Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 503601
UNCLASSIFIED	

TITLE

THE SOCIO-ECONOMIC ASSESSMENT OF MILITARY INSTALLATIONS USING AN INTEGER PROGRAMMING MODEL

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

DEPARTMENT OF NATIONAL DEFENCE
CANADA

OPERATIONAL RESEARCH AND ANALYSIS

DIRECTORATE OF SOCIAL AND ECONOMIC ANALYSIS

RESEARCH NOTE 2/94

THE SOCIO-ECONOMIC ASSESSMENT OF MILITARY INSTALLATIONS USING AN INTEGER PROGRAMMING MODEL

by

B. SOLOMON

MARCH, 1994

OTTAWA, CANADA

National Defence Défense nationale

Operational Research and Analysis

CATEGORIES OF PUBLICATION

ORA Reports are the most authoritative and most carefully considered publications issued. They normally embody the results of major research activities or are significant works of lasting value or provide a comprehensive view on major defence research initiatives. ORA Reports are approved by DGOR and are subject to peer review.

ORA Project Reports record the analysis and results of studies conducted for specific sponsors. This category is the main vehicle to report completed research to the sponsors and may also describe a significant milestone in ongoing work. They are approved by DGOR and are subject to peer review. They are released initially to sponsors and may, with sponsor approval, be released to other agencies having an interest in the material.

Directorate Research Notes are issued by directorates. They are intended to outline, develop or document proposals, ideas, analysis or models which do not warrant more formal publication. They may record development work done in support of sponsored projects which could be applied elsewhere in the future. As such they help serve as the corporate scientific memory of the directorate.

DEPARTMENT OF NATIONAL DEFENCE

CANADA

OPERATIONAL RESEARCH AND ANALYSIS

DIRECTORATE OF SOCIAL AND ECONOMIC ANALYSIS
RESEARCH NOTE 2/94

THE SOCIO-ECONOMIC ASSESSMENT OF MILITARY INSTALLATIONS USING AN INTEGER PROGRAMMING MODEL

by
B. SOLOMON

OTTAWA, CANADA **MARCH, 1994**

ABSTRACT /RÉSUMÉ

This paper proposes an integer programming model to assess Canadian military installations by selected socio-economic attributes (variables).

RÉSUMÉ

Cet article présente un modèle de programme entier utilisé pour l'évaluation des installations militaires au Canada par variables socio-économiques.

TABLE OF CONTENTS

INTRODUCTION

1. This research note is the second in a series on the use of cluster analysis in economic studies. The first paper introduced the theoretical foundations of cluster analysis along with some of the drawbacks associated with the techniques used to construct objective grouping of relevant information (Solomon and Fegan, 1993). In this paper, cluster analysis is employed to examine formally its use in grouping military installations by selected socio-economic attributes.

2. The recent and dramatic changes in the geostrategic environment have triggered a series of downward revisions in defence budgets in Canada and elsewhere. Furthermore, a large federal deficit, coupled with a sluggish economic recovery, have contributed to cutbacks in Canadian military spending. The recent announcement of base closures and reductions is part of the spending cutbacks, and economic analysis can be used in a number of cases to determine the cost-effectiveness of various defence decisions.

3. Recently DSEA completed a study on the socio-economic impacts of 43 military installations on their host communities (Wait, et al., 1992). While a number of studies had been conducted in preceding years on the impact of military installations on the surrounding communities, the Wait et al. (1992) study is the only one that provides a comprehensive analysis of all the major bases in Canada[1]. Because of the scope of the study, the report concentrated on readily quantifiable economic variables such as demographic, income and labour impacts.

4. While it is relatively safe (empirically) to conclude that installations located in major urban centres such as Toronto and Montreal have small impacts on the community, it is difficult to classify the impacts of other installations without some

---

[1] Wait and Parai (1992) provides a list of DSEA (DND) publications on socio-economic impact of military installations.

form of analytical method that uses subjective and objective criteria. In this paper clustering methods (based on *integer programming*) are used to classify a selected number of bases by their socio-economic impact ratios (estimated by the DSEA study mentioned above), to establish a baseline for comparison and to identify and understand the socio-economic impacts. It is anticipated that the theoretical framework developed in this study can also be applied to military operational or other decision variables.

5. The clustering method is formulated as an integer programming problem in order to take advantage of specialized discrete optimization algorithms. For example, adding a simple constraint can limit the number of objects in a cluster or prevent some objects from representing a median cluster. Integer programming is a branch of a general field of *mathematical programming* that deals with the development of modelling and solution procedures for the purpose of maximizing the extent to which the objectives of the decision maker are realized. In integer programming, the decision variables are constrained to be integer. Given most real-world problems require whole number solution, such restrictions are justifiable. Mulvey and Crowder (1979) initially introduced integer programming formulation to cluster analysis and Ng's (1994) subsequent application shows that such formulations are a viable alternative to other well known clustering algorithms.

6. It should also be pointed out that the Minister's Advisory Group on Defence Infrastructure showed some interest in categorizing installations when it proposed that bases be classified by their socio-economic impact ratios, using terms such as "low", "moderate" and "severe". Parai (1992) attempted to quantify the categories provided by arbitrarily choosing two percent as the cut-off point between the "low" and "moderate" categories. While the two percent cut-off was shown to be robust, by using three of the four impact measures, the author acknowledged that the study should be considered a preliminary analysis and that more information is required on how and why these categories are to be used (Parai, 1992).

7. The reminder of this note is divided as follows: Section I presents and discusses the socio-economic attributes, while in Section II, the clustering model is formally presented. The cluster subroutines are based on an optimization approach to allow the flexible and robust formulation of the clustering problem. In Section III, the cluster results are presented and Section IV concludes the study.

I. ECONOMIC ATTRIBUTES

8. To analyse the economic impacts of an expenditure either by a base[2] or any other economic agent on its host community, a detailed account of the community's tradeflow (what is produced locally and what is imported), industrial production and other macroeconomic variables such as government expenditures have to be taken into account. Ideally, a sub-provincial input-output (I-O) model should be used to analyze such activities. However, these models are difficult to build and Statistics Canada provides only provincial level I-O models.

9. In the absence of such a model, one may modify the I-O methodology to capture the **direct** and first-round **indirect** impacts. The direct impact (i.e. production and revenue) can be seen as the income generated by the base expenditure on goods and services and the employment income of the base employees. On the other hand, activities by industries that provide machinery and services to the base represent the indirect impact (Wait and Parai, 1992). Other impacts, such as the spending amount remaining in the host community as a result of the base O&M expenditure and the wage and salaries spent in the host community by the base employees, can also be considered indirect impacts.

[2] In this study the terms military installations and bases are used interchangeably.

10. The underlying algebraic presentation of the income impacts is as follows:

$$AE_{hc} = AE * LPR$$

$$AE^{i}_{hc} = AE_{hc} * CE_{f} * LPR_{f}$$

$$CE_{hc} = CE_{w} * CE_{f} * LPR$$

$$CE^{i}_{hc} = CE_{hc} * CE_{f} * LPR$$

Where AE represents the aggregate expenditure (base O&M expenditure in this case), AE^{i} and AE_{hc} represent the first round indirect (i) and host community (hc) level aggregate expenditure, CE represents the wages and salaries (w) expenditure, CE_{f} represents the family expenditure categories as defined in the SC family (f) expenditure statistics, LPR and LPR_{f} denote the local area production coefficient of the specific commodity based on interprovincial trade statistics. The Wait-Parai methodology uses the availability of data on expenditure by postal code area. Such information, coupled with interprovincial trade statistics, clearly provides a reliable estimate of what remains in the community. All host community data were based on the 1986 (the most recent at the time) census statistics, while base expenditures and other financial statistics were from DND sources (Wait, 1992) for fiscal year 1990-91.

11. The demographic impacts included in the Wait-Parai study consist of the proportion of military personnel and their family members to the total population of the community. Other quantifiable social impacts such as education (the proportion of military personnel children in the community school enrolment), housing (the impact of military personnel on the demand for housing in the community) and grants in lieu of taxes (the amount bases pay to the municipalities as a proportion of the total tax base) were also included in the study (Wait, et al., 1992). While there are other important variables that have to be considered, such as the military and strategic importance of the base, these were deemed to be beyond the scope of the study. Furthermore, the involvement of base personnel in community activities and the policing and emergency service provided by the military personnel were not analyzed

since these are difficult to quantify and required extensive collection of data (Wait, et al., 1992).

Impact Ratios as Economic Attributes

12. Within the constraints of the Wait-Parai study, the impact ratios are treated as the economic attributes for subsequent clustering processes. It should be pointed out that the clustering methodology employed in subsequent sections can be modified to include other attributes depending on the scope of similar base impact studies. The estimated impact ratios can be grouped into four main economic attributes namely, **demographic, labour, income** and **socio-economic.**

13. Within the demographic impacts reside the proportion of military personnel to the total population and the proportion of military family members to the host community population (see Figure 1). Similarly, the labour impact ratio consists of civilian and military employment.

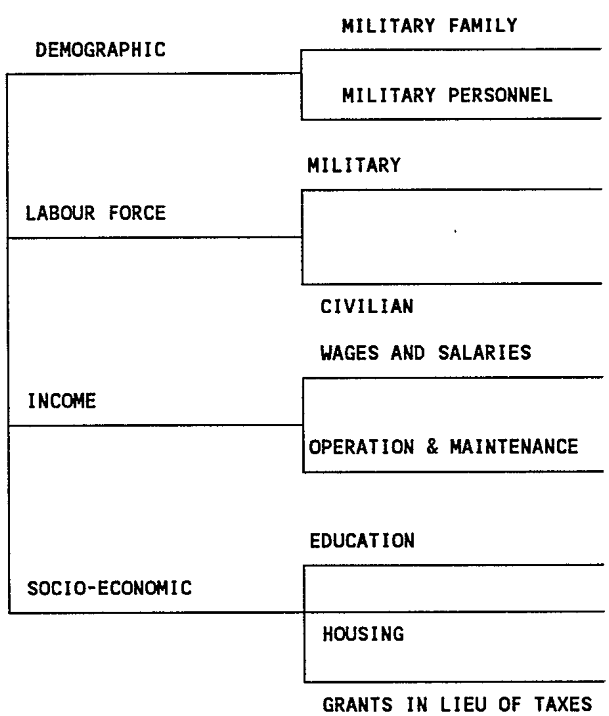

Figure 1 Economic Attributes

While civilian employment directly affects the community's employment and unemployment figures, the military employment may be transferable and subsequently may affect the unemployment rate quite differently than the civilian figures.[3]

14. The income impacts are represented by the O&M expenditures and Wages & Salaries (as proportion of community income) ratios. As pointed out earlier, the direct O&M expenditures that remain in the community are an important source of income to various medium and small businesses, while expenditures from net wages helps the local retail and service industries. Of the four main economic attributes in the study, the income impact is largely based on the algebraic representation shown on page 3.

TABLE I
ECONOMIC ATTRIBUTES

ECONOMIC ATTRIBUTE	DESCRIPTION
E1. Military Family	Military spouse and children as a percentage of the community population.
E2. Military Personnel	Military personnel (officers and NCMs) as a percentage of the community population.
E3. Military Labour Force	Military personnel (officers and NCMs) as a percentage of the community labour force.
E4. Civilian Labour Force	Civilian personnel (DND and NPF*) as a percentage of the community labour force.

[3] It should be pointed out that the Census statistics include military employees as part of the community's labour force. While the summary section of the Wait-Parai (1992) study aggregates the impact ratios for presentation purposes, the Wait, et al. (1992) study presents the impact ratios in a disaggegated form. Furthermore, military personnel and family are not split in the base study.

ECONOMIC ATTRIBUTE	DESCRIPTION
E5. Wages & Salaries	Total estimated direct and indirect spending of net wages (Military, DND and NPF personnel) as a percentage of host community income.
E6. O&M Expenditures	Total estimated direct and indirect expenditures (including NPF) as a percentage of host community income.
E7. Education	DND students attending primary and secondary schools in the community as a percentage of local school population.
E8. Housing	Demand for local housing (after accounting for on-base housing facilities) as a percentage of local housing.
E9. G.I.L.T. (Grants in lieu of taxes)	Grants paid by the base to the municipalities as a percentage of the local tax base.

Source: Wait et al. (1992) It should be pointed out that the economic attributes are not used separately in the project report.
* Non-Public Funds

15. The fourth and last attribute is an amalgam of quantifiable socio-economic impacts such as housing, education and grants-in-lieu of taxes. The housing ratio provides a proxy measure on the demand for local housing and it is derived by subtracting the occupied base housing facilities (married and single quarters typically, unless training is a major function of the base) from the number of military personnel on the base (the difference in personnel represents the demand for housing).

II. THE MODEL[4]

16. An earlier research note on cluster analysis introduced and discussed a number of clustering algorithms (Solomon and Fegan, 1993). These algorithms differ from each other in the way they measure the similarity between two objects and in the way they use the measure of similarity to form clusters. Although clustering algorithms, such as *iterative partitioning* and *hierarchical agglomerative*, remain popular, there has been increasing interest in *integer linear programming*. The method was first used in a clustering problem by Mulvey and Crowder (1979), and the authors concluded that integer programming can be considered a viable clustering technique. In this note the clustering of military installations is formulated as an integer programming (similar to Ng, 1994).

17. The measurement of similarity between every pair of military installations is based on the values of the four main economic attributes (i.e. demographic, labour, income and socio-economic). These attributes or evaluation criteria are further subdivided into a decision-tree type sub levels as depicted in Figure 1. While the clustering algorithms mentioned above can be applied to the attributes to form groups, it is often desirable to weigh some attributes more than others if the researcher has a priori information based on some theory or "expert judgement". It is always necessary to include and measure all important tangible and intangible factors in socio-economic models where human behaviour has consequential impact but these factors are often left out when one makes simplifying assumptions.

18. In weighting attributes, however, one has to work with independent attributes. More formally, in ascribing a preference between attributes, the set of attributes is dependent if the measurement of numerical scores (objective or subjective) with respect to one attribute implies or restricts a particular attainment of scores by all

---

[4] The model presented below is based on K. Ng (1994).

other attributes of the set. For example *cost* and *price* or *military family* and *military personnel* can be considered dependent attributes. In the next section the number of attributes is reduced to four independent subcriteria using the principal components method.

19. After the identification of the subcriteria is completed, the base's mean and standard deviation of each subcriterion are computed. These computed values are then used to "normalize" each of the base's attribute in the following manner:

$$\frac{E_j-\mu_E}{\sigma_E}$$

where μ_E and σ_E represents the attribute's mean and standard deviation and E_j represents the attribute's actual value.

20. Suppose for base i, e_{ik} denotes the normalized value of the base's k^{th} evaluation criteria (k = 1,2,...,M), then,

$$d_{ij}=\left|\sum_{k=1}^{M} w_k(e_{ik}-e_{jk})\right| \tag{1}$$

where w_k is the weighting factor corresponding to attribute k. The measure of similarity in this study (i.e d_{ij}) is the weighted absolute value distance between base i and base j. Depending on the type of data, any other distance measure such as the Euclidean distance may also be used (Ng, 1994).

The Data

21. As mentioned earlier, the impact ratios representing the subcriteria were extracted from the Wait et al. (1992) study. Although 43 military installations were examined in the study, only 10 bases are considered here for illustrative purposes. The 10 bases were selected randomly, with one from each province and territory (PEI is not included as there is no base there).

22. The subcriteria identified in Table I have to be weighted to reflect the relevant importance of a given impact ratio with respect to other ratios. Since the nine attributes shown in Table I are dependent, the subcriteria are reduced using information derived from a correlation matrix and the principal component method. The four dependent attributes are shown in Figure 2 and Table II.

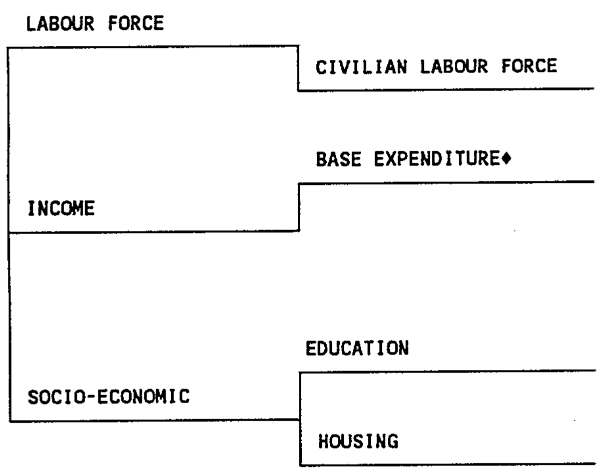

Figure 2 Independent Attributes

♦ Linear combination of Wages & salaries and O&M expenditures ratios

23. Since the Wait et. al (1992) study used demographic impact ratios primarily to estimate education and housing demand, this attribute is captured through the housing and education variables. Furthermore the impact ratios for labour and income attributes show a high level of correlation and thus civilian labour force and O&M ratios or a linear combination of the income impact and civilian labour force were selected via principal component analysis. Weights can now be assigned either by incorporating one's own beliefs or "expert opinion" on the importance of one subcriteria from another, or by using statistical methods such as regression or factor analysis. Once the relative weights of the criteria have been decided, a more systematic approach to assign the weighting factors can be provided through a pairwise evaluation of the relative importance of the estimated impact ratios. This method is employed in this study (Saaty, 1979). Thus if one supposes that subcriterion (impact ratio) i is more important than subcriterion j, the e_{ij} is assigned a number that reflect the intensity of importance.

24. A scale of importance of 1 for example indicates equal importance, while a 3 shows weak importance of one factor over another etc.[5] In this paper we assume that consultation with community officials provided the following preferences on the three main economic attributes, labour (l), income (i) and socio-economic (se) :

l is more important than se and the intensity is 3,
i is more important than l and the intensity is 2,
i is more important than se and the intensity is 4,

These ordinal numbers are then recorded in a matrix to compute the eigenvector associated with the maximum eigenvalue of the matrix. The values of the eigenvector

[5] Note that if j is more important than i, e_{ji} is assigned the value. Alternatively, either e_{ij} or e_{ji} is derived first and the other is obtained through the relationship $e_{ji} = 1/e_{ij}$. The nine scale preference indicator is discussed in Saaty (1979).

represent the intensities of the judgements as they are recorded. The decision hierarchy setup and the eigenvalue calculations are performed using Expert Choice (1986) software.

25. The weights from the eigenvalue calculations are multiplied together to represent the "effective weight" of each attribute. In general, if the effective weight of the k^{th} attribute is u_k and if the adjustment for the corresponding impact ratio (subcriterion) is v_k, the "adjusted effective weight " w_k , becomes:

$$w_k = u_k v_k / (u_1 v_1 + u_2 v_2 + \ldots + u_M v_M) \text{ and } M = \# \text{ of attributes (Ng, 1994).}$$

The data along with the "adjusted effective weight" are presented in Table II.

TABLE II

IMPACT RATIOS FOR SELECTED CANADIAN MILITARY INSTALLATIONS (%)

NAME	CLFORCE	BASE EXPENDITURE	EDUCATION	HOUSING
GOOSE BAY	26.0	35.9	17.5	0
HALIFAX	3.3	7.1	31.3	5
MONCTON	1.2	0.93	1.1	0.1
BAGOTVILLE	0.6	3.2	2.6	1.3
PETAWAWA	12.6	68.3	1.6	11.9
SHILO	3.9	6.1	1.3	0.1
MOOSE JAW	2.7	7.0	8.1	5.4
EDMONTON	0.2	0.68	0.8	0.2
NANAIMO	0.2	0.6	0.2	0.3
YELLOWKNIFE	0.2	1.5	1.7	0.5
WEIGHTS	0.320	0.558	0.061	0.061

Source: Derived from Wait et al. (1992)
The weights are based on the eigenvalue calculations.

The Integer Program Formulation

26. All integer models, as the name implies, require some or all of their decision variables to be integers. Such a restriction is desirable for most real-world problems that require whole number solutions (one cannot buy 1.25 cars, etc.). Furthermore, in "0-1 integer programming", adding a simple constraint can limit the number of objects in a cluster or prevent some objects from representing a median cluster. In this study a "0-1 integer" (i.e. variables assuming values of 0 or 1) model is used. Thus, we define the following binary variables:

Y_i $i = 1,2,\ldots,N$ (N is the number of bases) where,

Y_i = 1 denotes that base i is a cluster median,

0 denotes that base i is not a cluster median.

X_{ij} , $1 \leq i \leq N$, $1 \leq j \leq N$, $i \neq j$, where

X_{ij} = 1 denotes that base i is assigned to a cluster median at base j,

0 denotes that base i is not assigned to a cluster median at base j.

27. The program is expected to assess first the bases (Y_i's equal to 1) that serve as cluster medians each of which will be a focal point for a different cluster. Second, it will assess those bases forming clusters by assigning the remaining bases to one of the cluster medians (X_{ij}'s that are non-zero). The objective of the problem then becomes minimizing the sum of, over all the cluster medians, the distance from each non-median base to the cluster median to which it is assigned (Ng, 1994). Symbolically, the model is presented as:

$$\min \sum_{\substack{i,j \\ i\neq j}}^{N} d_{ij} x_{ij}$$

subject to

$$\sum_{\substack{i \\ i\neq j}}^{N} x_{ij} \leq (N-1) y_j \qquad \forall j \tag{2}$$

$$\sum_{\substack{i \\ i\neq j}}^{N} x_{ij} \geq y_j \qquad \forall j \tag{3}$$

$$y_i + \sum_{\substack{j \\ i\neq j}}^{N} x_{ij} = 1 \qquad \forall i \tag{4}$$

$$A \leq \sum_{j=1}^{N} y_j \leq B \tag{5}$$

where A and B denote the lower and upper bound on the number of clusters to be formed.

28. Equation 2 (or constraint 2) restricts the assignment of bases to base j unless base j is a cluster median. Equation 3 states that a cluster median ($Y_j = 1$) is assigned at least one other base while equation 4 restricts base i as either a cluster median or it is assigned to a cluster median. The upper and lower bounds of the number of clusters formed is expressed by equation 5.

III. THE CLUSTER RESULTS

29. The integer programming problem is formulated using the Lindo (1992) software with the branch and bound method option. The method calculates the optimal solution by examining only a portion of all possible combinations of integer values. The procedure basically divides the set of all solutions into a smaller group and examines only "promising" subgroups of solutions. The method is quite powerful for a small problem set such as the one examined here.

30. The first simulation is based on the constraint that the number of clusters should not be smaller than two or greater than three (i.e. the values for A and B in equation 5 are 2 and 3 respectively). The optimal objective value for the simulation is 2.338 and the solution corresponds to the following clusters:

{Goose-bay, Petawawa♦}
{Halifax, Shilo, Moose Jaw♦}
{Moncton, Bagotville, Edmonton♦, Nanaimo, Yellowknife}

where ♦ indicates a cluster median. A comparison of the results with the raw data provided in Table II shows that the first cluster represented by Goose-bay and Petawawa has a relatively larger impact ratio for the expenditure attribute and also exhibits similar ratios for the labour force variable than the other bases. Similarly, the bases identified in the third cluster are relatively distinct from the first cluster and the remaining bases aggregated in the second cluster.

31. The flexibility of integer programming in the restriction of the number of clusters to be formed is further illustrated when compared to the K means clustering method found in most statistical software. The K means clustering partitioned the bases into two clusters with Petawawa forming one cluster and the rest forming the other. When 3 clusters are imposed the K means again provided a trivial partitioning solution, where Petatwawa and Bagotville each formed a single cluster with the remaining bases forming a third cluster. Furthermore, the integer programming formulation permits the weighing of individual cases more readily than that is available in most statistical software.

32. Now suppose the researcher is interested in a particular case, for example, the O&M impact ratio. Because the variable is of some importance to the researcher he (or she) may specify a restriction to ensure that the bases and cluster medians have similar O&M values[6]. Symbolically,

$$\left|\frac{b_i^k - b_j^k}{b_j^k}\right| x_{ij} \leq \alpha y_j \qquad \forall i, \forall j \qquad (6)$$

$k=1,2,\ldots,K$; $K \subset M$, $0 < \alpha < 1$ (recall that M is number of attributes)

with b_i^k, b_j^k being the impact ratios of the key attributes k, and x_{ij}, y_{ij} are binary variables.

[6] It is possible that a base might be assigned to a cluster where the cluster median's key attribute ratio is quite dissimilar from that of the base's.

For the following simulation K = 1 and the allowable percentage deviation of the O&M impact ratio α = 0.50.

TABLE II

O&M EXPENDITURE IMPACT FOR SELECTED CANADIAN MILITARY INSTALLATIONS

NAME	O&M
GOOSE BAY	6.2
HALIFAX	1.6
MONCTON	0.13
BAGOTVILLE	0.6
PETAWAWA	1.5
SHILO	5.4
MOOSE JAW	0.6
EDMONTON	0.13
NANAIMO	0.3
YELLOWKNIFE	0.2
WEIGHTS	0.558

Source: Derived from Wait et al. (1992)

33. The second simulation based on the above constraints provided feasible integer solutions after respecifying the upper limit of the clusters to four. The solution corresponds to the following clusters:

{Goose-bay, Shilo♦}
{Bagotville, Moose Jaw♦}
{Halifax, Petawawa♦}
{Moncton, Edmonton, Nanaimo, Yellowknife♦}

Once again ♦ indicates a cluster median. The O&M values for Bagotville and Moose Jaw are identical and obviously satisfy the conditions specified by equation 6. Similarly, the O&M values for the remaining cluster members fall within the specified α (see Table III also).

Implications

34. In this study subjective judgements were incorporated in the weighting of the attributes to show an example of a model that can incorporate unquantifiable variables. Thus, the clustering problem can be formulated without weighting factors and the associated independence restriction. As pointed out earlier, most statistical software do not provide systematic weighting schemes and the model employed here is a viable alternative. Furthermore, the integer programming model can also be modified to include restrictions on the number of clusters that can be formed and the variables which are used as the basis for a cluster median.

35. Various exercises with different weighting factors and variables were performed to assess the sensitivity of the clusters. In situations where all the available impact ratios were examined with various weight assignments the clusters remained fairly consistent. However, as shown above, when the O&M expenditures ratio is used independent of other income ratios, only six of the ten bases formed the same clusters as before. Given the fact that the O&M variable has different distribution across bases than the other variables the results should not be considered surprising. Nevertheless, the researcher has to be aware of the data employed and how sensitive they are to variation in weight assignments.

36. The integer programming model discussed in this paper also has important implications for future decision-making scenarios. For example, the model can be used to classify existing military installations by operational readiness, availability of infrastructure and airspace, and their ability to accommodate future expansion of the CF. Alternatively, the model can also be used to compare the condition of the infrastructure between bases following specific architectural criteria.

37. As it is often the case, a successful use of a model in a particular case prompts high levels of expectation that the methodology is a panacea for all types of problems. As indicated in Solomon and Fegan (1993), clustering algorithms, like any other

statistical method, have limitations and caution should be exercised when using them. As shown earlier the use of the K means algorithm was not suitable for the base clustering problem and the integer model had to be specified. Furthermore, data can be grouped without employing clustering algorithms. For example, when faced with a classification problem one may consult a multi-variate graphical presentation of the problem to visually detect groups in the data. One such method is known as the Chernoff "face'. Since people understand facial expressions, Chernoff (1973) suggested representing p-dimensional observations as a two-dimensional face with characteristics such as face shape, curvature of the mouth, nose length, eye size, pupil position, and so forth determined by measurements of the p variables. Furthermore, one can associate the variables responsible for distinguishing clusters with prominent features, such as the face shape, nose length, etc., so as to facilitate comparisons between clusters based on these variables. For example Ng (1994) included a Chernoff face representation of the raw data to verify the consistency of the clusters formed.

CONCLUSIONS

38. In this paper, an **integer programming model** is proposed to assess Canadian military installations by selected socio-economic attributes. However, the proposed integer programming model and methodology is not restricted to this particular problem. The model can also be an important tool for other military (operational) decision making processes such as classifying existing military installations by operational readiness, by availability of infrastructure and airspace, and by their ability to accommodate future expansion of the CF if and when the ever evolving strategic environment dictates.

39. Future studies that utilize optimizing models (such as the integer program) may also include Chernoff face representations (Ng, 1994). This method provides a means to display pictorially the final groupings produced by the clustering methodology.

REFERENCES

Chernoff, H. (1973), "Using faces to represent points in K-dimensional space graphically" Journal of the American Statistical Association, Vol. 68, no. 342, pp. 361-368.

Expert Choice (1986), Decision Support Software, Inc.

LINDO (1992), 5th edition, Schrage, L., The Scientific Press.

Mulvey, J.M. and H.P. Crowder (1979), "Cluster Analysis: An Application of Lagrangian Relaxation", Management Science, Vol. 25, no. 4, pp. 329-340.

Ng, K,Y,K (1994), "Environmental Assessment of Canadian Cities: A Clustering Analysis Approach", Memo.

Parai, L. (1992) " One Word Summary of the Socio-Economic Impacts of Military Installations on their Host Communities", DSEA Memo.

Saaty, T.L. (1979), "Applications of analytical hierarchies", Mathematical Computation and Simulation, Vol. 21, pp. 1-20.

Solomon, B and S. Fegan (1993), "Cluster Analysis: Some Economic Applications", DSEA Research Note 8/93.

Wait, T and L. Parai (1992), "The Socio-Economic Impacts of Military Installations on Their Host Communities", ORAE Project Report 582, Ottawa.

Wait, T, Louis Parai, Lcdr. M.N. Rodgers and Binyam Solomon (1992), " The Socio-Economic Impacts of Military Installations on Their Host Communities: Individual Assessments", ORAE Project Report 583, Ottawa.

Wait, T. (1992), "Military Installations and Host Communities: Data, Sources and Definitions", ORAE Project Report 584, Ottawa.

DOCUMENT CONTROL DATA
(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared e.g. Establishment Sponsoring a contractor's report, or tasking agency, are entered in Section 8).

DIRECTORATE OF SOCIAL AND ECONOMIC ANALYSIS

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title)

THE SOCIO-ECONOMIC ASSESSMENT OF MILITARY INSTALLATIONS USING AN INTEGER PROGRAMMING MODEL

4. AUTHORS (last name, first name, middle initial)

SOLOMON, B.

5. DATE OF PUBLICATION (month Year of Publication of document)

MARCH, 1994

6a. NO OF PAGES (total containing information. Include Annexes, Appendices, etc.)

23

6b. NO OF REFS (total cited in document)

11

7. DESCRIPTIVE NOTES (the category of document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

DSEA RESEARCH NOTE

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address).

DEPARTMENT OF NATIONAL DEFENCE, NDHQ, OTTAWA

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

ACTIVITY 45732

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR's document number (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors: further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the test is bilingual).

This paper proposes an integer programming model to assess Canadian military installations by selected socio-economic attributes (variables).

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified . If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Distance measures
Hierarchical Agglomerative methods, Cluster Analysis
Integer programming, Mathematical programming
Hierarchical decesion matrix, Eigen values
principal components method
Economic impact analyses and attributes, host community O&M expenditures

Canada

#503601