REPORT DOCUMENTATION PAGE

(2)

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): ARO 22406.2-MA
6a. NAME OF PERFORMING ORGANIZATION: Rice University; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: U. S. Army Research Office
6c. ADDRESS (City, State, and ZIP Code): P.O. Box 1892, Houston, TX 77251	7b. ADDRESS (City, State, and ZIP Code): P. O. Box 12211, Research Triangle Park, NC 27709-2211
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U. S. Army Research Office; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: DAAG29-85-K-0212
8c. ADDRESS (City, State, and ZIP Code): P. O. Box 12211, Research Triangle Park, NC 27709-2211	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

AD-A200 846

11. TITLE (Include Security Classification)

Some Problems in Data Based Computer Graphics & Simulation

12. PERSONAL AUTHOR(S)
James R. Thompson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 8/1/85 TO 7/31/88	October 21, 1988	11

16. SUPPLEMENTARY NOTATION
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
SIMDAT, SIMEST, simulation, density estimation, time series, statistical computing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This has been a three year study in the investigation of topics in the general area of statistical computing. We have developed, jointly with Dr. Malcolm Taylor of BRL the SIMDAT algorithm for using a set of projectile data to generate other data sets which could have happened. Also, we have developed a simulation based technique, SIMEST, to handle problems not tractable via the closed form maximum likelihood approach. We have obtained automated procedures for selecting bandwidths in the kernel density estimation. We have developed graphical techniques for demonstrating density estimates in dimensions as large as 6. We have done extensive work to build algorithms whose function is the creation of multicolored graphical displays of density estimates. We have done work in parameter estimation for time series problems using both frequency based and time based approaches.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
James R. Thompson	713 527 4828	

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

Statement of Problem

This has been a three year study in the investigation of topics in the general area of statistical computing. For example, joint work with Dr. Malcolm Taylor at BRL has led to the SIMDAT algorithm for using a set of projectile data to generate other data sets which could have happened. Also, we have developed a simulation based technique, SIMEST, to handle problems not tractable via the closed form maximum likelihood approach. This technique, whose applicability we have amply demonstrated on oncological models where classical procedures had stalled, is now being examined in econometric contexts. We have also developed epidemiological models of AIDS and discovered that the epidemic's sudden appearance is very likely due to the enhancement factor provided by bathhouses. We have built our earlier work in density estimation to give automated procedures for selecting bandwidths in the kernel estimation approach. We have developed graphical techniques for demonstrating density estimates in dimensions as large as 6. We have done extensive work to build algorithms whose function is the creation of multicolored graphical displays of density estimates. We have done work in parameter estimation for time series problems using both frequency based and time based approaches. We have also done work for using in the general smoothing of physical processes.

Summary of Work

July 1, 1985-December 31, 1985

The work of Chiu has been concerned with the problem of estimating the variable time delays of a signal arriving at an array of sensors. A procedure to estimate the parameters of a linear time delay model has been developed. The procedure compares the Fourier transforms at different frequencies (thereby taking the Doppler effect into consideration). Under regularity conditions, the estimate obtained is shown to be consistent and asymptotically normal. Simulations were carried out and the results were found to agree well with the theoretical results. The procedure was applied to the records of the Imperial Valley earthquake of October 15, 1979 as recorded by the El Centro differential array.

Also, Chiu has considered the problem of smoothing noisy images. His procedure preserves some important characters of the images, such as edges and lines. The smoothed images can then be used to detect and locate the edges of the images.

Scott has continued his investigations of practical and theoretical issues in nonparametric density estimation. He has introduced new powerful density estimation algorithms for multivariate estimation. He has examined the theoretical asymptotics of several important density estimators. Algorithms for automatically calibrating these density estimators have been introduced and examined using asymptotic analysis and Monte Carlo techniques. Practical experiences with these methods have been reported in several papers and a chapter of a book.

Thompson has been concerned with the development of simulation based algorithms for parameter estimation in applied stochastic process models. Also, he has continued the application of such models to oncological situations. He was one of four statisticians selected to overview recommendations of EPA scientists to institute draconian controls on sulfur dioxide emissions from factories in Ohio and other states; this review found that such a move was contraindicated by the facts; the proposed controls were shelved. An historical analysis of the development of quality control in the United States has been used as a means of analyzing current difficulties and proposing solutions.

January 1, 1986 - June 30, 1986

The doctoral dissertation of Husseman demonstrates the use of variable tile width histograms for the nonparametric estimation of two dimensional densities. It is shown how the procedure can be used to obtain robust estimates in a variety of situations.

Scott has examined several classical cross validation techniques for density estimation. An integrated and highly automated generalization of earlier techniques is proposed.

A procedure for using limited data to generate simulated outcomes, SIMDAT, in a tank battle is proposed and demonstrated by Taylor and Thompson. The procedure represents a significant improvement over the bootstrap in many situations.

The book *Cancer Modeling* is the first to bring modern statistical computer based modeling and simulation techniques together for use in oncological situations.

The SIMEST algorithm of Thompson *et al* enables the user to go directly from micro-axioms to parameter estimation without attempting to obtain a "closed form."

July 1, 1986 - December 31, 1986

The work of Chiu was concerned mainly with improved techniques for parameter estimation of time series in the frequency domain. Additional work was also carried out for the design of tests for the detection of periodic components in a white noise time series.

Scott prepared 200 pages of handouts for his shortcourse (given jointly with Bernard Silverman of the University of Bath) in nonparametric density estimation at the national meetings of the American Statistical Association. He also examined the possibility of the construction of expert systems for the design of smoothing parameters of nonparametric density estimators.

Thompson wrote a book for his shortcourse at the Army Design of Experiments Conference at Monterey, which was distributed to the participants in spiral bound form. Demonstration software for the shortcourse was programmed by Kauffman and MacIntosh diskettes made available to the participants. Thompson also created an epidemiological model for AIDS, which arrived at the conclusion that the disease, in the United States, was fragile and dependent on the presence of a highly promiscuous subgroup of the homosexual population. The model presented predictions for the course of the disease.

January 1, 1987 - June 30, 1987

Kauffman has been working on a dissertation (Thompson, thesis advisor) on simulation based techniques for parameter estimation in stochastic processes.

Chiu has been working on procedures to determine the existence of signal sinuosoidal waves embedded in a white series. This problem is important in signal processing the records of an array of sensors (spatial-temporal processes) when one is interested in detecting signals propagating through the array. He is also extending his one dimensional procedure for smoothing images to the two dimensional case.

Scott is pursuing his work in graphical based techniques for the estimation of densities in dimensions of three and greater. He continues to work closely with the DOD and NASA in the analysis of real world data sets.

Thompson is working on density estimation in the nongraphical mode. He is also working on simulation based techniques for parameter estimation. His October Army shortcourse notes in model building have been accepted by John Wiley and Sons as

the basis for a book; he is currently at work on putting the notes into book form.

July 1, 1987 - December 31, 1987

Kauffman continues his work on a dissertation extending the SIMEST algorithm of Thompson *et al* for parameter estimation in applied stochastic processes without the necessity of finding closed form solutions.

Chiu has been extending his smoothing procedure for two dimensional images. Also, he has been working on procedures to determine the existence of signal sinuosoidal waves embedded in a white series. This problem is important in signal processing the records of an array of sensors (spatial-temporal processes). He is also extending his one dimensional procedure for smoothing images to the two dimensional case.

Scott is extending his average shifted histogram technique for the estimation of densities of higher dimension. He also is concerned with cross-validation procedures for his density estimation algorithms. He works closely with the DOD and NASA in the analysis of real world data sets.

Thompson has edited a book *Cancer Modeling* with Barry Brown in which modern data analytical procedures are developed for dealing with the problem of cancer origination, progression and control. The SIMEST algorithm is presented in that book. Thompson has written a paper which argues that the current AIDS epidemic would not have occurred if public health authorities had not permitted the licensing of gay bath houses in major American cities. Thompson's 1986 Army shortcourse notes in model building have been accepted by John Wiley and Sons as the basis for a book; it is hoped that the final manuscript can be given to the publisher shortly.

January 1, 1988 - June 30, 1988

Kauffman is extending the SIMEST algorithm of Thompson *et al* for parameter estimation in applied stochastic processes without the necessity of finding closed form solutions. The extension, which deals with the inclusion of covariate data, will be Kauffman's doctoral dissertation.

Chiu continues his frequency domain approach to time series analysis. He has been dealing with smoothing (bandwidth) considerations. He is also examining frequency domain based techniques for parameter estimation.

Scott has been engaged in an collaborative investigation with Dr. Wolfgang Haerdle of the University of Bonn. This has involved visits by Scott to Bonn and by Haerdle to Rice. The work involved has included topics in robust regression and in multivariate nonparametric density estimation. The utilization of these techniques in mathematical economics has been a major thrust of their work.

Thompson has been examining multiresponse variable versions of the SIMEST algorithms. The copy editing step of the publication of Thompson's book *Empirical Model Building* has been completed by John Wiley & Sons. The target date for publication is the end of 1988. Thompson has consulted with scientists at Fort Ord on the

subject of data based war gaming.

Publications

(1) Atkinson, E.Neely, Brown, Barry W. and Thompson, James R. (1988). "SIMEST and SIMDAT: differences and convergences," to appear in *Proceedings of the Twentieth Interface Conference on Statistics and Computing Science*, Wegman, Ed., Amsterdam: North Holland.

(2) Chiu, S.T.,"Smoothing Noisy Images", *Proceedings of the 1985 Conference on Applied Analysis (Mathematics & Statistics) in Areospace, Industry and Medical Sciences*, 131 - 136.

(3) Chiu, S.T.,"Estimation of the speed of a moving source from time delay estimates," IEEE, *Trans Acoustics, Speech, & Signal Processing, Vol. Assp-35, 1987, pp. 696-697.*

(4) Chiu,S.T., "Statistical Estimation of the Parameters of a Moving Source from Array Data,"*Annals of Statistics*, ,1986, pp. 559-578.

(5) Chiu,S.T., "A test for periodic components in white time series," submitted to *JASA*.

(6) Chiu, Shean-Tsong, "Robust estimation of parameters of the spectrum of a time series containing periodic components," submitted.

(7) Chiu, Shean-Tsong, "Estimating the parameters of the noise spectrum for a time series," submitted.

(8) Chiu, Shean-Tsong, "A linear estimation procedure of the parameters of autoregressive-moving average processes," submitted.

(9) Chiu, Shean-Tsong, "Bandwith selection for nonparametric regression: a Fourier analysis approach." submitted.

(10) Chiu, S.T., "Weighted least squares estimators on the frequency domain for the parameters of a time series," *Annals of Statistics*, , 1988, Vol 16, pp. 1315-1326.

(11) Haerdle, Wolfgang and Scott, David W., "Smoothing in low and high dimensions by weighted averaging using rounded points," submitted.

(12) Scott, D.W., "Experiences with Examining Large Multivariate Data Sets with Graphical Nonparametric Methods," *Proceedings of the Third Symposium on Mathematical Pattern Recognition and Image Analysis*, 1985, in press.

(13) Scott, D.W., "Choosing Smoothing Parameters for Density Estimators." To appear in *Proceedings of the 17th Symposium on the Interface of Computer Science and Statistics*, North-Holland.

(14) Scott, D.W., "Handouts for the ASA short course in density estimation."

(15) Scott,D.W., " Comment on 'How far are automatically chosen regression smoothing parameters from their optimum' by Hardle, Hall and Marron," to appear in JASA.

(16) Scott, D.W. and Terrell, G.R. "Biased and unbiased cross-validation in density estimation," *J. American Statistical Association*, 82:1131-1146, December, 1987.

(17) Scott, D.W. "A Note on choice of bivariate histogram bin shape." Submitted to *J. of Official Statistics*, Sept. 1987, Rice Technical Report No. 85-311-3.

(18) Scott, David W. and Schmidt, Heinz-Peter, "Calibrating histograms with applications to economic data," accepted by *Empirical Economics.*

(19) Scott, David W. (1988). "Comment on paper by Haerdle, Hall and Marron," in *J. American Statistical Association,*, **83**, pp. 96-98.

(20) Taylor, Malcolm and Thompson, James R., "A data based algorithm for the generation of random vectors," in *Computational Statistics and Data Analysis*,1986, pp. 93-101.

(21) Thompson, J.R.(with M. Landau, J. Van Ryzin and B. Swindel), "A FORAST Review Report" *American Statistician*, (1985) v.39, no.4, 250-254.

(22) Thompson, J.R. (with E.Atkinson and B. Brown), "Simulation versus numerical approximation" in *Proceedings of the 45th Session of the International Statistical Institute* (1985) 503-4.

(23) Thompson, J.R., "American quality control: What went wrong? What can we do to fix it?" in *Proceedings of the 1985 Conference on Applied Analysis (Mathematics and Statistics) in Aerospace, Industry and Medical Sciences*, R. Chhikara, ed., University of Houston: Houston, 247-255

(24) Thompson, J.R. (with R. Alexanian, B. Brown, B. Drewinko, J. Hokanson and B. Jannson), "The correlated resistance model for human myeloma," 1985, *The Journal of Tumor and Cell Kinetics*, 1985. V. 19, pp. 501-510.

(25) Thompson, J.R., *Density Estimation, Modeling and Simulation: Studies in Empirical Model Building*,, lecture notes for the shortcourse during the 1986 Conference on the Design of Experiments in Army Research Development and Testing.

(26) Thompson, J.R."AIDS: the mismanagement of an epidemic," to appear in *Mathematical Population Dynamics.*

(27) Thompson, J.R., Atkinson, E.N., and Brown, B.W. "SIMEST: an algorithm for simulation-based estimation of parameters characterizing a stochastic process," in *Cancer Modeling*, Thompson and Brown, eds., New York: Marcel Dekker, pp. 387-415, 1987.

(28) Thompson, J.R. and Brown, B.W. *Cancer Modeling*, New York: Marcel Dekker, 422 pages, 1987.

(29) Thompson, James R. (1988). "Modular wargaming," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 182-206.

(30) Thompson, James R.(1988). "Predation and immune response systems," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 207-213.

(31) Thompson, James R.(1988)."Pyramid clubs for fun and profit," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 214-217.

(32) Thompson, James R.(1988)."A model based examination of AIDS: its causes and likely progression," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 218-229.

(33) Thompson, James R., (1988)"A glimpse at exploratory data analysis,"in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 230-251.

(34) Thompson, James R.(1988)."Nonparametric density estimation," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research

Development and Testing, U.S. Army Research Office, pp. 252-267.

(35) Thompson, James R. (1988)."Stein's Paradox," in Proceedings of the Thirty-Third Conference on the Design of Experiments in Army Research Development and Testing, U.S. Army Research Office, pp. 268-275.

(36) Thompson, James R. (1989). *Empirical Model Building*, John Wiley and Sons, 244 pages.

PERSONNEL SUPPORTED

Chiu, Shean-Tsong: Assistant Professor of Statistics

Ensor, Katherine Bennett: Assistant Professor of Statistics

Husemann, Joyce Ann: Graduate Student--awarded Ph.D. in May of 1986.

Kauffman, Tom: Graduate Student

Scott, David W.: Professor of Statistics

Thompson, James R.: Professor of Statistics

Wang, Ferdinand: Graduate Student

Wiener, Ronald: Graduate Student