



	Comparison of alternative methods of updating input-output tables : a note	
Auteurs: Authors:	Daniel Leblanc, & Maurice Queyranne	
Date:	1979	
Type:	Rapport / Report	
Keielelice.	Leblanc, D., & Queyranne, M. (1979). Comparison of alternative methods of updating input-output tables: a note. (Technical Report n° EP-R-79-21). https://publications.polymtl.ca/6002/	
Document en libre accès dans PolyPublie Open Access document in PolyPublie		
URL de PolyPublie: PolyPublie URL:		https://publications.polymtl.ca/6002/
Version:		Version officielle de l'éditeur / Published version
Conditions d'utilisation: Terms of Use:		Tous droits réservés / All rights reserved
Document publié chez l'éditeur officiel Document issued by the official publisher		
Institution:		École Polytechnique de Montréal
Numéro de rapport: Report number:		EP-R-79-21
URL officiel: Official URL:		
	n légale: gal notice:	



Génie Industriel

COMPARISON OF ALTERNATIVE METHODS OF UPDATING INPUT-OUTPUT TABLES: A NOTE

DANIEL LEBLANC AND MAURICE QUEYRANNE

DEPARTEMENT DE GENIE INDUSTRIEL ECOLE POLYTECHNIQUE DE MONTREAL

JUNE 1979

RAPPORT TECHNIQUE NO. EP-79-R-21

Ecole Polytechnique de Montréal

CA2PQ UP4 79R21

Campus de l'Université de Montréal Case postale 6079 Succursale 'A' Montréal, Québec H3C 3A7 COMPARISON OF ALTERNATIVE METHODS OF UPDATING INPUT-OUTPUT TABLES: A NOTE

DANIEL LEBLANC AND MAURICE QUEYRANNE

DEPARTEMENT DE GENIE INDUSTRIEL ECOLE POLYTECHNIQUE DE MONTREAL

À CONSULTER SUR PLACE

JUNE 1979
RAPPORT TECHNIQUE NO: EP-79-R-21

COMPARISON OF ALTERNATIVE METHODS OF

UPDATING INPUT-OUTPUT TABLES: A NOTE*

Daniel LEBLANC and Maurice QUEYRANNE

ECOLE POLYTECHNIQUE DE MONTREAL DEPARTEMENT DE GENIE INDUSTRIEL

June 1979.
RAPPORT TECHNIQUE NO. EP79-R-21

^{*} This research was motivated by a study on the Input-Output table of Québec carried out for the Office de Planification et de Développement du Québec. The authors also thank B. T. Smith for his comments on a first draft of this note.

ABSTRACT: Various methods for updating Input-Output tables have been proposed and comparisons have been made on statistical bases. In this note we show on theoretical bases that several linear programming formulations are not suitable to this problem since a large proportion of the resulting coefficients assume arbitrarily preset values.

The problem of updating Input-Output tables from an initial table and known row an column totals is frequently faced by government agencies in their policy-making activities. The same problem also arises in statistical studies [14],[6] as well as in the estimation of telephone traffic [7], [11], [4] and urban trip distribution [5, chap. 4]. Various methods for solving these problems have been proposed (see the surveys by Lecomber [8] and Stemmeulen [13] for the application to Input-Output tables).

Several comparisons have been made between these methods on an experimental or statistical basis. In particular Davis et al. [2] have compared the well-known RAS method and two variants of a Linear Programming (LP) method proposed by Matuszewski et al. [9]. They have used the 1963 US National Input-Output table to derive a 1967 updated table which was compared with the actual 1967 data. They concluded that there is a clear superiority of the RAS technique over the LP approaches.

In this note theoretical considerations are used to establish the inadequacy of these LP approaches to this problem. We show how the proposed LP method can be formulated as a variant of the transportation

problem of linear programming, the solution of which involves a large proportion of updated coefficients assuming arbitrarily preset values.

In mathematical notations, the problem considered in [9] and [2] is the following: given an m x n table of reference values \mathbf{x}_{ij} , row and column totals \mathbf{w}_i^t (i=1,...m) and \mathbf{u}_j^t (j=1,...n) and lower and upper bounds \mathbf{r}_{ij}^t and \mathbf{s}_{ij}^t respectively for cell (i,j) for year t, find values \mathbf{x}_{ij}^t for all cells (i,j) such that $\mathbf{x}_{ij} \neq 0$ which

minimize
$$\Sigma \mid x_{ij} - x_{ij}^t \mid / x_{ij}$$
 (1) (x_{ij}^t)

subject to
$$\sum_{i} x_{ij}^{t} = u_{j}^{t}$$
, (2)

$$\sum_{i} x_{ij}^{t} = w_{i}^{t}, \qquad (3)$$

$$r_{ij}^{t} \leq x_{ij}^{t} \leq s_{ij}^{t} . \tag{4}$$

Here |y| denotes the absolute value of y and it is assumed that $\sum u_j^t = \sum w_i^t$. Here and in the sequel, only cells (i,j) such that $\mathbf{x}_{ij} > 0$ are considered.

Since the values x_{ij} , r_{ij}^t and s_{ij}^t are given for cell (i,j) before solution of (1) - (4), we call them the *preset values* related to variable

 \mathbf{x}_{ij}^{t} . It will be shown that in an optimal solution to (1) - (4) most of these \mathbf{x}_{ij}^{t} 's will assume one of the preset values.

Problem (1)-(4) is a nonlinear programming problem which can be reduced to the LP format in different ways. Matuszewski et al. [9] proposed such a reduction. We introduce here a slightly different transformation providing more insight into the structure of optimum solutions to problem (1)-(4).

Consider the following change of variables,

$$x_{ij}^{t} = y_{ij}^{t} + z_{ij}^{t}$$
(5)

with (i) if $r_{ij}^t \le x_{ij}^t \le s_{ij}^t$ then

$$r_{ij}^t \le y_{ij}^t \le x_{ij}$$
 and (6a)

$$0 \leq z_{ij}^{t} \leq s_{ij}^{t} - x_{ij}; \tag{7a}$$

(ii) if
$$r_{ij}^t \leq s_{ij}^t \leq x_{ij}$$
 then

$$r_{ij}^{t} \leq y_{ij}^{t} \leq s_{ij}^{t}$$
 and (6b)

and
$$z_{ij}^t = 0;$$
 (7b)

(iii) if
$$x_{ij} \le r_{ij}^t \le s_{ij}^t$$
 then

$$y_{ij}^t = x_{ij}$$
 and (6c)

$$r_{ij}^{t} - x_{ij} \leq z_{ij}^{t} \leq s_{ij}^{t} - x_{ij}^{t}; \qquad (7c)$$

and define the LP problem:

minimize
$$\sum_{\substack{(y_{ij}^t, z_{ij}^t)}} ((x_{ij}^t - y_{ij}^t) + z_{ij}^t) / x_{ij}$$
 (8)

subject to
$$\sum_{i} y_{ij}^{t} + z_{ij}^{t} = u_{j}^{t}$$
, (9)

$$\sum_{i} y_{ij}^{t} + z_{ij}^{t} = w_{i}^{t}, \qquad (10)$$

$$(6a) - (7c)$$
. (11)

Clearly the constraints (9) - (11) are equivalent to (2) - (4) and problem (1) - (4) is feasible if and only if problem (8) - (11) is feasible.

We claim that in an optimum solution $(\bar{y}_{ij}^t, \bar{z}_{ij}^t)$ to (8) - (11), we cannot have both $\bar{y}_{ij}^t < x_{ij}$ and $\bar{z}_{ij}^t > 0$. For otherwise adding a constant δ to \bar{y}_{ij}^t and subtracting it from \bar{z}_{ij}^t for the same cell (i,j) violating this condition defines another feasible solution whenever $0 < \delta \leq \min\left\{x_{ij} - \bar{y}_{ij}^t; \bar{z}_{ij}^t\right\}$ with objective function value (8) smaller by $2 \delta/x_{ij}$.

Hence under conditions (9) - (11), minimizing (8) is equivalent to:

minimize $\sum_{i,j} |(x_{ij} - y_{ij}^t) + z_{ij}^t| / x_{ij}$ (12)

which in turn is equivalent to minimizing (1).

The LP problem (8) - (11) is a generalization of the classical transportation problem of linear programming in which parallel arcs (or cells) and both lower and upper bounds on these arcs are allowed. Since this problem has m + n constraints (in addition to the bound constraints (11)) and since m + n - 1 of these constraints are linearly independent [1], [12], it follows from LP theory (ibid.) that when the problem is feasible there exists an optimum (basic) feasible solution in which at most m + n - 1 variables y_{ij}^t and z_{ij}^t assume a value distinct from their bounds. Thus there is an optimum solution to problem (1) - (4) in which at most m + n - 1 variables x_{ij}^t assume a value distinct from the preset values x_{ij}^t , r_{ij}^t and s_{ij}^t .

In the experiment reported in [2], the size of the table was m = n = 50 and there were 1899 nonzero $x_{ij}^{\ \ \ \ }$ s. From the observations made in the previous paragraph it can be expected that at least 1800 coefficients x_{ij}^t would be equal to some of their preset values. Indeed in this study [2], the authors solved problem (1) - (4) through a linear programming formulation LP(1) distinct from (8) - (11). They made interesting comparisons between the computed values and the actual 1967 coefficients and they also gave in their Table 2 some details on the numbers of computed values which deviated from the base year coefficients (corresponding to x_{ij} in (1) - (4). From these data, we infer that 1801 coefficients did not change from their initial value. Furthermore among the remaining 98 changing coefficients there may be still some which hit their lower or upper bounds. The authors also used another formulation LP(2) in

which all $r_{ij}^t = 0$ and $s_{ij}^t = +\infty$. The corresponding total number of computed coefficients keeping their base year value was then 1790, that is again more than 94% of the nonzero cells.

The major drawback of the LP approach to this problem seems to be that it implies that too many coefficients remain to their initial values while it allows quite large variations for the others. Indeed this is the reason for the introduction of lower and upper bounds in (4), "admittedly very arbitrary" [9 p. 206]. In addition these methods do not seem appropriate for incorporating additionnal information limiting the possible variation of some coefficients: indeed most of the coefficients will either keep the initial value or assume an extreme value. The same observation also applies for a convex piecewise linear objective function in (1). In that case most of the coefficients will assume a value corresponding to a breakpoint of the objective fonction, that is, again a quite arbitrary value.

In conclusion it seems advisable, on the grounds of these theoretical remarks, to use a more smooth convex objective function instead of (1). The procedures of Almon, Friedlander and the classical RAS method itself are of this type - see [8 p.3] for a survey. According to Lecomber, "available evidence (...) suggests that, provided the initial discrepancies are not too large, the results are not very sensitive to the method used. The comparative simplicity of RAS over the other methods that

preserve signs is an overriding asset", [8 p.2]. In addition such methods can easily take advantage of additional information obtained through complementary methods such as sampling and interview methods [10], [3] or the interesting proposal recently made by Stemmeulen [13], of using methods of factorial analysis.

REFERENCES

- [1] DANTZIG, G.B. <u>Linear Programming and Extensions</u>. Princeton University Press, Princeton, N.J. (1963).
- [2] DAVIS, H.C., LOFTING, E.M. and SATHAYE, J.A. "A Comparison of Alternative Methods of Updating Input-Output Coefficients". Technological Forecasting and Social Change 10, 79-87 (1977).
- [3] FONTELA, E., DUVAL, A., GABUS, A., BORLIN, M. and VELAY, C.

 "Forecasting Technical Coefficients and Changes in relative prices", in A.P. Carter and A. Brody (eds.). Application of Input-Output Analysis. North-Holland, Amsterdam (1970).
- [4] GIRARD, A. "Etude de méthode de prédiction du trafic téléphonique." INRS Technical Report 78-04, INRS Télécommunications, Verdun, Québec, Canada (1978).
- [5] HUTCHINSON, B.G. <u>Principles of Urban Transport Systems Planning</u>. Scripta Book Co. Washington D.C. (1974).
- [6] IRELAND, C.T. and KULLBACK, S. "Contingency Tables with Given Marginals." <u>Biometrika</u> 55, 179-188 (1968).
- [7] KRUITHOF, J. "Telefoonverkeerscrekening." De Ingenieur 52 (8) 15-25 (1937).
- [8] LECOMBER, J.R.C. "A Critique of Methods of Adjusting, Updating and Projecting Matrices," in R.I.G. Allen and W.F. Gossling (eds.). Estimating and Projecting Input-Output Coefficients, Toput-Output Publishing Co., London (1975).
- [9] MATUSZEWSKI, T.I., PITTS, P.R. and SAWYER, J.A. "Linear Programming Estimates of Changes in Input Coefficients". Canadian Journal of Economics and Political Science 30(2), 203-210 (1964).

- [10] MIERNYK, W.H. "Sampling techniques in making regional industry forecasts," in A.P. Carter and A. Brody (eds.), Application of Input-Output Analysis. North-Holland, Amsterdam (1970).
- [11] NIVERT, K, and von SCHANTZ, C. "Some Applications of Quadratic Programming to the Calculation of Traffic Matrices."

 Proc. 8th International Teletraffic Congress, pp. 333-1 to 333-5 (1976).
- [12] SIMONNARD, M. <u>Programmation linéaire</u>, vol. 1 and 2. Paris, Dunod, (1972).
- [13] STEMMEULEN, E. "Tableaux d'échanges: description et prévision."

 Cahiers du Bureau Universitaire de Recherche Opérationnelle,
 Série Recherche, no. 28, 1-83 (1977).
- [14] STEPHAN, F.F. "An Iterative Method of Adjusting Sample Frequency Tables when Expected Marginal Totals Are Known." Annals of Mathematical Statistics 13, 168-178 (1942).



I