

George Ostrouchov

Statistics and Data Sciences Group
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Bethel Valley Road
P. O. Box 2008, Bldg 6012
Oak Ridge, Tennessee 37831-6367

Tel: (865) 574-3137
Fax: (865) 574-0680
ost at ornl dot gov
<http://www.csm.ornl.gov/~ost>

Education

Ph.D.	Statistics	1984	Iowa State University
M.Sc.	Statistics	1981	Iowa State University
B.Math., Honours Co-op	Mathematics/Statistics	1978	University of Waterloo

Appointments

Since 1983	Senior Research Staff Member (and previous titles), Statistics and Data Sciences Group, Computer Science and Mathematics Division, ORNL
Since 2003	Adjunct Professor of Statistics, University of Tennessee, Knoxville
1994-1996	Adjunct Faculty, Great Lakes Colleges Association
1983-1983	Instructor, Department of Statistics, Iowa State University
1981-1983	Free-lance software consultant, Computing Center, Iowa State University
1979-1983	Research Assistant, Department of Statistics, Iowa State University
1978-1979	Statistician/Analyst, Informetrica Ltd., Ottawa, Canada

Areas of Interest

Computational aspects of statistical methods; analysis and visualization of massive and distributed data sets; large scale and multivariate modeling; graph theoretic aspects of statistical modeling; matrix and sparse-matrix computation; spatial statistical methods.

Memberships and Honors

American Statistical Association (Section on Statistical Computing, Section on Physical and Engineering Sciences), International Association for Statistical Computing, Society for Industrial and Applied Mathematics, Mensa, Phi Kappa Phi, Mu Sigma Rho

Professional Activities and Recognitions

Organizer and Chair: Minisymposium on Matrix Computations in Statistics at The Third SIAM Conference on Applied Linear Algebra, Madison, WI, 1988.

Associate Editor: Journal of Statistical Computation and Simulation, 1988–1994.

Award: Martin Marietta Energy Systems government-use invention award for “Cost Matrix Software using Sparse Matrix Technology” 1994.

Associate Editor: Technometrics, 1995–2002

Organizer and Chair: Dimension Reduction for Simulation Science Data at the Joint Statistical Meetings, Atlanta, GA, 2001.

Program Committee: C. Warren Neel Conference on Statistical Data Mining, Knoxville, 2002.

Organizer and Chair: Distributed Data Mining at C. Warren Neel Conference on Statistical Data Mining, Knoxville, June 22-25, 2002.

Task Force Invitee: High End Computing Revitalization Task Force (HECRTF), Washington, DC, June 16-18, 2003.

Program Committee: 6th International Workshop on High Performance Data Mining: Pervasive and Data Stream Mining, 2003.

Panelist: Science Case for Large Scale Simulation (SCaLeS) Workshop, Washington, DC, June 24-25, 2003

Management Committee: Spring Research Conference on Statistics series co-sponsored by the Section on Physical and Engineering Sciences of the American Statistical Association and the Institute of Mathematical Statistics, 2003-2005.

Conference Co-Chair: Joint Research Conference on Statistics in Quality, Industry, and Technology, (Chair of Spring Research Conference Component), Knoxville, 2006

Task Force: American Statistical Association Presidential Task Force on Interactions with Other Organizations, 2006–2007

Search Committee: Governor's Chair in Statistics UT-ORNL Search Committee, 2007-2008

Organizing Committee: ACM Workshop on Knowledge Discovery from Sensor Data at KDD 2007

Program Committee: 8th SIAM International Conference on Data Mining, SDM 2008

Program Committee: 2008 International Workshop on Resiliency in High Performance Computing

Grants and Awards

Principal Investigator: National Institutes of Health grant, ‘Dose Estimation from Daily and Weekly Dosimetry Data,’ 1996-1997, \$150,000.

Principal Investigator: Strategic Environmental Research and Development Program (SERDP), “Spatial Statistical Models and Optimal Survey Design for Rapid Geophysical Characterization of UXO Sites,” 2001-2002, \$663,000.

Principal Investigator: Laboratory Directed Research and Development Program (LDRD), “Computing Transition States on High Dimensional Potential Surfaces with Application to Chemistry in Nanospaces,” 2001-2002, \$100,000.

Principal Investigator: Laboratory Directed Research and Development Program (LDRD), “Scalable Tools for Petascale Distributed Data Analysis,” 2002-2003, \$630,000.

Technical Co-Leader, ORNL subproject: Scientific Data Management Integrated Software Infrastructure Center (DOE SciDAC), 2002-2004, \$1,660,000.

Principal Investigator: Laboratory Directed Research and Development Program (LDRD), “Bringing Statistical Visualization to the Terascale and Beyond: Visual Analysis in Full Context,” 2004-2005, \$565,000.

Subproject Co-PI: SciDAC Center for Enabling Technologies, “Visualization and Analytics Center for Enabling Technologies (VACET),” 2007-2011, \$11,000,000 (\$2,000,000 for ORNL)

Languages

Fluent in English, Russian, and Czech. Can function in German.

Publications

Refereed and Solicited Publications

- [1] George Ostrouchov. Symbolic Givens reduction and row-ordering in large sparse least squares problems. *SIAM J. Scientific and Statistical Computation*, 8:248–264, 1987.
- [2] George Ostrouchov and W. Q. Meeker, Jr. Accuracy of approximate confidence bounds computed from interval censored Weibull and log-normal data. *J. Statistical Computation and Simulation*, 29:43–76, 1988.
- [3] George Ostrouchov. ANOVA model fitting via sparse matrix computations: a fast direct method. *SIAM J. Scientific and Statistical Computation*, 10:58–71, 1989.
- [4] George Ostrouchov. Computer communication: electronic bulletin boards. *Statistical Computing & Graphics*, 1(1):14–15, 1990.
- [5] George Ostrouchov. Computer communication: software distribution libraries. *Statistical Computing & Graphics*, 1(2):17–18, 1990.
- [6] George Ostrouchov. Computer communication: anonymous ftp. *Statistical Computing & Graphics*, 2(1):15–16, 1991.
- [7] George Ostrouchov. Computer communication: what's your e-mail address? *Statistical Computing & Graphics*, 2(2):22–23, 1991.
- [8] George Ostrouchov. Computer communication: alternate connections and white pages. *Statistical Computing & Graphics*, 3(2):22–23, 1992.
- [9] George Ostrouchov. Computer communication: resource discovery. *Statistical Computing & Graphics*, 3(1):21–22, 1992.
- [10] George Ostrouchov and Edward L. Frome. A model search procedure for hierarchical models. *Computational Statistics & Data Analysis*, 15:285–296, 1993.
- [11] George Ostrouchov. Gopher and other resource discovery tools. *Statistical Computing & Graphics*, 4(1):16–17, 1993.
- [12] Toby J. Mitchell, George Ostrouchov, Edward L. Frome, and George D. Kerr. A method for estimating occupational radiation dose to individuals, using weekly dosimetry data. *Radiation Research*, 147:195–207, 1997.
- [13] George Ostrouchov. Accounting for bias and measurement error in occupational studies. *Radiation Research*, 151:107–108, 1999.
- [14] Jingqian Jiang, Michael W. Berry, June M. Donato, George Ostrouchov, and Nancy W. Grady. Mining consumer product data via latent semantic indexing. *Intelligent Data Analysis*, 3:377–398, 1999.

- [15] D. J. Downing, V. V. Fedorov, W. F. Lawkins, M. D. Morris, and G. Ostrouchov. Large data series: Modeling the usual to identify the unusual. *Computational Statistics & Data Analysis*, 32:245–258, 2000.
- [16] Nagiza F. Samatova, George Ostrouchov, G. Al Geist, and Anatoli Melechko. RACHET: A new algorithm for clustering multi-dimensional distributed datasets,. In *SIAM Third Workshop on Mining Scientific Datasets*, pages 16–24, 2001.
- [17] Nagiza F. Samatova, G. Al Geist, George Ostrouchov, and Anatoli Melechko. Parallel out-of-core algorithm for genome-scale enumeration of metabolic systemic pathways. In *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS.02)*, pages 8–17, 2002.
- [18] Faisal N. Abu-Khzam, Nagiza Samatova, George Ostrouchov, Michael A. Langston, and Al Geist. Distributed dimension reduction algorithms for widely dispersed data. In *Parallel and Distributed Computing and Systems*, pages 174–178. ACTA Press, 2002.
- [19] Y. M. Qu, G. Ostrouchov, N. F. Samatova, and G. A. Geist. Principal component analysis for dimension reduction in massive distributed data sets. In *Workshop on High Performance Data Mining at the Second SIAM International Conference on Data Mining*, pages 4–9, 2002.
- [20] Nagiza F. Samatova, George Ostrouchov, G. Al Geist, and Anatoli Melechko. RACHET: An efficient cover-based merging of clustering hierarchies from distributed datasets. *Distributed and Parallel Databases*, 11:157–180, 2002.
- [21] G.-X. Yu, G. A. Geist, G. Ostrouchov, and N. F. Samatova. An SVM-based algorithm for identification of photosynthesis-specific genomes features. In H. Caussinus, P. Ettinger, and R. Tomassone, editors, *Proceedings of the IEEE Bioinformatics Conference*, pages 235–243, Stanford, CA, 2003.
- [22] B.-H. Park, N. Samatova, G. Ostrouchov, and G. A. Geist. XMap: Fast dimension reduction algorithms for multivariate streamline data. In *Proceedings of the 6th International Workshop on High Performance Data Mining: Pervasive and Data Stream Mining*, pages 1–6, 2003.
- [23] G. Ostrouchov and N. F. Samatova. Embedding methods and robust statistics for dimension reduction. In Jaromir Antoch, editor, *COMPSTAT2004*, pages 359–370. Physica-Verlag, 2004.
- [24] B.-H. Park, G. Ostrouchov, and N. F. Samatova. Reservoir-based random sampling with replacement from a data stream. In *Third SIAM International Conference on Data Mining*. April 22-24, 2004, Kissimmee, FL, 2004.
- [25] George Ostrouchov and Nagiza F. Samatova. On FastMap and the convex hull of multivariate data: Toward fast and robust dimension reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:1340–1343, 2005.
- [26] W Bethel, C Johnson, C Hansen, S Parker, A Sanderson, C Silva, X Tricoche, V Pascucci, H Childs, J Cohen, M Duchaineau, D Laney, P Lindstrom, S Ahern, J Meredith, G Ostrouchov, K Joy, and B Hamann. VACET: Proposed SciDAC2 visualization and analytics center for enabling technologies. *Journal of Physics: Conference Series*, 46:561–569, 2006.
- [27] S. Ahern, J. R. Daniel, J. Gao, G. Ostrouchov, R. J. Toedte, and C. Wang. Multi-scale data visualization for computational astrophysics and climate dynamics at Oak Ridge National Laboratory. *Journal of Physics: Conference Series*, 46:550–555, 2006.

- [28] S. Khan, A. R. Ganguly, S. Bandyopadhyay, S. Saigal D. J. Erickson III, V. Protopopescu, and G. Ostrouchov. Nonlinear statistics reveals stronger ties between ENSO and the tropical hydrological cycle. *Geophysical Research Letters*, 33, 2006. L24402, doi:10.1029/2006GL027941.
- [29] B-H. Park, G. Ostrouchov, and N.F. Samatova. Sampling streaming data with replacement. *Computational Statistics & Data Analysis*, 52:750–762, 2007.
- [30] E. W. Bethel, C. Johnson, K. Joy, S. Ahern, V. Pascucci, H. Childs, J. Cohen, M. Duchaineau, B. Hamann, C. Hansen, D. Laney, P. Lindstrom, J. Meredith, G. Ostrouchov, S. Parker, C. Silva, A. Sanderson, and X. Tricoche. SciDAC visualization and analytics center for enabling technology. *Journal of Physics: Conference Series*, 78:012032 (5pp), 2007.
- [31] S. Khan, G. Kuhn, A. R. Ganguly, III D. J. Erickson, and G. Ostrouchov. Spatio-temporal variability of daily and weekly precipitation extremes in South America. *Water Resources Research*, 43, 2007. W11424, doi:10.1029/2006WR005384.
- [32] E. W. Bethel, C. Johnson, C. Aragon, Prabhat, O. Rbel, G. Weber, V. Pascucci, H. Childs, P.-T. Bremer, B. Whitlock, S. Ahern, J. Meredith, G.Ostrouchov, K. Joy, B. Hamann, C. Garth, M. Cole, C. Hansen, S. Parker, A. Sanderson, C. Silva, and X. Tricoche. DOEs SciDAC visualization and analytics center for enabling technologies - strategy for petascale visual data analysis success. *CTWatch Quarterly*, 3(4), November 2007.
- [33] S. Khan, S. Bandyopadhyay, A. R. Ganguly, S. Saigal D. J. Erickson III, V. Protopopescu, and G. Ostrouchov. Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Physical Review E*, 76:1–15, 2007.
- [34] Kenneth I Joy, Mark Miller, Hank Childs, E Wes Bethel, John Clyne, George Ostrouchov, and Sean Ahern. Frameworks for visualization at the extreme scale. *Journal of Physics: Conference Series*, 78:012035 (10pp), 2007.
- [35] G. Ostrouchov, W. E. Doll, L. P. Beard, M. D. Morris, and D. A. Wolf. Multiscale structure of UXO site characterization: Spatial estimation and uncertainty quantification. *Stochastic Environmental Research and Risk Assessment*, page In Press, 2007.
- [36] George Ostrouchov. A matrix computation view of the FastMap and RobustMap dimension reduction algorithms. *SIAM J. Matrix Anal. Appl.*, page Submitted, 2007.

Other Publications

- [1] George Ostrouchov. Parallel computing on a hypercube: an overview of the architecture and some applications. In Richard M. Heiberger, editor, *Proceedings of the 19th Symposium on the Interface of Computer Science and Statistics*, pages 27–32. American Statistical Association, 1987.
- [2] George Ostrouchov and Edward L. Frome. A model search procedure for hierarchical log-linear models. In *1989 Proceedings of the Statistical Computing Section*, pages 277–282. American Statistical Association, 1989.
- [3] T. A. Vineyard, D. J. Downing, R. C. Durfee, J. J. Edwards, D. M. Flanagan, M. C. Fletcher, R. T. Goeltz, G. Ostrouchov, J. A. Rome, M. J. Saltmarsh, J. L. Smyre, and W. R. Wing. Assessment of en route sector and terminal air traffic control performance. Technical Report

K/DSRD-733, limited distribution, Martin Marietta Energy Systems, Oak Ridge, TN 37831, 1991.

- [4] George Ostrouchov. HModel: An X tool for global model search. In Yadolah Dodge and Joe Whittaker, editors, *Computational Statistics, Volume 1*, pages 269–274. Physica-Verlag, 1992.
- [5] S. Keller-McNulty and G. Ostrouchov. Error-free least squares based on LU factorization applicable to sparse problems. Technical Report I-95-1, Department of Statistics, Kansas State University, 1995.
- [6] D. J. Downing, W. F. Lawkins, M. D. Morris, and G. Ostrouchov. A method for detecting changes in long time series. Technical Report ORNL/TM-12879, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1995.
- [7] D. J. Downing, V. Fedorov, W. F. Lawkins, M. D. Morris, and G. Ostrouchov. Large datasets: Segmentation, feature extraction, and compression. Technical Report ORNL/TM-13114, Oak Ridge National Laboratory, 1996.
- [8] D. J. Downing, V. V. Fedorov, W. F. Lawkins, M. D. Morris, and G. Ostrouchov. Analysing perturbations and nonstationarity in time series using techniques motivated by the theory of chaotic nonlinear dynamical systems. Technical Report ORNL/TM-13115, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1996.
- [9] George Ostrouchov. Review of: “S+SpatialStats: User’s Manual for Windows and Unix” by Stephen P. Kaluzny, Silvia C. Vega, Tamre P. Cardoso, and Alice A. Shelby. *Short Book Reviews*, 18(2):26–27, August 1998.
- [10] George Ostrouchov, Gregory P. Zimmerman, John J. Beauchamp, Valerii V. Fedorov, and Darryl J. Downing. Evaluation of statistical methodologies used in U.S. Army ordnance and explosive work. Technical Report ORNL/TM-13588, Oak Ridge National Laboratory, Oak Ridge, TN 37830, 1998.
- [11] George Ostrouchov, Edward L. Frome, and George D. Kerr. Dose estimation from daily and weekly dosimetry data. Technical Report ORNL/TM-1999/282, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1999.
- [12] W. M. Putman, J.B. Drake, and G. Ostrouchov. Statistical downscaling of United States regional climate from transient GCM scenarios. In *15th Conference on Probability and Statistics in the Atmospheric Sciences, Asheville, North Carolina*, pages J8–J11, 2000.
- [13] T.H. Dunigan and G. Ostrouchov. Flow characterization for intrusion detection. Technical Report ORNL/TM-2001/115, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 2001.
- [14] R.D. Burris, S. Cholia, T.H. Dunigan, F.M. Fowler, M.K. Gleicher, H.H. Holmes, N.E. Johnston, N.L. Meyer, D.L. Million, G. Ostrouchov, and N.F. Samatova. Probe project status and accomplishments - year two. Technical Report ORNL/TM-TBD, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 2002.
- [15] G. Ostrouchov, W. E. Doll, D. A. Wolf, M. D. Morris, L. P. Beard, D. K. Butler, and J. E. Simms. Spatial statistical model and optimal survey design for rapid geophysical characterization of uxo sites. Technical Report Project CU-1201, SERDP, 2003.

- [16] A. Martino, A. Gorin, T. Lane, S. Plimpton, N. Samatova, Y. Xu, H. Al-Hashimi, C. Strauss, B.-H. Park, G. Ostrouchov, A. Geist, W. Hart, and D. Roe. Analysis of protein complexes from a fundamental understanding of protein binding domains and protein-protein interactions in *synechococcus wh8102*. Technical report, Genomes to Life Contractor-Grantee Workshop I, Arlington, Virginia, February 9-12 2003.
- [17] B.-H. Park, G. Ostrouchov, G.-X. Yu, A. Geist, A. Gorin, and N. F. Samatova. PICUPP: Protein interaction classification by unlikely profile pair. *Journal of Computational Biology*, Submitted, 2003.
- [18] R.D. Burris, S. Cholia, T.H. Dunigan, F.M. Fowler, M.K. Gleicher, H.H. Holmes, N.E. Johnston, N.L. Meyer, D.L. Million, G. Ostrouchov, and N.F. Samatova. Probe project status and accomplishments. Technical Report ORNL/TM-2003/140, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 2003.
- [19] A. R. Ganguly, S. Khan, D. J. Erickson, R. W. Katz, G. Ostrouchov, V. A. Protopopescu, S. Bandyopadhyay, and S. Saigal. Multivariate dependence in complex systems. In *Fifth Symposium on Understanding Complex Systems*. University of Illinois at Urbana-Champaign, 2005.
- [20] R. Ganguly, T. Hsing, R. Katz, D. Erickson, G. Ostrouchov, T. Wilbanks, and N. Cressie. Multivariate dependence among extremes, abrupt change and anomalies in space and time for climate applications. In *Workshop on Data Mining for Anomaly Detection at The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2005.
- [21] G. Ostrouchov. Inner layer threat with two layer protection systems. Technical report, OOU report to DOE/SC Office of Safety, Security and Infrastructure (SC-31), Oak Ridge, TN, 37831, 2006.
- [22] George Ostrouchov. Odds of more math articles increase. *ORNL Reporter*, page 3, 2006. (A note on probability of football score matching winning lottery numbers).
- [23] G. Ostrouchov. Analysis by threat decomposition. Technical report, OOU report to DOE/SC Office of Safety, Security and Infrastructure (SC-31), Oak Ridge, TN, 37831, 2007.
- [24] M. Anitescu, G. Ostrouchov, and L. Pytak-Nolte. Priority research direction: Improved media parameterization and reconstruction. In *Report of the Computational Subsurface Sciences Workshop*, pages 242–250. DOE Office of Science, 2007.