International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:10, No:10, 2016

Hybrid Approach for Country's Performance Evaluation

C. Slim

Abstract**—**This paper presents an integrated model, which hybridized data envelopment analysis (DEA) and support vector machine (SVM) together, to class countries according to their efficiency and performance. This model takes into account aspects of multi-dimensional indicators, decision-making hierarchy and relativity of measurement. Starting from a set of indicators of performance as exhaustive as possible, a process of successive aggregations has been developed to attain an overall evaluation of a country's competitiveness.

Keywords**—**Artificial neural networks, support vector machine, data envelopment analysis, aggregations, indicators of performance.

I. Introduction

THE current world economic dynamics, which is characterized by market globalization and a liberalization of international transactions, is continuously encouraging nations to provide firms with the most favorable social, political and economic environments to face foreign competition. In this context, every country is bound to pay special attention to its competitiveness and have a well-developed place in the world markets.

A review of the literature revealed the existence of different approaches to competitiveness evaluation. In their analyses, these approaches include three major aspects, namely: the hierarchical analysis level (micro-economic level vs macro-economic level [1]-[3], the aspect of multi-dimensional analysis [4], [5], and the relativity of competitiveness measurement [6], [7].

We will focus on the development of a hybrid model based on the DEA-SVM approach of Competitiveness Analysis, which may be applied in an environment characterized by the large number of decision-makers belonging to different decision-making hierarchical levels.

DEA method product empirical area by pieces that, in economic terms, represents the production frontier of best practice revealed in [8]. Effective farms are located on the empirical efficiency frontier indicating the maximum production that can be produced with different combinations of factors for a given technology. The major drawbacks of this method are:

- Inefficiencies deducted have no statistical property;
- Measurement errors and/or omission of variables can affect the measurements of inefficiency.

C. Slim is the Head of MOCFINE Laboratory, Higher Institute of Accounting and Administration Enterprise, University of Manouba, Tunisia (phone: 0021671600705, fax: 0021671602404, e-mail: chokri.slim@ iscae.rnu.tn).

SVMs have been proposed as a novel technique. SVMs are a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution; established on the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error. SVMs are shown to be very resistant to the overtting problem, eventually achieving high generalization performance in solving various problems.

In this paper, the SVM learning algorithms are applied to the DEA networks and a SVM. Furthermore, the proposed DEA-SVM retains most of the advantages of the original DEA and SVM systems.

The DEA method and the SVM approach are briefly summarized in Section II. The hybrid DEA-SVM model is presented in Section III. Section IV illustrates the empirical analysis to evaluate the competitiveness of 22 European countries. Finally, in Section V, some conclusions are presented.

II. Background

A. DEA

DEA is to determine the efficiency benchmarks (reference units) and to place all units against these benchmarks. It proceeds by data envelopment. The units are located on the envelope (or empirical frontier) and thus constitute the reference points. A distance of the other units to this boundary is a measure of their inefficiency. More detailed information can be found elsewhere in [8] and [9].

Here, we outline the mathematical formulation of the DEA method, which is a presentation from [10].

The DEA formulation is given as follows. Given a set of n decision-making units (DMUs) to be analyzed, each uses m common inputs x_i with u_i weight and s common outputs y_j with v_j weight. Let k (k=1, 2…, n) denote the DMU whose relative efficiency or productivity H_k is to be maximized.

$$\text{Maximize } H_k = \frac{\sum_{r=1}^{s} u_{rk} y_{rk}}{\sum_{i=1}^{s} v_{ik} x_{ik}} \tag{1}$$

Subject to:

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1;\ i= 1, 2, ...,m; j= 1, 2, ..n$$

$u_r, v_i \geq \varepsilon$

B. SVM

SVM developed by [11] has gained popularity due to many attractive features and excellent generalization performance. It is one kind of new machine learning algorithm in the statistical learning theory. SVM formulation is given as: The problem addressed is that of the Binary discrimination. It is to find a way to build a decision function associating to each observation its class. We are going to deal with this problem in a probabilistic framework and ask that the forms to discriminate are vectors x $IR^{\cdot p}$. The probabilistic framework of the problem is to assume the existence of an unknown function on ($IR^{\cdot p,}$ {-1, 1}). To build this estimator we assume the existence of a sample set $\{(x_i, y_i)\}$.

A problem of discrimination is said linearly separable when there is a linear function f (also called linear separator) that satisfies:

$$f(x) = \sum_{i=1}^{n} w_i x_i + b = 0 \tag{2}$$

where w_i is a n-dimension vector and b is a scalar value. Meanwhile, each sample follows:

$$\begin{cases} w_i x_i - b \geq 1 & \text{for } x_i \text{ of the first class} \\ w_i x_i - b \leq -1 & \text{for } x_i \text{ of the second class} \end{cases} \tag{3}$$

The problem then is to find w and b such that $\frac{2}{\|w\|}$ is maximum $\forall$ (x_i, y_i). For non-linear case, the idea is to add an adjustment variables ξ_i (slack variables) in the formulation to take into account the classification errors or noise. In equivalent manner, the problem can be written more simply as the minimization of:

$$Min\ (p, b, \zeta) = \frac{1}{2} w.w + \frac{C}{2} \sum_{i=1}^{n} \zeta_i^2 \tag{4}$$

C is a constant to control the compromise between number of classification errors, and the margin width. Primal formulation can be converted into dual formulation using multipliers Lagrange αi. SVM requires solving the following optimization problem:

$$V(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i . x_j) \tag{5}$$

Subject to:

$$\sum_{i=1}^{l} \alpha_i y_i = 0, \quad C \geq \alpha \geq 0, \ i = 1, 2 \ldots l$$

The linearly separable case is somewhat interesting because classification problems are often nonlinear. To solve this classical method is to project the data in a higher dimensional space called feature space. The idea is that by increasing dimensionality of the problem we find ourselves in the linear case seen previously. We will thus apply a nonlinear transformation **Φ (.)** the input vectors xi such that xi IR and $\Phi(x_i)$ IR. This change will lead to pass a scalar product in the original space xi xj to a scalar product $\Phi(x_i).\Phi(x_j)$ in the high space. The trick is to use a kernel function denoted K avoids the explicit calculation of the scalar product in the feature space. The kernel functions must satisfy the Mercer theorem. We then have $K(x_i\ x_j) = \Phi(x_i).\Phi(x_j)$.

There are many kernel functions, the radial based functions (RBF) are used in this work.

The final decision function for a new sample has the following form:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x_j) + b \right) \tag{6}$$

where b is a threshold term computed as:

$$b = \sum_{i=1}^{n} \alpha_i y_i k(x_i, x_j) \text{ ; for any } j \quad \{1,2,\ldots n\}.$$

III. The Methodology Description

A. Evaluation Competitiveness

In this study the fields refer to: Overall Performance, Macro and market dynamism, Financial Dynamism, Infrastructure & Investment Climate and Human and Intellectual Capital. Each field covers a set of sub-fields. Thus, the Macro and Market Dynamism covers the sub-fields: Investment and Productivity Growth, Overall Trade Dimension, Export Competitiveness, Export Structure, Trade Policy and Government Involvement in the Economy. Each sub-field is characterized by a set of performance indicators (63 indicators considered in 2001). As an example, we mention the sub-field Export Structure which includes the performance indicators: Manufactured Export (% of total exports), Percent Change in Share of Manufactured Exports, High Tech. Exports (% of manufactured exports). In addition, [12] adopts a hierarchical structure which may also provide an implementation context for the model presented. In fact, competitiveness is grasped through five dimensions, namely the productive, technological, commercial, financial & monetary dimension and finally the political and institutional one. According to this conception, each dimension includes several aspects. Thus, the productive dimension integrates the aspect of innovation flexibility and restructuring old productions, the coherence effects of the productive tissue, the firms' ability to raise productivity, the firms' ability to manage the social component, the evolution of market shares and finally employment growth rate. The problem posed by the European Economic Union (EEU) consists in performing an analysis and a diagnosis of the producers' strengths and weaknesses in the European Union, the factors accounting for success and the potential improvements. To this end, a set of performance indicators are considered by EUROSTAT (Statistical Office of the European Commission) in order to create a clear picture for the achievement of such a diagnosis. The indicators consist of two classes. The first one is related to industrial competitiveness and the second is related to economic competitiveness. Each class includes a set of sub-classes. Thus, the first class consists of the sub-classes Market share, Contribution to the economy, External trade indicators, Financial indicators while the second class consists of the sub-classes Competitiveness in cost and design, Labor productivity, Productivity of capital, Competitiveness in marketing, Price competitiveness, General factors affecting a country. Each sub-class is grasped by a set of indicators. For instance, for the sub-class Competitiveness in marketing, we identify the indicators: Pricing strategy, Quality of marketing effort, Market power, Competitive delivery times, After-sales service. The hierarchical and multidimensional data structure considered by the "Global Competitiveness Report" (GCR) and the World Competitiveness Yearbook (WCY) is also about to be implemented for the suggested modeling approach. According to this structure, a country is characterized by eight fields of activities called factors, and which are incidentally: Domestic Economy, Internationalization, Government, Finance, Infrastructure, Management, Science & Technology and People. Each factor among the latter includes a set of sub-factors. For example, we mention the factor Domestic Economy which includes the sub-factors: Value added, Investments, Savings, Final consumption, Economic sectors performance, Cost of living, Adaptiveness. Each sub-factor is grasped by a set of elementary indicators. Thus, the sub-factor Value added is grasped by the indicators: Gross Domestic Product (GDP), GDP per capita. Finally, implementing DEA-SVM may be envisaged by considering a sectorial decomposition of a country's activity (manufacturing, textile, tourism, handicraft, etc.). Each sector may in turn be characterized by a set of indicators that reflect the performance level of the sector involved.

The methodology suggested by the DEA-SVM takes place according to several stages. We start determining an aggregate measure (score) of the elementary indicators by exploiting the DEA technique at the level of each homogeneous grouping of elementary indicators (first aggregation level). It's worth noting that a homogeneous grouping of elementary indicators is a set of indicators describing an identical aspect of competitiveness. In a second stage, we seek to aggregate the resulting measurements from the previous stage, using the SVM technique. A one-dimensional measurement of competitiveness is to be calculated from the multi-dimensional expression worked out in the previous stage. We expect to position the countries in a metric space and to compare them with a hypothetical country supposed to be perfectly competitive. The objective pursued being to classify these countries on a one-dimensional continuum and to maintain the degree of similarity reflected between these countries in relation to the competitiveness measurements. To set up a system for measuring the degree of similarity between countries, we will have to express a similarity index. Thus, in this procedure, we get back our methodological line of relative evaluation in relation to the best practice. Finally, our last stage consists in arranging the countries to be studied according to their competitive level.

B. Hybrid Model

In this work, a hybrid method combining DEA and SVM is proposed to evaluate countries competitiveness performance in different levels. The proposed model for evaluating competitiveness consists of two steps that we develop analytically (Fig. 1).

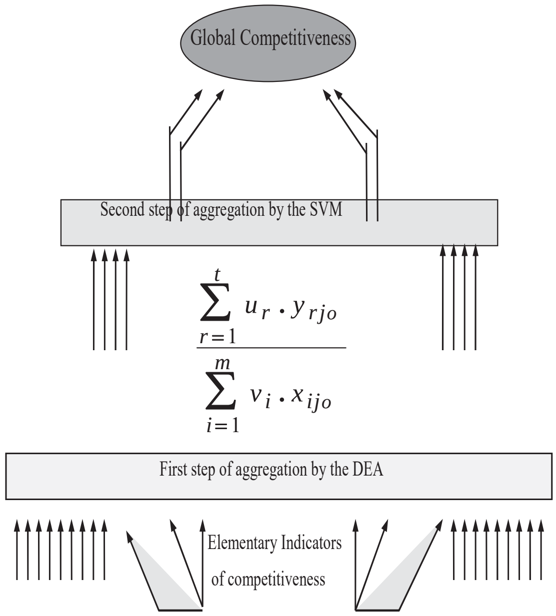

Fig. 1 The DEA-SVM Hybrid model

In the first step corresponding to the first hierarchical level, competitiveness evaluation is achieved by exploiting the DEA technique. This technique estimates the efficient frontier of all the DMU and measures their efficiency in relation to the optimal situation through the first equation. We adopt the below notation:

- y_{rj}: «favorable» Indicator r of the country j
- x_{ij}: «unfavorable» Indicator i of the country j
- u_r: weight of the «favorable» indicator r
- v_i: weight of the «unfavorable» indicator i
- n: Number of countries
- t: Number of «favorable» indicators
- m: Number of «unfavorable» indicators
- ε: very small number

Two orientations may characterize the developments of the DEA models: we seek either to reduce the inputs or to increase the outputs. On this basis, we consider extending the DEA methodology relating to the replacement of "input / output" dichotomy with another "favorable / unfavorable" dichotomy [13]. In this context, each DMU is going to consider a set of homogeneous competitiveness indicators noted E. From such a set, the DMU will seek to maximize a sub-set of indicators (variables) deemed "favorable" noted F and to minimize another sub-set of indicators (variables) deemed "unfavorable" noted D.

As it has just been presented, the DEA model is nonlinear. It may be transformed into a linear program while considering the dichotomy "favorable / unfavorable" (7):

$$\text{Maximize Ho} = \sum_{r \in F} u_r \, y_{rj_o} \tag{7}$$

Subject to: $\sum_{i \in D}^{m} v_i x_{ij_o} = 1$; $\sum_{r \in F} u_r y_{rj} - \sum_{i \in D}^{m} v_i x_{ij} < 0$ for j = 1,2... n; $u_r > \varepsilon$, $v_i > \varepsilon$

This model seeks a performance value by maximizing the weighted sum of the "favorable" indictors, under the constraint that the weighted sum of "unfavorable" indicators is equal to the unit value (1). The balances u_l and vi are the unknowns of the model. Thus, the most performing country is the one which manages, up to a point, to maximize its quotations at the level of the "favorable" indicators.

If set I, which expresses the homogeneous grouping of competitiveness indicators, is only made up of indicators of "favorable":

$$\text{Maximize Ho} = \sum_{r \in F} u_r \, y_{rj_o} \tag{8}$$

Subject to : $\sum_{r \in F} u_r y_{rj} < 1$ For j =1, 2,.. j_o ,..n

$u_r > \varepsilon$

If set I, is only made up of indicators of "unfavorable" type:

$$\text{Maximize Ho} = u \tag{9}$$

Subject to : $\sum_{r \in D} v_r y_{rj_o} = 1$; u - $\sum_{r \in D} v_r x_{rj} < 0$ pour j =1,2,..n

$u > 0$, $v_i > \varepsilon$

At the end of this stage, we reach the performance evaluation of all the countries to be studied at the level of the m homogeneous groupings of elementary indicators. Each country may be considered as a vector in a metric space of dimension m. In this case, the co-ordinates of each vector will be represented by the performance scores (competitiveness) drawn up by the DEA technique.

The second step regarded the first level as a new feature and added it into the previous feature vector. Based on the new feature set, SVM was applied to evaluate competitiveness at a higher hierarchical level. We are going to broaden the DEA logic as a measurement based on a geometrical interpretation. The scores obtained will be used to classify the countries according to their competitive level, at the decision-making hierarchical level involved.

We raise the situation that two countries can be arranged differently although they present a difference of little significance between the values of the scores calculated. Taking this perspective into account, we suggest a new competitiveness evaluation concept based on a stratified arrangement which allows the integration of every country into pre-defined classes of categories describing the measurement of competitiveness. Thus, for a decision-maker who defines five classes of competitiveness measurement, we can identify the following scale: very competitive country, competitive country, fairly competitive, poorly competitive and not competitive.

IV. Empirical Study

A. Data Set

As an empirical application, we have exploited the DEA-SVM to evaluate the competitiveness of 22 European countries that are Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, Turkey, and United Kingdom. The considered data source, which appears as a set of elementary performance indicators, is taken from the World Competitiveness Yearbook 2000 (WCY 2000).

The model (DEA-SVM) developed in this paper is integrated (programmed) into the models base of a Decision Support System (DSS). These models are coupled with the LINDO software for resolution linear program. Such a system allows the automatic generation of the DEA model outputs formulated at the first hierarchical level (Fig. 1) and gives the DEA-SVM a convivial application context.

We implemented the algorithm proposed in Fig. 1, from Stage 1 to Stage 3:

- Stage 1: DEA global level generation. We chose capability attributes as input and performance attributes as output for our model, and used software LINDO to allow the automatic generation of the DEA model outputs formulated at the first hierarchical level.
- Stage 2: Training. We mainly applied SVM to train the model with the same data using the DEA feature set. Suitable selection of kernel function and the related parameters may largely improve the prediction accuracy. To this end, we performed grid search to optimize the parameters C, γ associated with RBF kernel, and d associated with polynomial kernel based on 5-fold cross validation with 20 runs, where one run represents a new random subset split of the entire data set. Multiple runs are to aim at eliminating the instability of predictions arising from the small size of the data set. Here, the SVM was performed using LIBSVM and other methods were implemented using WEKA.
- Stage 3: Evaluation and classification. Given a new country and the corresponding attribute values, the DEA

score is firstly calculated based on Step 1. Then, the model trained on the entire data set is performed to identify the class that the new country belongs to.

B. Results

The result of the DEA-SVM treatments (the second and the third aggregation level) allowed us to provide the classification of the countries at the level of the eight factors and at the global level (sorted according to global level) (Table I) where the eight factors are coded as:

(1) Domestic Economy
(2) Internationalization
(3) Government
(4) Finance
(5) Infrastructure
(6) Management
(7) Science & Technology
(8) People

Taking these arrangements into account, we can analyze every country's competitive position in a relative multi-dimensional way and according to two hierarchical levels. In fact, we can characterize each country by a set of factors presenting the sources of power on the one hand, and another set of factors presenting the sources of weakness on the other hand. With this analysis, it will be possible to guide each country towards determining the actions (field of activity or factor) required for the improvement of its competitive position or its preservation and its reinforcement. It should be noted that a factor is considered as a source of power for a country if the latter is classified among the first three countries with respect to this factor. In the same way, a factor is considered as a source of weakness for a country if the latter is classified among the last three countries. It should be stated that the choice of the number of countries checking the selection condition must be specified a priori by the decision-maker.

The suggested procedure of stratified arrangement is applied to obtain the arrangement of the 22 European countries. We note that our model gives similar results as presented by WCY 2000 (Table II).

We conclude that the hybrid DEA-SVM model is a promising method that can be utilized as a competitive solution in the country classification area. An important advantage of this method is that it can be applied to identification on new country evaluation.

V. Conclusion

We presented a DEA-SVM model with the purpose of competitiveness evaluation and classifying the countries into four categories: (i) very competitive, (ii) competitive, (iii) fairly competitive, and (iv) not competitive.

The development of the present model is motivated by a twofold reason. On the one hand, the practical need to evaluate competitiveness while taking into account the multi-dimensional aspect of analysis, the relative measurement and the decisional hierarchical structure for a country. On the other hand, the search for a model that grasps the diversity of the multi-dimensional data structures and the data that may be aggregated according to a hierarchical structure.

TABLE I
ARRANGEMENT OF THE TWENTY-TWO EUROPEAN COUNTRIES ACCORDING TO THE FIRST LEVEL

Nations \ Factors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Global Level
Switzerland	7	8	3	1	7	14	2	1	1
Germany	2	1	5	6	13	20	1	6	2
France	1	2	8	10	12	10	4	8	3
Sweden	13	12	13	13	3	17	3	5	4
United Kingdom	5	3	7	2	16	8	9	16	5
Netherlands	10	5	10	4	11	13	5	4	6
Finland	6	9	11	7	4	16	6	11	7
Luxembourg	4	15	2	3	9	1	22	7	8
Norway	15	14	6	15	1	21	11	3	9
Austria	12	11	14	9	2	7	12	2	10
Ireland	3	10	15	14	6	2	10	15	11
Belgium	14	4	22	12	5	18	7	17	12
Spain	11	7	1	8	14	9	13	18	13
Denmark	17	13	19	5	10	19	8	14	14
Italy	8	6	21	18	15	15	14	19	15
Portugal	18	18	18	16	8	6	20	12	16
Hungary	16	19	12	19	17	5	15	10	17
Czech Republic	20	17	9	17	18	12	16	13	18
Slovenia	19	22	17	11	22	4	17	9	19
Greece	9	21	20	20	19	3	18	20	20
Poland	21	20	4	22	20	22	19	21	21
Turkey	22	16	16	21	21	11	21	22	22

TABLE II
ARRANGEMENT OF THE TWENTY-TWO EUROPEAN COUNTRIES ACCORDING TO THE GLOBAL LEVEL

Nation	Nation Ranking (Five classes)	Ordinal Rank DEA-SVM	Ordinal Rank WCY2000
Finland	competitive	1	1
Netherlands	competitive	2	2
Switzerland	very competitive	3	3
Ireland	competitive	4	5
Luxembourg	competitive	5	4
Denmark	fairly competitive	6	8
Germany	very competitive	7	6
Sweden	competitive	8	7
United Kingdom	competitive	9	9
Norway	competitive	10	10
Austria	competitive	11	11
France	very competitive	12	12
Belgium	competitive	13	13
Spain	competitive	14	14
Italy	fairly competitive	15	17
Portugal	fairly competitive	16	16
Czech Republic	poorly competitive	17	20
Hungary	poorly competitive	18	15
Greece	Not competitive	19	18
Slovenia	Not competitive	20	19
Poland	Not competitive	21	21
Turkey	Not competitive	22	22

To verify the feasibility of the proposed DEA-SVM model, countries evaluation is performed on an existing dataset of WCY 2000 for the 22 European countries. The contribution of this study can be summarized as follows: Firstly, the DEA method does provide valuable information in the country competitiveness evaluation. Secondly, the proposed DEA-SVM hybrid method provides a similar classification results given by WCY 2000. Hence, the SVM method has capacity on handling competitiveness evaluation problems on a small dataset. Although the dataset of countries competitiveness is very small, the results show that a very small-sized data set can give meaningful results in training DEA-SVM. The above-mentioned findings suggest that the DEA-SVM model should be a better alternative to conduct the countries evaluation tasks.

REFERENCES

[1] A. Asensio, Compétitivité et contrainte extérieure: comparaison de dix pays de l'OCDE sur la période 1970-1989, Problèmes économique, (1991), 2-252.
[2] A. Asensio and J. Mazier, Compétitivité, avantages coûts et hors coûts et spécialisation, Revue d'économie industrielle, (1991),55.
[3] M. Fouquin, La compétitivité prix, Compétitivité des nations, Edition Economica, (1998).
[4] S. Lall, Promouvoir la compétitivité industrielle dans les pays en développement, centre de développement, OCDE, (1991).
[5] B. Belon, La compétitivité, Traité d'économie industrielle, Economica (1991).
[6] S. Garelli and M. L. De Gurertechin, La compétitivité mondiale, Economie et statistique, (1995), 57-67.
[7] M. Porter, The competitive advantage of nations, the Free Press. A division of Macmillan Inc., New York, (1990).
[8] A. Charnes, W. W. Cooper, and E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research, (1978) vol.2, pp.429-442.
[9] R. D. Banker, A. Chames, and W. W. Cooper., Some models for estimating technical and scale inefficiencies in data envelopment analysis, Management Science, (1984), vol.30, pp.1078-1092.
[10] F. Y. Lin, and V. Lall, Using data envelopment analysis to compare suppliers for supplier selection and performance improvement, Supplier Chain Management: An International Journal, 2000vol.5, pp.143-150, March
[11] S. Talluri and J. Sarkis, A model for performance monitoring of suppliers, International Journal of Production Research, 2002 vol.40, pp.4257-4269, November.
[12] V. N. Vapnik, Statistical learning theory. Springer, New York (1998).
[13] H. Chabchoub, La modélisation de la compétitivité des nations: une approche de programmation mathématique, Ph.D, Université Laval, Juin (1998).