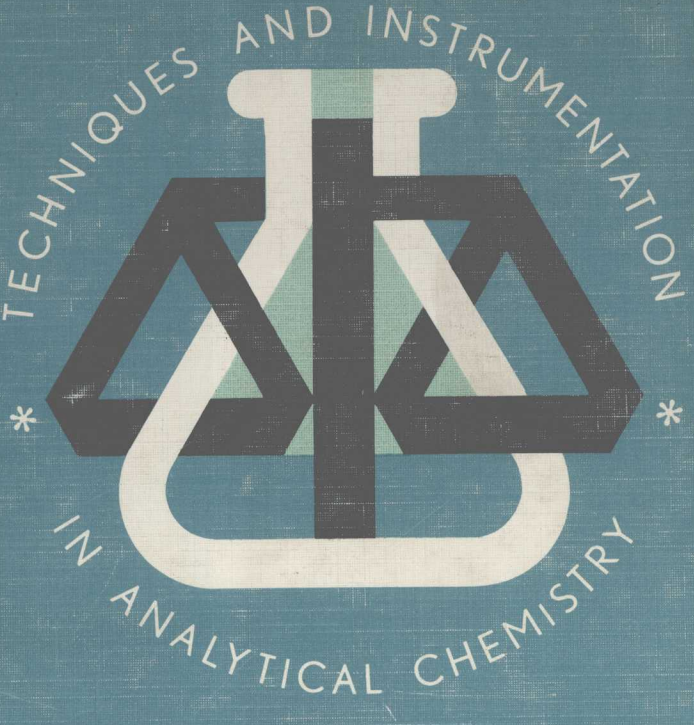
1



EVALUATION AND OPTIMIZATION OF LABORATORY METHODS AND ANALYTICAL PROCEDURES

A Survey of Statistical and Mathematical Techniques

D.L. Massart, A. Dijkstra and L. Kaufman

with contributions by S. Wold, B. Vandeginste and Y. Michotte

ELSEVIER SCIENTIFIC PUBLISHING COMPANY

TECHNIQUES AND INSTRUMENTATION IN ANALYTICAL CHEMISTRY – VOLUME 1

Evaluation and Optimization of Laboratory Methods and Analytical Procedures

A Survey of Statistical and Mathematical Techniques

Desiré L. Massart
Farmaceutisch Instituut, Vrije Universiteit Brussel

Auke Dijkstra
Laboratorium voor Analytische Chemie, Rijksuniversiteit Utrecht

Leonard Kaufman
Centrum voor Statistiek en Operationeel Onderzoek, Vrije Universiteit Brussel

with contributions by **Svante Wold**, **Bernard Vandeginste** and **Yvette Michotte**

ELSEVIER SCIENTIFIC PUBLISHING COMPANY
Amsterdam – Oxford – New York 1978

ELSEVIER SCIENTIFIC PUBLISHING COMPANY
335 Jan van Galenstraat
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the United States and Canada:

ELSEVIER/NORTH-HOLLAND INC.
52, Vanderbilt Avenue
New York, N.Y. 10017

First edition 1978
Second impression 1980

Library of Congress Cataloging in Publication Data

Massart, D L
Evaluation and optimization of laboratory methods and analytical procedures.

(Techniques and instrumentation in analytical chemistry ; v. 1)
Bibliography: p.
Includes index.
1. Chemistry, Analytic--Statistical methods. I. Dijkstra, Auke, joint author. II. Kaufman, Leonard, joint author. III. Title. IV. Series.
QD75.4.S8M37 543 78-12667
ISBN 0-444-41743-5

ISBN 0-444-41743-5 (Vol. 1)
ISBN 0-444-41744-3 (Series)

Printed in The Netherlands

Errata

Page 99, 2nd equation, 3426^2 should read 3436^2
Page 536, 2nd line, $2\pi\omega_0$ should read $\omega_0/2\pi$

3rd line, $\sum_{n=1}^{\omega_m/\omega_0}$ should read $\sum_{n=-\omega_m/\omega_0}^{\omega_m/\omega_0}$

Page 545, Fig. 27.1, e^{t/τ_x} (ordinate) should read e^{-t/τ_x}
Page 549, 4th line, 3% should read 1%

INTRODUCTION

One of the major problems with which the analytical chemist is confronted is how to make the best possible use of the large amount of information on analytical principles, methods and procedures that is available in the analytical literature. As stated by Laitinen and Harris (1975), an analytical chemist can be judged in part by his skill in the critical selection of methods. Almost daily the analytical chemist is confronted with problems of optimizing analytical procedures or related problems such as the selection of the best procedure for solving a given problem. This situation is represented by the most recent definitions of analytical chemistry, stating that analytical chemists have to produce qualified, relevant information on materials and processes in an optimal way (Gottschalk, 1972 ; Kaiser, 1974). It is therefore surprising to note that analytical chemists in general do not seem to have taken pains to develop strategies for optimization. Until recently it was common for the choice of the best procedure for a given analytical problem to be made largely intuitively and based upon experience.

Recently, a number of papers have appeared that express the concern of an increasing number of analytical chemists with this situation. It has been our intention to discuss in this book the formal methods that are available at present for the optimization and selection of analytical methods.

Before it is possible to make a selection or to carry out an optimization, one must have criteria according to which this may be done. Consequently, the performance of analytical procedures has to be evaluated by determining one or more performance characteristics of the procedure that is to be optimized or of the procedures from which the best one is to be selected. The set of criteria has to be defined for each problem and will include quantities such as precision, accuracy, limits of detection and interferences. Up to now, most of these criteria have been used in quantitative analysis, and it is probable that another set of characteristics will be required for qualitative analysis. Measures of information may become important in this respect and, therefore, a large section is devoted to information theory. Performance characteristics are discussed

in Part I.

The next two parts of the book are devoted to the optimization of procedures. One can discern different levels of optimization (Beveridge and Schechter, 1970), and for the purpose of this book we consider three such levels :

(1) Selection of one existing procedure from several alternatives. This is the simplest possible stage. There are several alternatives ; each of these is evaluated and the one which corresponds best to the exigencies of the application is selected. The main problem here is the evaluation of the procedures. This is described in Part I.

(2) Optimization of a procedure for which the outline is given. For example, given that the determination will be carried out colorimetrically with dithizone, select the optimal wavelength, the best concentration of dithizone, the pH, etc. This kind of problem usually consists in the selection of the optimal value of one or more continuously adjustable parameters. Occasionally, one may also have to include discrete parameters such as the kind of detector to be used. These problems are discussed in Part II.

(3) Optimization of combinations of analytical procedures or attributes of methods. There are many instances in which analytical procedures are combined, for instance

the combination of tests in a clinical laboratory to yield the optimal diagnostic or discriminatory power ;

the combination of GLC stationary phases to form a preferred set ;

the combination of elementary steps in a separation procedure to yield an optimal multicomponent separation scheme.

Such combinatorial problems are discussed in Part III. The analytical laboratory as a whole may be considered as a combination of methods, apparatus, etc. Some optimization problems concerning the functioning of an analytical laboratory are therefore also included in Part III.

Laitinen (1973), in an Editorial in *Analytical Chemistry*, stated that an analytical method is a means to an end, and not an end in itself. Analytical chemists tend to overlook this and sometimes develop methods that are more precise, faster, etc., while forgetting the intended applications. In such

instances, one can hardly speak of optimization. The analytical procedure must be chosen in relation to the goal and questions such as "how useful is it to increase the precision of a procedure used for a particular purpose" then arise. Some problems of this kind are discussed in Part IV.

Problems of optimization in analytical chemistry are often related to other optimization problems. However, such analogies will only be recognized if problems are formulated in a more or less formal and generalized way. The analytical procedure and the analytical laboratory should be considered from the point of view of systems theory. Some aspects of such a generalizing approach are discussed in Part V. In fact, one observes that analytical chemists concerned with optimization problems intuitively follow a systems approach. It would therefore have been more logical to start this book with a discussion of systems theory, but as yet it is not possible to construct a complete systems theoretical picture of the analytical procedure and the analytical laboratory, and to some extent the topic is of academical interest. Therefore, we have started the discussion with those points which clearly are of direct value in analytical practice.

The trend towards a more formal approach of the selection of analytical methods is not really new, but it has definitely grown stronger in the last few years. It is one of the principal concerns of the very recent field of chemometrics (Kowalski, 1975). It is not only felt among those who make this their research field in general analytical chemistry, but also by analytical chemists who are concerned more directly with analytical practice, such as clinical chemists and by those who need the results of analytical determinations, such as physicians using laboratory tests for medical diagnosis. At about the same time that concepts such as information were introduced into general analytical chemistry, clinical chemists began to use multivariate data analysis techniques to investigate which analytical methods yield the most diagnostic information. If one looks at the literature cited by the "generalists" and the clinical chemists, one finds that they cite different literature and that,

in general, there seems to be very little communication between the two groups.

In some applications, formal methods for the investigation of the performance of analytical methods were introduced many years ago. This is the case, for example, with official analytical chemists, who have developed methods for the evaluation of errors likely to occur in analytical procedures. Many analytical chemists from other specialities, but who are also concerned with the evaluation of analytical methods, seem, however, to ignore the existence of such methods.

We have tried to combine the knowledge stored in these (and other) different specialities in the hope of stimulating a more systematic application of formal selection methods in analytical chemistry. Our first idea was to limit this book to newer methods or concepts, such as information theory and operational research, but it was soon clear that it would be meaningless to try and make a synthesis and not include classical statistical concepts. Therefore, a number of chapters on classical statistical methods were added. Because, on the other hand, we did not want to duplicate the material already available in several books on statistics in chemical analysis, we have tried to eliminate statistical methods designed for the evaluation of results rather than of procedures, and we have not tried to discuss the subject exhaustively. This is also true for all other chapters. More specialized knowledge should be sought in the original literature or specialized books and chapters to which we refer. Since this book was written to introduce a number of formal optimization techniques to analytical chemists and not for specialists, we have not tried to cover the existing literature exhaustively. Instead, we usually have given some references to books, review articles and a few illustrative articles.

In writing this book, we have started from the belief that some of the newer mathematical methods or theories, such as pattern recognition, information theory, operational research, etc., are relevant to some of the basic aims of analytical chemistry, such as the evaluation, optimization, selection, classification, combination and assignment of procedures or sub-procedures - in short all those processes that intervene in determining exactly which analytical procedure or

programme should be used. Unfortunately, most chemists are daunted by the task of learning how these mathematical methods function and this is not helped by the difficulty of establishing a link between the formal mathematics found in most books on this subject and analytical problems.

We have tried to treat the mathematical topics as lucidly as possible and to illustrate the text with examples, in some instances abandoning a rigorous mathematical treatment. However, we have tried to compensate for this by including a series of mathematical sections. The level of the mathematics is higher in Chapters such as 2, 3 and 4, where the subject treated is not completely unfamiliar to most analytical chemists. In those chapters where the subject matter is probably new to most analytical chemists, only the most elementary explanations are given, often in words, because we think it more important to emphasize the underlying philosophy than to explain the mathematics. In doing so, we hope we have removed the barriers of applying formal methods to optimization problems in analytical chemistry. One major difficulty encountered when writing this book was the mathematical symbolism. We have tried to present a coherent set of symbols throughout the book but, because of the diversity of the methods described, we have not been entirely succesful in this respect. Nevertheless, we think that the symbols used should be sufficiently clear.

Parts of the draft of this book were read by E. Defrise-Gussenhoven (Vrije Universiteit Brussel ; Chapters 2 and 3), H.C. Smit (University of Amsterdam ; Chapter 10), D. Coomans (Vrije Universiteit Brussel ; Chapter 20), G. Kateman (Catholic University of Nijmegen ; Chapters 26 and 27) and P.M.E.M. van der Grinten (DSM, Heerlen ; Chapters 26 and 27). Their suggestions were very valuable. The authors are also indebted to S. Peeters, who typed the final text, A. Langlet-De Schrijver and L. Maes, who prepared the figures, and C. Uytterhaegen-Hendrickx, G. De Boeck, A. Van Gend, F. Gheys and M. Segers-Geeroms who typed the manuscript. The research leading to the idea of writing this book was helped financially by organizations such as the Fonds voor Geneeskundig Wetenschappelijk Onderzoek and the Fonds voor Kollektief en Fundamenteel Onderzoek and was stimulated by collaboration with members from the Studiegroep voor

Laboratoriumoptimalisering and the Centrum voor Statistiek en Operationeel Onderzoek of the Vrije Universiteit Brussel, and with the following students, who obtained degrees based on research on subjects covered in this book : H. De Clercq, M. Detaevernier, J. Smeyers-Verbeke, J.H.W. Bruins Slot, P.F. Dupuis, G. van Marlen, P. Cley, T. Koppen and A. Eskes. The proofs were read by A. Kaufman. The authors express their thanks to all of these persons and organizations.

REFERENCES

G.S.G. Beveridge and R.S. Schechter, Optimization : Theory and Practice, McGraw-Hill, New York, 1970.
G. Gottschalk, Z. anal. Chem., 258 (1972) 1.
R. Kaiser, Z. anal. Chem., 272 (1974) 186.
B.R. Kowalski, J. Chem. Inf. Computer Sci., 15 (1975) 201.
H.A. Laitinen, Anal. Chem., 45 (1973) 1585.
H.A. Laitinen and W.E. Harris, Chemical Analysis, McGraw-Hill, New York, 1975.

CONTENTS

PART I

EVALUATION OF THE PERFORMANCE OF ANALYTICAL PROCEDURES