Deconvolving the Productivity of Salespeople via Constrained Quadratic Programming

Gautam K. Bhat
IBM Global Technology Services
Ramapuram, Chennai, India
gautambh@in.ibm.com

Kush R. Varshney
IBM Thomas J. Watson Research Center
Yorktown Heights, New York, USA
krvarshn@us.ibm.com

***Abstract*—With the present market trend, businesses and organisations with large salesforces are experiencing much turnover among their sellers. Movement of salespeople from one company to another is a continual process as long as there is market demand. In the traditional sense, a salesperson's productivity is directly proportional to the revenue that he or she brings to the company. Importantly, the senior leaders in organisations are interested in knowing the variations in sales productivity as a result of hiring and attrition in the salesforce. In this paper we focus our attention on the characterisation of sales productivity based on four categories. When an existing salesperson leaves, what is the sales productivity over time if replaced by a new hire from a university, an experienced new hire, or a transfer from another division in the company? In addition if an organisation ventures into acquisition, what is the anticipated sales productivity from this? We model the sales productivity of new hires as a linear time-invariant system and estimate productivity profiles with a least-squares deconvolution formulation. By applying business constraints on productivity profiles for regularisation, we are left with a constrained quadratic program to solve. We demonstrate the estimation technique on real-world sales data from a global enterprise, finding productivity profiles under the four different cases listed above.**

Keywords-business applications; deconvolution; estimation theory; least-squares; salesforce analytics

I. Introduction

Systems theory, broadly construed, is concerned with modelling, analysing, and optimising a set of interacting components that form an integrated whole. Those components could be mechanical machines, electrical or electronic elements, or even human beings [1]. In this paper, the particular system that we focus on is a salesforce whose interacting components are individual salespeople. Such study of sellers falls under the scope of business analytics, specifically salesforce analytics [2]–[4].

When new salespeople join an enterprise, whether hired directly from a university, as a result of a merger or acquisition, through an internal transfer from a different division, or hired with experience from the industry at-large, it takes time for them to get acclimated to the organisation, learn about the product and service offerings, and build contacts. Thus sellers have a period of little to no productivity followed by a ramping up period until they have reached the productivity of sellers that have been with the enterprise for a longer duration. The amount of revenue earned is the traditional measure of productivity for salespeople.

The problem that we study herein is characterising the productivity of newly-hired salespeople, specifically determining productivity profiles as a function of time after hiring (for different types of new hires). Such characterisations are important for planning purposes because the head count of sellers is not a true indication of the productivity of the salesforce; if a target productivity is desired at a particular time in the future, then hiring decisions must be made in the present based on the productivity profiles [5]. Brooks' Law, which states that "adding manpower to a late software project makes it later" [6] applies equally well to sales productivity because of ramp-up time [7]: if there is a shortage of sellers today, then hiring more sellers today will not improve productivity because they will have no initial productivity, and may in fact transiently decrease productivity of the overall salesforce as they are integrated into the enterprise.

In studying novel systems, Willsky asks [8], "How can we extend existing mathematical methodologies? How can we use existing methodologies in the context of a specific physical problem to obtain a tractable formulation which addresses the issues of interest in the more ill-defined physical problem?" The novel system study here is that of sales productivity characterisation. In the context of this specific 'physical' problem, the tractable existing framework within which we pursue our methodology is that of convolutive discrete-time linear time-invariant (LTI) systems [9]. We seek to identify the system that transforms head count at various post-hiring times to sales productivity or revenue. Specifically, we use a least-squares formulation of deconvolution to do so [10], [11], which leads to a quadratic programming optimisation.

The extension of the basic least-squares identification that we present here further models the system through business constraints for purposes of regularisation. We include non-negativity constraints, which often arise in statistical learning optimisations [12]. Additional business constraints lead us to additional mathematical constraints related to monotonicity and smoothness of the seller productivity profiles as well as a saturation level of the profiles. Thus overall, we optimise a constrained quadratic program to characterise the productivity of new sellers as a function of the time since they were hired by the enterprise.

Constrained quadratic programming for time profile estimation also arises in application domains other than salesforce analytics. For example, there is a constrained quadratic programming formulation in [13] to estimate transcriptional profiles of clinical blood samples. Also, there is a similar formulation in [14] to characterise event profiles from functional magnetic resonance imaging data sources. Although there have been theoretical, mathematical models of salesforces in the marketing literature previously [15], systems-theoretic thinking and analysis has not been applied to the problem of new seller productivity. In our previous work [7], we did not adopt a linear time-invariant system and deconvolution perspective to the problem.

This paper discusses the specification, identification, and estimation of sales productivity of salespeople. This involves measurements on a set of criteria at different levels of skill and experience. The numbers of measurements are taken over a long period of time to get as accurate results as possible. In optimising the sales productivity result, we have considered all the requisite constraints that are applicable to the salesforce. Optimisation techniques need to be performed on a huge data set and over a period of time. Once we have the result, that is not the end of the world; it needs to be validated and verified from a logical perspective. Several constraints are taken into account so that a definite logical result is derived. In the context of optimising sales productivity result, it is imperative to say that the result must be non-negative. A salesperson irrespective of being a fresh graduate, experienced, or so on, will only generate revenues, hence the optimised result must always be non-negative. The deconvolution method assumes that: 1) productivity of the fresh sales graduate may not be steep in the initial phase of his career, hence exceptional skills are not taken into consideration here, and 2) same two salespeople may or may not contribute to the same level of sales productivity. Again, recent studies have shown that the workforce productivity may vary from trial to trial.

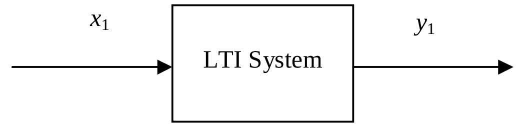
Here we have divided the paper into several smaller parts. We begin with the system model to give a background of what is being done and a high level understanding of the need to go for deconvolution. Followed by this we provide the algorithm associated with the derivation followed by experiments and results section with details on every step performed and analysed. Finally we have the discussion section ending with conclusion.

II. System Model

In this section, we first describe LTI systems and how sales revenue fits into that framework. Then we describe the system model of sales productivity that we employ.

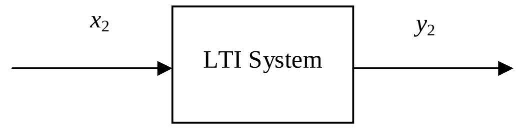
A. Linear Time-Invariant Systems

LTI systems produce same amount of output for the similar amount of input passed. In other words, if x_1 amount of input is fed to the system, it will produce y_1 amount of output. If x_2 amount of input is applied, then, y_2 amount of output is produced. The following pictorial representation will make it clearer.



(1)

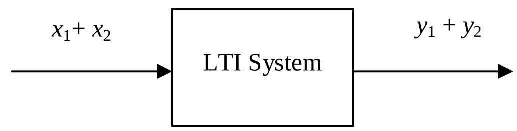
The above diagram can be interpreted such that for an input x_1, the output generated is y_1.



(2)

For an input x_2, the output generated is say y_2.

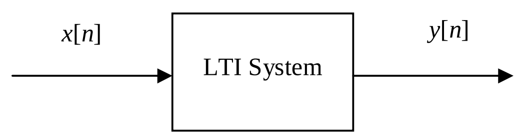
Similarly, if the above two inputs are added, the output will also be summed up as shown.



(3)

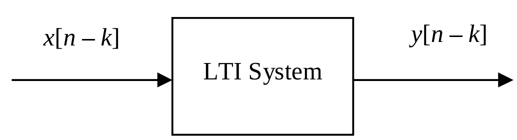
The above analogy applies to sales productivity in that if we double the number of sellers, the revenue will double.

Along similar lines, time-invariant systems always produce the same output irrespective of when the input is sent or applied to the system. This means that if we hire someone in August 2010 or in April 2012 or in January 2020, if seven months have elapsed since hiring, his or her productivity will be the same. This can be visualised through a diagram below.



(4)

In the above diagram, input x is applied at time n which generates output y for the same time n.



(5)

Input x applied at time $n – k$ still generates the same output y at time $n – k$.

The transformation of inputs to outputs by LTI systems is described by the convolution operation. The output is the convolution of the input with the unit pulse response of the system. Convolution can be expressed in matrix-vector form by taking either the input signal or the unit pulse response signal of the system as a vector with the other used to construct a convolution matrix.

B. Sales Productivity Model

As discussed in the previous section, we can model sales productivity as an LTI system taking counts of sellers as input and revenue as output. The unit pulse response of the system is then the productivity profile. Let us first fix the notation so that the input signal $x[k]$ represents the number of sellers that were hired k months ago. Similarly, let $h[k]$ represent the unit pulse response of the system and also the productivity profile so that $h[-k]$ is the productivity of a seller that was hired k months ago. The output y is the total revenue produced by all sellers in the salesforce. Assuming that the system is causal with a finite impulse response of length m, the convolution leading to the output is

$$y[n] = \sum_{k=0}^{m} x[n-k]h[k] \tag{6}$$

Here, n represents different months in which we observe the system. We can also write this as the matrix-vector equation

$$\mathbf{y} = \mathbf{X}\mathbf{h}, \tag{7}$$

where $\mathbf{X}$ is the convolution matrix formed from x.

Since we are dealing with four different classes of new sellers, we in fact have four different productivity profiles h and four different counts of new sellers, but that can be represented similarly to (2):

$$\mathbf{y} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 & \mathbf{X}_4 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \\ \mathbf{h}_4 \end{bmatrix} \tag{8}$$

where the $\mathbf{X}_i$, $i = 1, \ldots, 4$, matrices are again convolution matrices. In the sequel, to keep the notation simple, we refer to (2) as the representation of the system.

We measure the revenues for the entire salesforce y and we also measure the head counts of new sellers each month. The productivity profiles are to be estimated.

III. Estimating Productivity Profiles using Quadratic Programming

Since the revenues $\mathbf{y}$ and the head counts $\mathbf{X}$ are measured, our task is to estimate the productivity profile $\mathbf{h}$. We take a least-squares approach by minimising the ℓ_2 norm between the measured revenue and the estimated output of the LTI system:

$$\|\mathbf{y} - \mathbf{X}\mathbf{h}\|_2^2 . \tag{9}$$

This can be written as a quadratic program with the following objective function to find the productivity profile solution,

$$\min_{\mathbf{h}} \tfrac{1}{2}\mathbf{h}^T\mathbf{X}^T\mathbf{X}\mathbf{h} - \mathbf{h}^T\mathbf{X}^T\mathbf{y} . \tag{10}$$

where $\mathbf{h}$ is the sales productivity vector.

One of the simplest forms of regularisers is given by the sum of squares of the unit pulse response vector elements: $\tfrac{1}{2}\mathbf{h}^T\mathbf{h}$ [11]. With the least-squares objective and this regulariser, the solution is the Moore-Penrose pseudoinverse

$$\mathbf{h} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y} . \tag{11}$$

However, in our sales analytics problem, we have further business constraints to motivate additional regularisation. The constrained quadratic program that we use to find $\mathbf{h}$ is the following:

$$\begin{aligned} \min_{\mathbf{h}} \quad & \tfrac{1}{2}\mathbf{h}^T\mathbf{X}^T\mathbf{X}\mathbf{h} - \mathbf{h}^T\mathbf{X}^T\mathbf{y} \\ \text{s.t.} \quad & \mathbf{h}_{lb} \le \mathbf{h} \le \mathbf{h}_{ub} \\ & \mathbf{A}_{in}\mathbf{h} \le \mathbf{A}_{ub} \end{aligned} . \tag{12}$$

We have four constraints motivated by the business application. First, productivity profiles are non-negative because sellers cannot produce negative revenue; if they sell nothing, their productivity is zero. Second, we assume that the productivity profiles are monotonically non-decreasing because over time, the sellers gain experience, knowledge, and contacts, so their productivity does not get worse over time. Third, we assume that the productivity does not rapidly jump from time step to time step, so we constrain increases in the productivity profile to not exceed a certain value. Last, we impose a maximum productivity L, which is a saturation level and the productivity of sellers that have been with the enterprise for more than m months.

Specifically, for the non-negativity constraint, we set $\mathbf{h}_{lb}$ to be a length m vector of all zeroes. For the saturation constraint, we set $\mathbf{h}_{ub}$ to be a length m vector with all entries equal to a parameter L. The other constraints we include are encoded through the matrix $\mathbf{A}_{in}$ and the vector $\mathbf{A}_{ub}$. It is straightforward to encode that successive values of $\mathbf{h}$ be monotonically non-decreasing and also that successive values of $\mathbf{h}$ not increase by more than another parameter value that we set. $\mathbf{A}_{in}$ has blocks that are Toeplitz matrices compose of positive and negative ones. Half of $\mathbf{A}_{ub}$ is all zeroes and the other half is equal to the parameter value indicating the increase limit per time in $\mathbf{h}$.

Having derived the optimisation problem (7), we find the optimised value of the sales productivity profiles $\mathbf{h}$ using the `qp` function of Octave. To construct the $\mathbf{X}$ matrix, we utilise the `convmtx` function of Octave. To begin with, we adopted the most simplified solution by assuming that there are no constraints in the problem statement. No constraint scenario is realised by considering empty vectors and matrices on all parameters in the above function.

First, we calculated the matrix with minimal constraints as below:

$\mathbf{h}$ = `qp(zeros(`m`,1)`, $\mathbf{X}^T\mathbf{X}$, $-\mathbf{X}^T\mathbf{y}^T$).

To get non-negative values we define a few more variables:

$\mathbf{h}$ = `qp(zeros(`m`,1)`, $\mathbf{X}^T\mathbf{X}$, $-\mathbf{X}^T\mathbf{y}^T$, [], [], `zeros(`$m$`,1)`, []).

Finally, the full constrained quadratic program is solved as

$\mathbf{h}$ = `qp(zeros(`m`,1)`, $\mathbf{X}^T\mathbf{X}$, $-\mathbf{X}^T\mathbf{y}^T$, [], [], `zeros(`$m$`,1)`, L`*ones(`m`,1)`, [], $\mathbf{A}_{in}$, $\mathbf{A}_{\text{ub}}$),

where we use the Octave function `toeplitz` to construct $\mathbf{A}_{in}$.

IV. Empirical Results

We collected data from one of the business units of International Business Machines (IBM). The data corresponds to one of the sales organisations from a few recent years. Hiring information comes from human resources (HR), and revenue details come from the finance section.

Fig. 1 shows the total head count of sellers in the organisation during the period of examination. Fig. 2 shows head count dynamics for the different classes of new sellers. The values shown indicate the number of sellers of each category that joined the organisation during the month. Fig. 3 displays the solutions for the productivity profiles that we obtained, i.e. it plots time duration in months and revenue generated by the sales person for 4 different cases. Fig. 4 shows the actual total revenue of the organisation and the revenue reconstructed using the actual head counts and the learned productivity profiles **h**.

Case 1 shown in blue colour in Fig. 3 indicates the plot for a salesperson who has been transferred from one unit to the other. As shown in the figure, this salesperson is expected to bring in good revenue over a short period of time and remains constant thereafter. Transfers in have one month of no

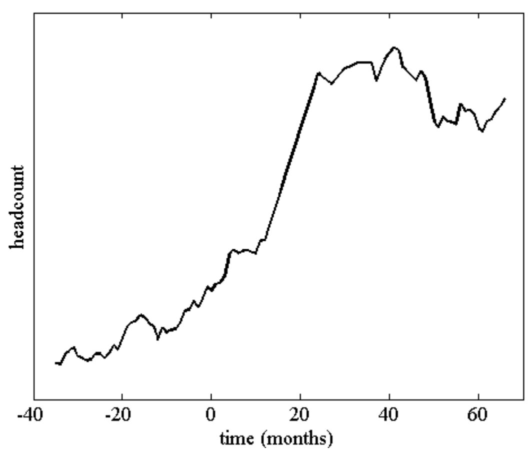


Figure 1. Total head count data of salespeople during the period of examination.

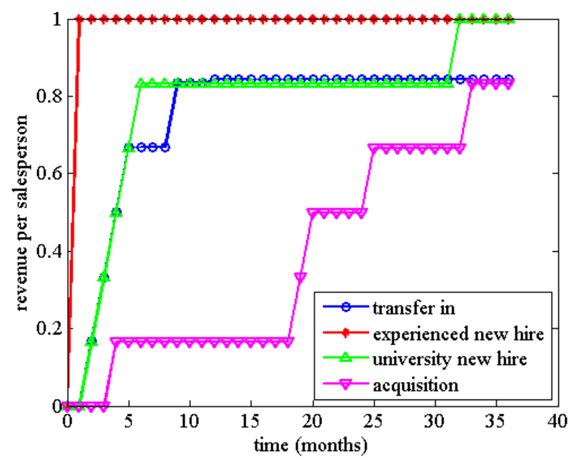


Figure 3. Productivity profiles for four different cases estimated from head count data and revenue generated by sellers for given time duration.

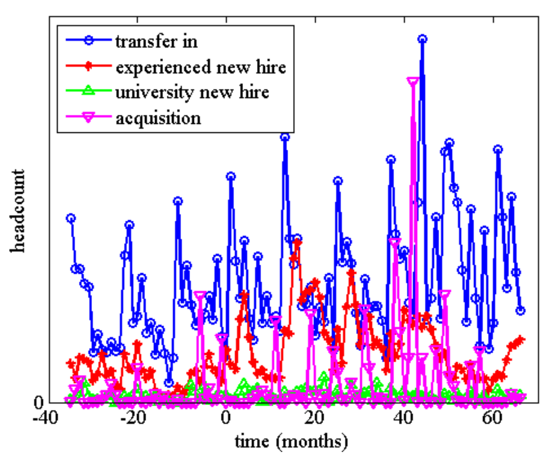


Figure 2. Head count dynamics for different classes of sellers during the period of examination.

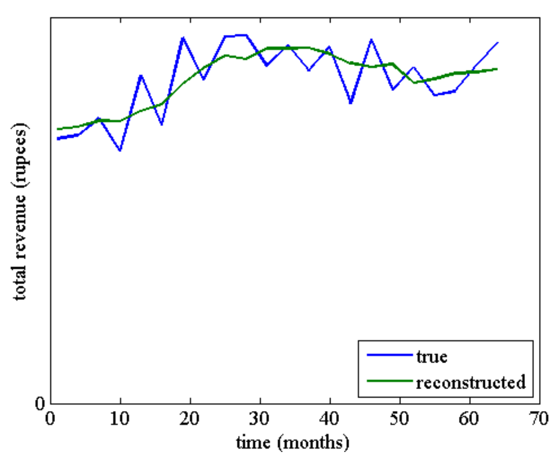


Figure 4. Revenue fit using revenue reconstructed from the actual head counts and the estimated productivity profiles.

productivity and do not reach the steady state productivity L.

Case 2 shown in red colour indicates the performance of an experienced salesperson who joins the organisation. The productivity of this category of salespeople is instantly high as they do not take much time to settle and perform. The vertical scale has values labelled between 0 and 1. This is so because we have normalised the value from 0 to 1 to maintain confidentiality of the sales data of the organisation.

Case 3 relates to new hires from colleges and universities who do not already have prior work experience. At this point all that they possess is theoretical knowledge requiring practical exposure. They, like the transfers in have one month of no productivity, but do reach the saturation value L. Their progression is similar to transfers in.

Case 4 indicates the nature of involvement of the salesforce as a result of acquisition. In this competitive world, acquisitions are not rare and the implications of this on an organisation's sales productivity are extremely important to understand. Accurate forecasting of sales productivity has a direct impact on organisations' income and revenue generation. Those that join due to acquisitions have several months of no productivity at the beginning and ramp up very slowly. They do not reach the steady-state productivity L like the transfers in.

V. Conclusion

In this paper we have discussed the impact of varied classes of salespeople on the sales productivity. We gathered data from IBM and based our analysis on linear time-invariant systems theory, mathematical quadratic programming, and pattern recognition techniques. We study problems in quadratic programming where the optimisation is confined to nonnegative constraint. For these problems, we might get a negative value which does not make sense in the sales world and so we perform numerical computations using that constraint. The remaining three constraints covered in this paper are profile values being monotonically non-decreasing, successive profile values not increasing by more than a certain amount, and profile values being less than or equal to a predefined upper limit value.

The gut instinct of sales leaders is accurate to a point, but through business analytics, we are able to produce more refined and exact productivity profiles for four cases, namely, transfer in, experienced new hires, university new hires and those sellers that join the organisation as a result of acquisition. The productivity of a salesperson is difficult to quantify accurately. Estimating the trend of how each of the 4 categories of salespeople behave is based on the assumption that the criticality, work environment and challenges faced, and complexity of the engagement is the same for all sellers under each of the categories and also for all of them in general.

The educational background of sellers from any of the four categories is not taken into account in this paper. There could be the possibility of the common notion that educational qualification from a premium university, college or institute results in higher productivity because such universities and institutes may infuse great amount of confidence in their outgoing students who become salespeople. These issues fall under behavioural and psychological aspects of business analytics. However, these issues are not considered among the constraints that we chose for calculations here because we are focused on the productivity trend over a period of time and not necessarily on the amount of productivity from the individual salespeople; also, we do not have data concerning them available readily. Additionally, this is a debatable subject; hence it is safe to keep such considerations out of the scope of this paper.

References

[1] H. S. Sekhon, R. K. Varshney, P. Kumar, J. M. Singh, and R. Prasad, eds., *Glimpses of Systems Theory and Novel Applications: Felicitation Volume in Honour of Prof. R. K. Varshney*. Aligarh, India: Navin Press, 2005.

[2] T. H. Davenport, J. Harris, and J. Shapiro, "Competing on talent analytics," *Harvard Bus. Rev.*, Oct. 2010.

[3] K. R. Varshney and A. Mojsilović, "Business analytics based on financial time series," *IEEE Signal Process. Mag.*, vol. 28, no. 5, Sep. 2011, pp. 83–93.

[4] M. Baier, J. E. Carballo, A. J. Chang, Y. Lu, A. Mojsilović, M. J. Richard, M. Singh, M. S. Squillante, and K. R. Varshney, "Sales-force performance analytics and optimisation," *IBM J. Res. Dev.*, vol. 56, no. 6, Nov./Dec. 2012.

[5] M. Singh, D. Bhattacharjya, L. Deleris, D. Katz-Rogozhnikov, M. Squillante, B. Ray, A. Mojsilović, C. Barrera, D. Kakrania, J. Richard, A. Saha, and J. Fu, "The growth and performance diagnostics initiative: A multidimensional framework for sales performance analysis and management," *Serv. Sci.*, vol. 3, no. 1, pp. 82–98, Spring 2011.

[6] F. P. Brooks, Jr., *The Mythical Man-Month: Essays on Software Engineering*. Reading, MA, USA: Addison-Wesley, 1995.

[7] K. R. Varshney, M. Singh, M. Sharma, and A. Mojsilović, "Estimating post-event seller productivity profiles in dynamic sales organisations," in *Proc. IEEE Int. Conf. Data Mining Workshops*, Vancouver, Canada, Dec. 2011, pp. 1191–1198.

[8] A. S. Willsky, "Some solutions, some problems, and some questions," *IEEE Contol Syst. Mag.*, vol. 2, no. 3, Sep. 1982, pp. 4–16.

[9] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & Systems*. Upper Saddle River, NJ, USA: Prentice Hall, 1997.

[10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[11] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Burlington, MA, USA: Academic Press, 2009.

[12] L. K. Saul, F. Sha, and D. D. Lee, "Statistical signal processing with nonnegativity constraints," in *Proc. Eur. Conf. Speech Comm. Tech.*, Geneva, Switzerland, Sep. 2003, pp. 1001–1004.

[13] T. Gong, N. Hartmann, I. S. Kohane, V. Brinkmann, F. Staedtler, M. Letzkus, S. Bongiovanni, and J. D. Szustakowski, "Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples," *PLoS ONE*, vol. 6, no. 11, Nov. 2011.

[14] Y. Lu, T. Jiang, and Y. Zang, "Single-trial variable model for event-related fMRI data analysis," *IEEE Trans. Med. Imag.*, vol. 24, no. 2, Feb. 2005.

[15] K. S. Moorthy, "Theoretical modeling in marketing," *J. Marketing*, vol. 57, no. 2, Apr. 1993.