. The rating of state BL-E is a little bit higher than that of state LG-O. State GL-O takes rank of state BL-E, and state GL-D ranks among these four states. Either the ascending order of interactions between state 2-5% in variable slope class, or state 2GF in factor climate area and the states in factor subgroup indicate the same order.

Figure 9 : Trends of contributions of factor ratings and interactions based on soil subgroup



Another example is shown in figure 10. It illustrates the same tendency mentioned in foregoing example. According to the magnitudes of factor ratings, the ranked order of states in variable slope class is illustrated as follows :

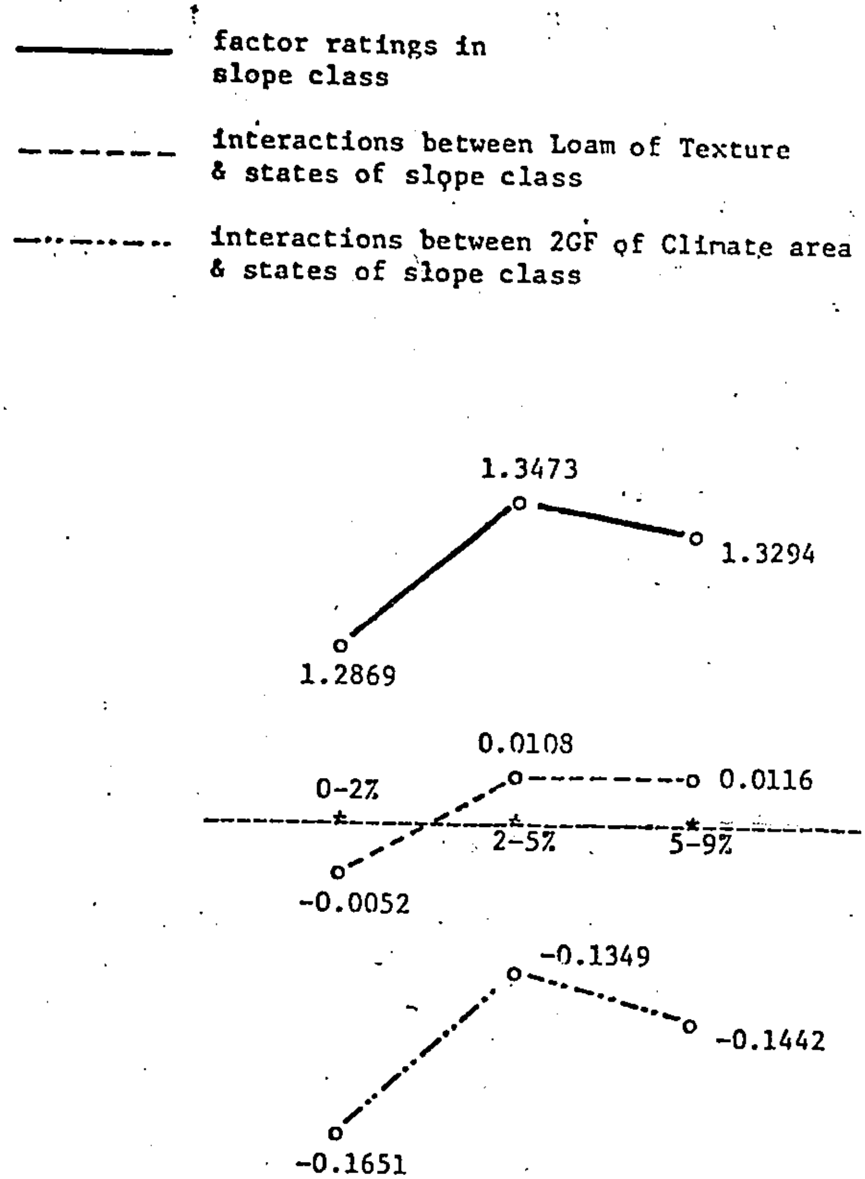
ratings ---	1.2869	<	1.3294	<	1.3473
states ---	0-2%		5-9%		2-5%

The interactions between either state Loam in factor surface texture or state 2GF in factor climate area and these three states in factor slope class pointed out the corresponding order.

magnitudes of interactions between Loam &	-0.0052 (0-2%)	<	0.0108 (2-5%)	<	0.0116 (5-9%)
magnitudes of interactions between 2GF &	-0.1651 (0-2%)	<	-0.1442 (5-9%)	<	-0.1349 (2-5%)

This trend also represented implicitly in an evaluated Use matrix of the LASTAN model in Lok's case study.

Figure 10 : Trends of contributions of factor ratings and interactions based on slope class



4.5 COMPUTED INDICES

Using this computed Use matrix for barley to re-install in the LASTAN model to substitute the evaluated Use matrix, and then through operations of the LASTAN model, a series of performance indices can be generated. Each index indicates the land suitability for barley under a possible soil condition when only four of the aforementioned factors are taken into account. Figure 11 displays parts of this series of performance indices. The highest index (8.2331) specifies the soil most suited to barley. On the contrary, the lowest index (3.2478) indicates its unsuitability to growing barley. The following two Site vectors (Site H and Site L) represent the highest and lowest indices in this study.

Factors	SLOP			TXTR					SBGP				CLIM			
States	1	2	3	1	2	3	4	5	1	2	3	4	1	2	3	Index
Site H	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	8.2331
Site L	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	3.2478

In comparison with the soil conditions of these two sites, although the state of soil subgroup occupied by Site L (GL-D) has much more contribution or effect on barley than that of being occupied by Site H (GL-O), and the soil texture are the same (Clay-Loam), the difference of computed indices between these two sites is very high, equal to 4.9853. Again, it proves the significant contributions of interactions on soil suitability evaluation.

Figure 11 : Parts of performance indices calculated by LASTAN model installing the computed Use matrix

Slope Class	Surface Texture	Soil Subgroup	Climate Area		
			2GF	3GF	5GF
		GL-D	5.8017	4.5190	5.3067
		LG-O	6.8261	6.2206	6.7916
	Loam	BL-E	7.2839	6.9907	7.2663
		GL-O	7.0688	6.5898	6.8795
		GL-D	5.8060	4.7750	5.4635
		LG-O	6.9417	6.5879	7.0597
	Sandy Loam	BL-E	6.9637	6.9222	7.0986
		GL-O	7.0524	6.8251	7.0156
		GL-D	5.5043	3.2478	4.9065
		LG-O	7.0217	5.4424	6.8844
0 - 2%	Clay Loam	BL-E	7.3152	6.0482	7.1948
		GL-O	8.0569	6.6041	7.7648
		GL-D	5.4754	4.2750	4.9635
		LG-O	6.9116	6.3884	6.8602
	Silt Loam	BL-E	7.2051	6.9942	7.1706
		GL-O	7.1450	6.7483	6.9388
		GL-D	5.9980	4.9744	5.6629
		LG-O	6.7063	6.3599	6.8317
	Loamy Sand	BL-E	7.5197	7.4856	7.6620
		GL-O	6.7389	6.5190	6.7095

Chapter V

CONCLUSIONS

5.1 OVERVIEW

Basically this thesis is concerned with lessening the drawbacks of parametric methods. Two major critical drawbacks in these methods are that:

1. interactions among factors are not taken into consideration; and
2. the values of physical factors are generated by evaluation.

The first drawback, not taking interactions into account, was solved when the LASTAN model was established. In the LASTAN model, there are two important components: the Site vector which represents the soil or climate conditions in a specific area; and the Use matrix which represents the land use requirements. Values assigned to the cells of this Use matrix are not only the contributions of physical factors, but also the two-way interactions among factors. Through a series of operations of the LASTAN model, the performance index of site suitability for a defined land use can be calculated, which incorporates those pertinent main effects of parameters and two-way interactions.

However, the Use matrix is generated on the basis of evaluation. That is, the method of generating the Use matrix is somewhat subjective. The objective of this thesis is to develop a mathematical method to generate the Use matrix defined in the LASTAN model in order to avoid the arbitrary setting up of the coefficients.

From the equations of the operations of the LASTAN model, there are three variable terms: the Site vector, the Use matrix and the performance index. Mathematically, if two of these three variable terms are known, then the third one can be calculated. In this case, the Site vector, representing the soil conditions, could be obtained from soil surveys; the Use matrix is the variable to be calculated, and the performance index is unknown before the operations of the LASTAN model. However, crop yields, which could be obtained from soil surveys as well, were used instead of the performance indices. Thus, the Site vector and the result (R) are known and the Use matrix can be calculated. It has been proven that the computer program COUSMAT is a program which achieves this objective.

5.2 COMMENTS

As previously stated, the goal of this study is essentially of a methodological nature. The work consists in finding an original solution to determine a Use matrix in a way which would avoid the arbitrary setting up of the coefficients. The computer program COUSMAT was thus developed in accordance with the inverse operations of the LASTAN model. It is therefore mathematical and less subjective. Although a test with field data in the Peace River area provides a crude evaluation of the success of this model, some unexpected consequences occured in this case study. Those exceptions are resulted from the ineligible data compilation. As discussed in section 4.3., for instance, state BL-E (Eluviated Black) in variable soil subgroup becomes the second lowest contribution on barley growth due to the fact that limited sampled sites (4 sites only) in this state results in the erroneous generalizations. Also, as discussed in 4.2.3., one exceptional site sample caused an extremely negative interaction value (-2.6969) among state 0-2% in variable slope class and state GL-D in variable soil subgroup. In short, it has been proven that this methodological demonstration is applicable. However, the validity of this model rests greatly on the validity of data compilation.

5.3 APPLICATION OF COUSMAT PROGRAM

Obviously, when a large number of nonmetric factors are taken into account, not only the two-way interactions but also three-way, four-way and higher way interactions among factors could be generated. However, in the LASTAN model, attention is placed on main effects of soil conditions (factor ratings) and two-way interactions. The others, such as three-way and higher-way interactions are ignored. Mathematically, the technique used in the COUSMAT program is not suitable for computing three- or higher-way interactions due to features inherent in the LASTAN model. In the LASTAN model, the Site vector is a dualized vector. It means that the values in entries of this vector are binary values, either 1 or 0. In COUSMAT, a homogeneous matrix is used to solve the linear equations in order to gain solutions of two-way interactions. Within these operations, its inverse matrix has to be calculated. Assuming that three-way interactions are to be calculated, such matrix can easily become a single matrix which does not have inverse matrix. Thus, no solution can be obtained. Therefore, if the three-way or higher-way interactions are required, a more complicated program must be designed instead of the COUSMAT program. Generally speaking, as described by Kim, the interpretation of higher-order interaction effects may become problemetic or may be of little interest (Kim, 1975), because when the orders increase their weights decrease.

In conclusion, when the aforementioned defects of data processing have been corrected, the COUSMAT program, is just the mathematical method needed to calculate two-way interactions for the LASTAN model.

BIBLIOGRAPHY

Canada Soil Survey Committee, The Canadian System of Soil Classification, Ottawa : Research Branch, Canada Department of Agriculture, 1978.

Cook A. H., Barley and Malt, New York : Academic Press, 1962.

Dumanski J., Marshall I. B., Huffman E. C., Soil Capability Analysis for Regional Land Use Planning---a study of the Ottawa urban fringe, Ottawa : Land Resource Research Institute, 1979.

Dumanski J., Marshall I. B., Huffman E. C., Physical land characteristics and land capability in the Ottawa urban fringe, in Wesche R., & Kugler-Gagnon M., (eds.) Ottawa Hull : Spatial perspectives and planning, Occassional papers, Dept. Geography & Regional Planning, Univ. of Ottawa, 1978.

Farley, A. L., Atlas of British Columbia -- people, environment and resource use, Vancouver : The University of British Columbia, 1979

Farstad L., Soil Survey of the Peace River area in British Columbia, Vancouver : The University of British Columbia, 1965

Goodchild, M. F., The determinants of land capability, Occassional paper 7, Environment Canada, 1976

Iversen G. R., Norpoth H., Analysis of Variance, Beverly Hills : Sage publications, Inc., California, 1976.

Kim J. O., Kohout F. J., Analysis of Variance and Covariance : Subprogram ANOVA and ONEWAY, in SPSS, New York : McGraw-Hill Book Company, 1975

Lipschtz S., Linear algebra, New York : McGraw-Hill Book Company, 1968.

Lok S. C., The use of numerical analysis techniques in land classification for resource management : the case of soil soitibility for agriculture, Ottawa : University of Ottawa, M. A. Thesis, 1979

Lok S. C., Phipps M., Calculating indices of land suitability for agricultural crops : an interactive numberical technique, in Journal of Soil and Water Conservation, 1981

Neimanis V., Mckechnie R., Urban Growth, Infrastructure and Land Capability : a Windsor example Ottawa : Environment Canada, 1980

Nix, H. A., The assessment of Biological productivity, in Stewart, G. A. Land Evaluation, Macmillan of Australia, 1968

Reichelderfer F. W., et. al. Climate and Man, Yearbook of Agriculture, Wachington D. C. : United States Department of Agriculture, 1941.

Riquier J., Approach to Land Classification, Rome : F. A. O. Soil Bulletin 22, 1974.

Robinson A., Sale R., Morrison J., Elements of Cartography, 4th edition, New York : John Wiley & Sons, 1978.

Royen W. V. The agricultural resources of the world, Atlas of the world's resources Vol. I, New York : Prentice-Hall Inc., 1954.

Small E., Dumanski J., Lendvay-Zwickl J., Climate and Soil Requirements for Economically Important Crop in Canada, Ottawa : Research Branch, Agriculture Canada, 1981.

Stewart G. A., Land Evaluation, Canberra : the Macmillan Company of Australia, 1968.

Schmidt J. M., Understanding and Using Statistics, Basic Concepts, 2nd edition, Lexington : D. C. Heath and Company, Massachusetts, 1979.

Schmude K. O., A perspective on prime farmland in Journal of Soil and Water Conservation, 1977.

Vink A. P. A., Land Use in advancing Agriculture, New York : Springer-Verley, 1975.

Walpole R. E., Statistics, New York : Macmillan Company, 1971.

Yantis R. P., Painter R. J., Elementary Matrix Algebra with Applications, 2nd edition, Boston : Prindle, Weber & Schmidt, Inc., Massachusetts, 1977.

Appendix A

COUSMAT PROGRAM

```
PROGRAM COUSMAT(INPUT,OUTPUT);
(* program of generating the Use matrix in the LASTAN model *)

CONST
  N = 15; (* number of total states *)
  A =  3; (* number of states in factor slope class *)
  B =  5; (* number of states in factor surface texture *)
  C =  4; (* number of states in factor soil subgroup *)
  D =  3; (* number of states in factor climate area *)
  E = 47;
  F = 12;
  X = 83; (* number of total interactions *)
  SLOPWT = 0.0289; (* weight of factor slope class *)
  TXTRWT = 0.0100; (* weight of factor surface texture *)
  SBGPWT = 0.0841; (* weight of factor soil subgroup *)
  CLIMWT = 0.0625; (* weight of factor climate area *)
  STWT = 0.039; (* weight of interactions --- SLOP & TXTR *)
  SSWT = 0.099; (* weight of interactions --- SLOP & SBGP *)
  TSWT = 0.104; (* weight of interactions --- TXTR & SBGP *)
  SCWT = 0.088; (* weight of interactions --- SLOP & CLIM *)
  GCWT = 0.129; (* weight of interactions --- SBGP & CLIM *)
  TCWT = 0.074; (* weight of interactions --- TXTR & CLIM *)
  GRANMEAN = 46.11; (* overall mean yield of barley *)

TYPE
  AAA = ARRAY (1..A) OF REAL;
  BBB = ARRAY (1..B) OF REAL;
  CCC = ARRAY (1..C) OF REAL;
  DDD = ARRAY (1..D) OF REAL;
  RAYONE = ARRAY (1..N) OF REAL;
  RAYTWO = ARRAY (1..X) OF REAL;
  SPT = ARRAY (1..A,1..B) OF REAL;
  SPS = ARRAY (1..A,1..C) OF REAL;
  TRS = ARRAY (1..B,1..C) OF REAL;
  SPC = ARRAY (1..A,1..D) OF REAL;
  TRC = ARRAY (1..B,1..D) OF REAL;
  SGC = ARRAY (1..C,1..D) OF REAL;
  ALPHA = ARRAY (1..X,1..N) OF INTEGER;
  BETA  = ARRAY (1..N,1..N) OF REAL;
  GAMA  = ARRAY (1..N,1..X) OF REAL;
```

```
VAR
  SLOP : AAA;
  TXTR : BBB;
  SBGP : CCC;
  CLIM : DDD;
  SLOPTXTR : SPT;
  SLOPSBGP : SPS;
  TXTRSBGP : TRS;
  SLOPCLIM : SPC;
  SBGPCLIM : SGC;
  TXTRCLIM : TRC;
  SMAT : ALPHA;
  CUMAT, UMAT : BETA;
  RMAT, MMAT, MINV : GAMA;
  DU : RAYONE;
  DR, R, V : RAYTWO;
  I, J, K, L, M : INTEGER;
  SUM : REAL;


**********************************************************************
*                                                                    *
* SMAT  : the hypothetical site matrix                               *
* UMAT  : the hypothetical use matrix                                *
* CUMAT : the computed use matrix                                    *
* RMAT  : the result matrix of the LASTAN model                      *
* MMAT  : the mediate matrix for computing interactions              *
* MINV  : inverse matrix of MMAT                                     *
* DU    : diagonal array of the use matrix or factor ratings         *
* SSU   : array computed by diagonal of SMAT & UMAT                  *
* R     : the mediate array for computing interaction matrix         *
* DR    : diagonal of result array or performance indices            *
* V     : the computed interaction array                             *
*                                                                    *
* The operations of the LASTAN model:                                *
*                                                                    *
*                            SMAT * UMAT = IMAT                      *
*                            SMAT * IMAT' = RMAT                     *
*                                                                    *
* The method of computing interaction array:                         *
*                                                                    *
*                            MMAT * V = DR - SSU = R                 *
*                            ====>  V = MINV * R                     *
*                                                                    *
**********************************************************************

(* procedure of calculating multiplication of two matrixes *)

PROCEDURE MULTIMATRIX(VAR  RES:RAYTWO;  ORI:GAMA;  MUL:RAYTWO);

VAR
  I, J, K : INTEGER;
  SUM : REAL;
```

```
BEGIN
  FOR I := 1 TO X DO
    BEGIN
      SUM := 0.0;
      FOR J := 1 TO X DO
        SUM := SUM + ORI(I,J) * MUL(J);
      RES(I) := SUM;
    END;(* for i := 1 to n *)
END;(* procedure multimatrix *)

(* procedure of computing inverse matrix *)

PROCEDURE FORINV(VAR  MATS:GAMA;  SMAT:GAMA);

CONST
  DN = 166;

TYPE
  DBARRAY = ARRAY (1..X,1..DN) OF REAL;

VAR
  CMAT : DBARRAY;
  IM, IP1, IST, I, J : INTEGER;
  CL : REAL;

(* procedure of establishing intermedia matrix *)

PROCEDURE DOUBLEREAD(VAR  ZMAT:DBARRAY;  TMAT:GAMA);

VAR
  I, J : INTEGER;

BEGIN
  FOR I := 1 TO X DO
    BEGIN
      FOR J := 1 TO DN DO
        BEGIN
          IF (J<= X) THEN
            ZMAT(I,J) := 1.0
          ELSE
            ZMAT(I,J) := 0.0;
        END;(* for j := 1 to dn *)
    END;(* for i := 1 to x *)
END;(* procedure doubleread *)

(* main program of forinv *)

BEGIN
  DOUBLEREAD(CMAT,SMAT);
  FOR IP := 1 TO X DO
    BEGIN
      IM := IP;
```

```
        IST := IP + 1;
        FOR I := IST TO X DO
          BEGIN
            IF (ABS(CMAT(IM,IP)) <ABS(CMAT(I,IP))) THEN
              IM := I;
          END;(* for i := ist to x *)
        IF (IM <> IP) THEN
          FOR J := IP TO DN DO
            BEGIN
              CL := CMAT(IP,J);
              CMAT(IP,J) := CMAT(IM,J);
              CMAT(IM,J) := CL;
            END;(* for i := ip to dn *)
        CL := CMAT(IP,IP);
        CMAT(IP,IP) := 1.0;
        FOR J := IST TO DN DO
          CMAT(IP,J) := CMAT(IP,J) / CL;
        FOR I := 1 TO X DO
          IF (I <> IP) THEN
            BEGIN
              IP1 := IP + 1;
              CL := CMAT(I,IP);
              CMAT(I,IP) := 0.0;
              FOR J := IP1 TO DN DO
                CMAT(I,J) := CMAT(I,J) - CL * CMAT(IP,J);
            END (* if i <> ip *)
      END;(* for ip := 1 to x *)
    FOR I := 1 TO X DO
      BEGIN
        FOR J := 1 TO X DO
          MATS(I,J) := CMAT(I,J+X);
      END;(* for i := 1 to x *)
  END;(* procedure forinv *)


  (* main program of COUSMAT *)

  BEGIN
    (* read in site matrix *)
    FOR I := 1 TO X DO
      BEGIN
        FOR J := 1 TO N DO
          READ(SMAT(I,J));
        READLN;
      END;(* for i := 1 to x *)
    (* read in data to compute factor ratings *)
    FOR I := 1 TO A DO
      BEGIN
        READ(SLOP(I));
        DU(I) := SLOP(I) * SLOPWT;
      END;
    READLN;
```

```
FOR I := 1 TO B DO
  BEGIN
    READ(TXTR(I));
    DW(A+I) := TXTR(I) * TXTRWT;
  END;
READLN;
FOR I := 1 TO C DO
  BEGIN
    READ(SBGP(I));
    DW(A+B+I) := SBGP(I) * SBGPWT;
  END;
READLN;
FOR I := 1 TO D DO
  BEGIN
    READ(CLIM(I));
    DW(A+B+C+I) := CLIM(I) * CLIMWT;
  END;
READLN;
(* read in data to compute performance indices *)
K := 1;
FOR I := 1 TO A DO
  BEGIN
    FOR J := 1 TO B DO
      BEGIN
        READ(SLOPTXTR(I,J));
        IF (SLOPTXTR(I,J) = 0) THEN
          SLOPTXTR(I,J) := GRANMEAN;
        DR(K) := SLOPTXTR(I,J) * STWT;
        K := K + 1;
      END;
    READLN;
  END;(* sloptxtr *)
FOR I := 1 TO A DO
  BEGIN
    FOR J := 1 TO C DO
      BEGIN
        READ(SLOPSBGP(I,J));
        IF (SLOPSBGP(I,J) = 0) THEN
          SLOPSBGP(I,J) := GRANMEAN;
        DR(K) := SLOPSBGP(I,J) * SSWT;
        K := K + 1;
      END;
    READLN;
  END;(* slopsbgp *)
FOR I := 1 TO A DO
  BEGIN
    FOR J := 1 TO D DO
      BEGIN
        READ(SLOPCLIM(I,J));
        IF (SLOPCLIM(I,J) = 0) THEN
          SLOPCLIM(I,J) := GRANMEAN;
        DR(K) := SLOPCLIM(I,J) * SCWT;
        K := K + 1;
      END;
```

```
      READLN;
    END;(* slopclim *)
  FOR I := 1 TO B DO
    BEGIN
      FOR J := 1 TO C DO
        BEGIN
          READ(TXTRSBGP(I,J));
          IF (TXTRSBGP(I,J) = 0) THEN
            TXTRSBGP(I,J) := GRANDMEAN:
          DR(K) := TXTRSBGP(I,J) * TSNT;
          K := K + 1;
        END;
      READLN;
    END;(* TXTRSBGP *)
  FOR I := 1 TO B DO
    BEGIN
      FOR J := 1 TO D DO
        BEGIN
          READ(TXTRCLIM(I,J));
          IF (TXTRCLIM(I,J) = 0) THEN
            TXTRCLIM(I,J) := GRANDMEAN:
          DR(K) := TXTRCLIM(I,J) * TCNT;
          K := K + 1;
        END;
      READLN;
    END;(* txtrclim *)
  FOR I := 1 TO C DO
    BEGIN
      FOR J := 1 TO D DO
        BEGIN
          READ(SBGPCLIM(I,J));
          IF (SBGPCLIM(I,J) = 0) THEN
            SBGPCLIM(I,J) := GRANDMEAN:
          DR(K) := SBGPCLIM(I,J) * SCNT;
          K := K + 1;
        END;
      READLN;
    END;(*sbgpclim*)

  (* computing mediate array for calculating interaction *)

  FOR I := 1 TO K DO
    BEGIN
      SUM := 0.0;
      FOR J := 1 TO N DO
        SUM := SUM + SMAT(I,J) * SMAT(I,J) * DU(J);
      R(I) := DR(I) - SUM;
    END;
```

```
(* computing mediate matrix and its inverse matrix *)

FOR I := 1 TO X DO
  BEGIN
    L := 1;
    FOR K := (A+1) TO N DO
      FOR J := 1 TO A DO
        BEGIN
          MMAT(I,L) := SMAT(I,J) * SMAT(I,K);
          L := L + 1;
        END;
    FOR K := (A+B+1) TO N DO
      FOR J := (A+1) TO (A+B) DO
        BEGIN
          MMAT(I,L) := SMAT(I,J) * SMAT(I,K);
          L := L + 1;
        END;
    FOR K := (A+B+C+1) TO N DO
      FOR J := (A+B+1) TO (A+B+C) DO
        BEGIN
          MMAT(I,J) := SMAT(I,J) * SMAT(I,K);
          L := L + 1;
        END;
  END;(* for i := 1 to x *)
FORINV(MINV,MMAT);

(* printout the computed use matrix *)

WRITELN;
WRITE(' THE COMPUTED USE MATRIX FOR BARLEY BY FOUR FACTORS ');
WRITELN;
WRITE(' ':6,'SLOPE CLASS',' ':20,'SURFACE TEXTURE');
WRITELN(' ':20,'SOIL SUB-GROUP',' ':6,'CLIMATE AREA');
WRITELN;
WRITE('0-2%':8,'2-5%':3,'5-9%':8,'LOAM':3,'S-LOAM':7);
WRITE(' CLAYLOAM SILTLOAM L-SAND','CL-D':8,'LG-O':8);
WRITELN('BL-E':8,'GL-O':3,'2CF':8,'3CF':3,'5CF':8);
WRITELN;

MULTIMATRIX(V,MINV,R);
K := 1;
L := 1;
M := 1;
FOR I := 1 TO N DO
  BEGIN
    FOR J := 1 TO N DO
      BEGIN
        IF I=J THEN
          CUMAT(I,J) := DU(I)
        ELSE
          BEGIN
```

```
            IF (I>A) AND (J<=A) THEN
              BEGIN
                CUMAT(I,J) := V(K);
                K := K + 1;
              END
            ELSE
              CUMAT(I,J) := 0.0;
          END;
        IF ((I>(A+B)) AND (J>A) AND (J<=(A+B))) THEN
          BEGIN
            CUMAT(I,J) := V(K-B+L);
            L := L + 1;
          END;
        IF ((I>(A+B+C)) AND (J>(A+B)) AND (J<=(A+B+C))) THEN
          BEGIN
            CUMAT(I,J) := V(K-B+M);
            M := M + 1;
          END;
        WRITE(CUMAT(I,J):8:4);
      END;(* for j := 1 to n *)
    WRITELN;
  END (* for i := 1 to n *)
END.
```