



SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

33 SOUTH 17TH STREET • PHILADELPHIA, PA. 19103 • (215) LOcust 4-2930

BOARD OF DIRECTORS

DONALD L. THOMSEN, JR.
PRESIDENT AND EXECUTIVE DIRECTOR
REPLY TO: 97 PARISH ROAD SOUTH
NEW CANAAN, CONNECTICUT 06840
(203) 966-1008

JOEL E. COHEN
ROCKEFELLER UNIVERSITY

WILLIAM T. GOLDEN
NEW YORK, NEW YORK

HOWARD W. JOHNSON
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

HERBERT B. KELLER
CALIFORNIA INSTITUTE
OF TECHNOLOGY

BROCKWAY McMILLAN
BELL TELEPHONE LABORATORIES

VAUN A. NEWELL
EXXON CORPORATION

EMANUEL R. PIORE
ROCKEFELLER UNIVERSITY

MINA REES
CITY UNIVERSITY OF
NEW YORK

DONALD L. THOMSEN, JR.
SIAM INSTITUTE FOR MATHEMATICS
AND SOCIETY

F. JOACHIM WEYL
HUNTER COLLEGE, CITY
UNIVERSITY OF NEW YORK

MASTER

NEWS :-

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

SIMS STUDY

Statistics and Environmental Factors in Health

Members of the Study:

SIMS (1974)

Stanford University (1974)

Columbia University (1975)

427
attached

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Volume 1 No. 1

October 1976

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

SIMS THREE YEAR STUDY ON

STATISTICS AND ENVIRONMENTAL FACTORS IN HEALTH

Members of the Study:

SIMS

97 Parish Road South
New Canaan, Connecticut 06840
Phone: (203) 966-1008

Columbia University
School of Public Health
Division of Biostatistics
600 West 168th Street
New York, New York 10032
Phone: (212) 694-3936

Stanford University
Department of Statistics
Stanford, California 94305
Phone: (415) 497-2879

* * *

The Study is supported by:

Alfred P. Sloan Foundation
Energy Research and Development Administration(ERDA)
National Science Foundation(NSF)
Rockefeller Foundation

This "NEWS" is published by SIMS, the SIAM Institute for Mathematics and Society, 97 Parish Road South, New Canaan, Connecticut 06840.
All rights reserved.

TABLE OF CONTENTS

	Page
Editorial Overview of the Study	1
Communication from Columbia	3
Communications from Stanford	3
Technical Reports	5
SIMS Seminars at Stanford	7
Presentations at Professional Conferences and Meetings	8
Participants in the Study	9
SIMS 1976 Research Application Conference on Environmental Health	10
SIAM/SIMS Publications: Ecosystems, Epidemiology, Energy(advt.)	11

Editorial Overview of the Study

With these first pages telling about the SIMS Study on Statistics and Environmental Factors in Health (SEFH), it seems appropriate to provide some introduction and background. To those unfamiliar with SIMS, it may be of interest to say a few words first about SIMS, the SIAM Institute for Mathematics and Society. SIMS was formally established in January 1973 by SIAM, the Society for Industrial and Applied Mathematics, with its own articles of incorporation and its own Board of Directors. (In legal terminology it is like unto a wholly owned subsidiary of SIAM.) SIMS' principal objectives are to: develop, promote, support and maintain research and foster education in the application of mathematics to societal problems.

One of SIMS' principal activities is its Transplant Program, whereby talented mathematicians are "transplanted" into interdisciplinary centers for two years; mathematicians participating in this Program become deeply involved in pressing societal problems, and to those problems they are expected to contribute solutions which without their involvement might otherwise not have been obtained. In this ongoing Program, four transplants have taken place thus far; the societal fields have been population growth, urban analysis, environmental health, and energy; support for this Program has been through grants to SIMS from the National Science Foundation(NSF), the Office of Naval Research, and the Alfred P. Sloan Foundation.

Another of SIMS' continuing programs has been a series of Research Application Conferences (RAC's) on topics of current interest; each conference lasts for five days; so far four such conferences have been held, with the principal subjects being ecosystems (1974), epidemiology (1974), energy (1975), and environmental health (1976); Proceedings have appeared for the first three and is in preparation for the fourth. The Conferences on ecosystems, energy, and environmental health were supported by NSF; the Conference on epidemiology was supported jointly by NSF and the Department of Health Education and Welfare, Bureau of Health Services Research.

The Study has been undertaken in response to an often expressed need for more mathematical involvement in the field of environmental health. The Study will focus on some of the most difficult problems associated with environmental health and will either use existing or develop new statistical methodologies to aid in solution of those problems. The Study consists of research and its applications, graduate education, and dissemination of results to interested investigators in the field of environmental health. The Study will last for three years.

After a feasibility study during 1975, the Study took its present form with participation by Columbia and Stanford Universities. Associated with the Study for Columbia are the Division of Biostatistics and Division of Epidemiology, Columbia School of Public Health, and the Department of Mathematical Statistics, Columbia University. At Stanford, the center is composed of members of the Department of Statistics. Stanford was a member of the Study during the feasibility study period during 1975; Columbia became a member of the Study late in 1975. Advising the members of the Study is a special Review Panel appointed by SIMS. Individual members of the Study at Columbia and Stanford and members of the Review Panel are noted on page 9.

The 1975 feasibility study was supported by grants to SIMS from IBM, the National Science Foundation (NSF), Research Corporation, and the Rockefeller Foundation. The three year Study is being supported initially by the Energy Research and Development Administration (ERDA), NSF, the Rockefeller Foundation, and the Alfred P. Sloan Foundation; requests for support to two other government agencies are currently in process.

The objectives of this news publication are to: (1) call attention to activities associated with the Study, such as reports and papers published, seminars and courses conducted, and presentations made at professional meetings; (2) recognize faculty and graduate research assistants participating; and (3) report from time to time any other newsworthy events which may be associated with the Study (such as the report on the 1976 SIMS Research Application Conference on Environmental Health in this issue).

Communication it is hoped will become a significant part of the Study; it is encouraged on the part of all those who may read these pages.

D. L. Thomsen, Jr.
President, SIMS

COMMUNICATION FROM COLUMBIA

Work is progressing on finding new ways to analyze the data on air pollution and health for New York City. Two particular areas have received special attention. The first is the study of how well the city's air pollution as a whole can be characterized by the 40 station aerometric network. Numerical work on correlation of various transforms of the data is being pursued as is principal components analysis. In July of this year, Dr. Inge Goldstein gave talks at Stanford University and at the Health Effects Section of the Air Pollution Control Association Meeting in Portland, Oregon on this subject.

The other area of study concerns the health data, which consists of emergency room visits at four hospitals in New York City. A doubly stochastic model for the daily number of visits is being considered, and was the subject of the paper given by Dr. Jack Cuzick at the SIMS Conference on Environmental Health held on July 5-9 at Alta, Utah. Dr. Cuzick also gave a talk, "On the number of crossings of a pollution threshold by a Gaussian process," at Stanford University in July.

* * *

COMMUNICATIONS FROM STANFORD

Pierre Delfiner of the Center for Geostatistics, Fontainebleau,
France
Visitor, Summer 1976 at Stanford University
Department of Statistics.

Summary of Summer Research, 1976

The method of "Kriging" developed by Professor G. Matheron in France provides an efficient way to extend the pollution information collected on a local basis (the monitoring stations) to the whole area of interest. The method employs an estimate of the mean-squared difference between two readings as a function of the spatial vector which separates those readings; this function is called the "variogram." Any linear functional of the data can be estimated, together with the variance of the estimation error. Models of different complexity are used according to whether the phenomenon shows a drift or not. Altogether the strength of the approach derives from the fact that it only considers increments of the data, viewed as part of a realization of a random field. Those increments filter out the mean of the process which, therefore, does not have to be known.

COMMUNICATIONS FROM STANFORD(continued)

Ray E. Faith
Department of Statistics
Stanford University

Ray E. Faith of Stanford during the past summer has prepared a report on the spatial correlation of oxidant concentrations in the San Francisco Bay area (as presented at the SIMS Research Application Conference on Environmental Health at Alta, Utah reported on page 10).

In addition, Dr. Faith has written programs for the Wang computer (1) to plot isopleths using the Kriging method (see above under Delfiner), (2) to compute the variogram for station means, and (3) to estimate a true mean from the censored mean.

* * *

Paul Switzer
Department of Statistics
Stanford University

Paul Switzer has examined the problem of estimating the area proportion of an air basin over which some pollutant exceeds a threshold concentration. In particular, how should such estimates depend on the geographic arrangement of measuring stations and the given threshold concentration? One answer is provided by regarding the pollutant concentration as a random function of latitude and longitude and then finding estimates which give small average errors. We may then see which properties of the random function are important and which are unimportant in the construction of good estimates. Finally, we may ask how the important random function properties may be approximated from available station data.

His current work concerns the robustness to spatial nonstationarity of statistical procedures derived from stationarity assumptions. For example, nonstationarity of most pollutant concentrations is obvious from station data, however, it is generally very convenient to use stationarity derived contouring algorithms, nevertheless. Such algorithms use estimated spatial autocorrelations, but if these autocorrelations are deliberately overestimated using a spatially constant mean, then we may effectively mimic a much more complicated nonstationary analysis.

TECHNICAL REPORTS

Thomas W. Sager
Department of Statistics
Stanford University

Technical Report No. 1 (July 15, 1976)

RELATING SPATIAL DISTRIBUTIONS OF POLLUTANTS TO HEALTH EFFECTS

SUMMARY

A new statistical tool for epidemiology is introduced and some of its elementary properties are considered. Starting with the contours of a spatially distributed explanatory variable (air pollution), the method produced a relationship between the explanatory variable(s) and the response variable (population-adjusted health effects) by accumulating the response within successively wider contours. The method has some of the flavor of regression in that it reduces the relationship between two spatial distributions to a single easily interpreted graph. It effectively utilized knowledge of the geographic location of data before discarding the location coordinates as nuisance parameters.

* * *

Ruth Sheshinski, Graduate Student
Department of Statistics
Stanford University

Technical Report No. 2 (July 22, 1976)
Technical Report No. 4 (August 16, 1976)

ROBUSTNESS TO MISSPECIFICATIONS OF CORRELATION MODELS FOR THE
PROBLEM OF INTERPOLATING AIR POLLUTION MAPS

SUMMARY

This first report concerns the problem of interpolation of pollutant concentrations between air monitoring stations. Knowing that the station values are not statistically independent, it seems important to study the precision of interpolation with respect to different assumptions for correlation models between stations.

A number of examples indicate that weighted averages of station values are robust to misspecifications of correlation models.

Report No. 4 is a continuation of Report No. 2.

We compare mean squared errors of interpolation estimators based on a correctly and incorrectly specified constant mean

TECHNICAL REPORTS(continued)

with estimators based on an unspecified constant mean. The ratios of the mean squared errors decrease with decreasing correlation between station observations, and are, in general, very close to one.

We also compare interpolations using a specified linear trend function versus an unspecified linear trend function, and the use of biased interpolation estimators versus universally unbiased estimators in the presence of a linear trend. In general, the biased and unbiased estimators have similar performance, except for very steep linear trends.

* * *

Yashaswini Mittal
Department of Statistics
Stanford University

Technical Report No. 3 (August 16, 1976) and Working Paper No. 1

MAXIMA OF PARTIAL SAMPLES

SUMMARY

Air pollution standards are usually set in terms of maximum concentrations of pollutants. Various practical difficulties due to the cost of sampling or due to missing data at times result in only a partial sample being available for making inferences. The methodology developed in this paper addresses the way the difference of two maxima behave and can be used to glean information about the actual maximum concentration of a pollutant when information is only available with respect to the maximum of a partial sample. The results were applied to oxidant data from nine Bay Area stations.

* * *

(N.B. Copies of these technical reports and working papers may be obtained by contacting the Statistics Department at Stanford University at the address given on the inside of the front cover.)

SIMS Seminars at Stanford

Dr. Jerry Wesolowski, California Department of Public Health
"Measurement of ambient air pollutants"
May 26, 1976.

Professor Martin Morf
"Fast algorithms for least squares estimation and time
series modelling"
June 25, 1976.

Professor June Juritz, University of Capetown
"Some suggestions for assessing the effect of weather on
air pollution"
July 2, 1976.

Professor Thomas W. Sager, Stanford University
"Relating spatial distributions of pollutants to health
effects"
July 9, 1976.

Professor Jack Cuzick, Columbia University
"Crossings of pollutant thresholds in the nonstationary
case"
July 16, 1976.

Professor Julian Peto, Oxford University
"Asbestos dust - related disease and standards settings"
July 19, 1976.

Professor Inge Goldstein, Columbia University
"Use of aerometric data for epidemiologic studies"
July 23, 1976

Professor Al Miller, San Jose State University
"Problems of air pollution prediction in the San Francisco
Bay area"
July 30, 1976.

Pierre Delfiner, Center for Geostatistics, Fontainebleau, France
"Geostatistical estimation and modeling of spatial variables
and applications"
August 6, 1976.

A. Norris, L. H. Robinson and J. S. Sandberg,
Bay Area Air Pollution Control District
"Detecting trends in air quality data and meteorological
data"
October 1, 1976.

Presentations at Professional Conferences and Meetings

- (1) Thomas W. Sager (Stanford)
"Relating spatial distributions of pollutants to health effects"
(invited), 9th International Biometric Conference,
Boston, August 26, 1976.
- (2) Ray E. Faith (Stanford)
"A statistical analysis of oxidant and carbon monoxide levels monitored in the San Francisco air basin 1971-1973: a comparison of certain spacial/temporal models for interpolating between stations"
SIMS Research Application Conference on Environmental Health, Alta, Utah, July 5-9, 1976.
- (3) Lincoln E. Moses (Stanford)
"Some thoughts about assessing adequacy of the fineness of a grid of monitoring stations"
SIMS Research Application Conference on Environmental Health, Alta, Utah, July 5-9, 1976.
- (4) Jack M. Cuzick (Columbia)
"Use of doubly stochastic processes in estimating health effects due to environmental factors"
SIMS Research Application Conference on Environmental Health, Alta, Utah, July 5-9, 1976.
- (5) Persi Diaconis (Stanford)
"Analysis of air pollution data"
AAAS Annual Meeting, Boston, February 22, 1976.
- (6) Paul Switzer (Stanford)
"Time and space variation of pollutant concentrations"
AAAS Annual Meeting, Boston, February 22, 1976.

PARTICIPANTS in the STUDY

COLUMBIA UNIVERSITY

Division of Biostatistics,
School of Public Health

Joseph L. Fleiss
Inge F. Goldstein
Sylvan R. Wallenstein

Graduate Research Assistants

Neng-Rong Lee(Robbins)*
Lawrence Rausch(Susser**)

Mathematical Statistics Department

Jack M. Cuzick
Herbert E. Robbins
Burton H. Singer

Consultant

Leon F. Landovitz(City University of New York)

STANFORD UNIVERSITY

Statistics Department

Persi Diaconis
Ray E. Faith
M. Vernon Johns, Jr.
Yashaswini D. Mittal
Lincoln E. Moses
Ingram Olkin
Thomas W. Sager
Paul Switzer

Graduate Research Assistants

Dick DeVeaux (Diaconis)*
Cyril Grivet (Switzer)
Bonnie Hole (Diaconis)
Ed Korn (Faith)
Jeannette Prietsch (Sager)
Ruth Sheshinski (Switzer)
Neil Willits (Mittal)
Suzanna Wong (Moses)
Chen-hui Foo (Switzer)

Visitors

Pierre Delfiner(Center for Geostatistics,
Fontainebleau, France)
Alice Whittemore(City University of New York)***
Dag Tjøstheim, NORSAR, Norway

SIMS

Review Panel

William G. Cochran(Harvard University)
Mark Kac(Rockefeller University)
Vaun A. Newill(Exxon Corporation)
Donald L. Thomsen, Jr., ex officio(SIMS)
John R. Totter(Oak Ridge National Laboratory, ERDA)
James L. Whittenberger(Harvard School of Public Health)

* ()'s indicate faculty member with whom graduate research assistant is working or by whom sponsored

** Dr. Mervyn Susser; Professor of Epidemiology

*** Under support of a Rockefeller Foundation Fellowship