

George Ostrouchov

Scientific Data Group
Computer Science and Mathematics Division
Oak Ridge National Laboratory
1 Bethel Valley Road
Oak Ridge, Tennessee 37831-6367

Tel: (865) 574-3137
ostrouchovg @ ornl . gov
<http://www.csm.ornl.gov/~ost>

Education

Ph.D., Statistics

Iowa State University, Ames, IA, December 1984

Dissertation: *Large Sparse Least Squares Computations*

M.Sc., Statistics

Iowa State University, Ames, IA, May 1981

Project: *Accuracy of Approximate Confidence Bounds Computed from Interval Censored Weibull and Lognormal Data*

B.Math., Honours Co-op, Statistics

University of Waterloo, Waterloo, Canada, May 1978

Current Appointments

Since 2012	Analytics Task Lead Scientific Data Group, Computer Science and Mathematics Division, Oak Ridge National Laboratory
Since 2003	Senior Research Staff Member Scientific Data Group (previously Statistics and Data Sciences Group), Computer Science and Mathematics Division, Oak Ridge National Laboratory
Since 2009	Joint Faculty Professor of Statistics Department of Business Analytics and Statistics, The University of Tennessee, Knoxville
Since 2008	Computational and Applied Mathematics Staff Member The University of Tennessee and Oak Ridge National Laboratory Joint Institute for Computational Sciences

Past Appointments

2003-2009	Adjunct Professor of Statistics Department of Statistics, Operations, and Management Science (initially Department of Statistics), The University of Tennessee, Knoxville
1993-2003	Research Staff Member II Statistics and Data Sciences Group (previously Statistics Group in Mathematical Sciences Section) Computer Science and Mathematics Division (previously Engineering Physics and Mathematics Division), Oak Ridge National Laboratory
1983-1993	Research Staff Member I

Statistics Group, Mathematical Sciences Section (initially Mathematics and Statistics Research Section), Engineering Physics and Mathematics Division (initially in Computer Sciences Division), Oak Ridge National Laboratory

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 Summer **Research Assistant**, Universität Hohenheim, Germany

1978 Summer **Research Assistant**, Swiss International Air Lines AG, Zurich, Switzerland

1974-1979 **Statistician/Analyst**, Informetrica Ltd., Ottawa, Canada
(Co-op + 1yr)

Memberships

American Association for the Advancement of Science (AAAS), - Sections: Statistics; Information, Computing, and Communication; Mathematics

American Statistical Association (ASA), Sections: Physical and Engineering Sciences; Statistical Computing; Statistical Graphics

International Association for Statistical Computing (IASC)

Society for Industrial and Applied Mathematics (SIAM)

The R Foundation

Association for Computing Machinery: Special Interest Group on High Performance Computing (ACM-SIGHP)

National Ski Patrol

Recognitions

Distinguished Contributor, Computer Science and Mathematics Division, Oak Ridge National Laboratory, 2012

Fellow, American Statistical Association, *“For excellent and sustained research and collaboration involving the statistical analysis of massive data sets, and for outstanding service to the mathematical sciences community.”* 2010

Certificate of Appreciation from US Undersecretary of Energy *“For exemplary performance in ensuring the success of the Terrorism Prevention Measures Optimization Project conducted on behalf of the Department of Energy’s Office of Science”*, 2007

Martin Marietta Energy Systems, Government-Use Invention Award for *“Cost Matrix Software using Sparse Matrix Technology”*, 1994. The first use of sparse matrix methods to enable new transparency in large accounting systems.

Phi Kappa Phi, 1981

Mu Sigma Rho, 1980

Professional Activities

Program Committee 3rd Workshop on Advances in Software and Hardware for Big Data to Knowledge Discovery (ASH) at IEEE International Conference on Big Data, Dec. 5-8, 2016, Washington DC

Organizing Committee: Conference on Data Analysis (CoDA2016), March 2-4, 2016, Santa Fe, NM

Guest Associate Editor Statistical Analysis and Data Mining, Special Issue, 2014-2015

Organizer: Birds-of-a-Feather Session on *Super-R: Supercomputing and R for Data-Intensive Analysis*, at The International Conference for High Performance Computing, Networking, Storage and Analysis, November 16-21, 2014, Denver, CO.

Program Committee Workshop on Advances in Software and Hardware for Big Data to Knowledge Discovery (ASH) at IEEE International Conference on Big Data, October 27-30, 2014, Washington, DC

Organizer: Birds-of-a-Feather Session on *Super-R: Supercomputing & R for Data-Intensive Analysis*, at The International Supercomputing Conference, June 22-26, 2014, Leipzig, Germany.

Organizing Committee: Conference on Data Analysis (CoDA2014), March 5-7, 2014, Santa Fe, NM

Organizer: Birds-of-a-Feather Session on *Super-R: Supercomputing and R for Data-Intensive Analysis*, at The International Conference for High Performance Computing, Networking, Storage and Analysis, November 17-22, 2013, Denver, CO.

Guest Associate Editor Technometrics, Special Issue on Data-Focused Research across the Department of Energy, 2012-2013

Organizing Committee: Conference on Data Analysis (CoDA2012), February 29 - March 2, 2012, Santa Fe, NM

Elected Program Chair: ASA SPES, Joint Statistical Meetings, July 31 - August 5, 2010, Vancouver, Canada

Program Committee: 4th International Workshop on Knowledge Discovery from Sensor Data at KDD 2010, July 25-28, 2010, Washington, DC

Program Committee: Workshop on Resiliency in High Performance Computing (Resilience 2010) at IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid 2010), May 17-20, 2010, Melbourne, Australia

Organizer: Advancing Clarity and Scale in Statistical Computing at Joint Research Conference on Statistics in Quality, Industry, and Technology, May 25 - 27, 2010, Gaithersburg, Maryland

Program Committee: Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes, and Impacts at ICDM 2009, December 6-9, 2009, Miami, Florida

Program Chair-Elect: American Statistical Association, Section on Physical and Engineering Sciences, Joint Statistical Meetings, August 2-6, 2009, Washington, DC

Program Committee: 3rd International Workshop on Knowledge Discovery from Sensor Data at KDD 2009, June 28-July 1, 2009, Paris, France

Program Committee: Workshop on Resiliency in High Performance Computing (Resilience 2009) at HPDC, June 9-13, 2009, Munich, Germany

Program Committee: Workshop on Radiation Effects and Fault Tolerance in Nanometer Technologies at ACM ICCF, May 5-7, 2008, Ischia, Italy

Panelist and Moderator: Workshop on Mathematics for Analysis of Petascale Data, sponsored by the Department of Energy's Office of Advanced Scientific Computing Research, 2008

Program Committee: Workshop on Resiliency in High Performance Computing (Resilience 2008) at CCGrid, May 19-22, 2008, Lyon, France

Program Committee: Society for Industrial and Applied Mathematics, Conference on Data Mining, April 24-26, 2008, Atlanta, GA

Organizing Committee: 1st International Workshop on Knowledge Discovery from Sensor Data at KDD 2007, August 12, 2007, San Jose, CA

Co-Chair: Search Committee for Governor's Chair in Statistics at ORNL/UT, 2007-2008

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

Conference Chair: 13th Spring Research Conference (SRC) on Statistics in Industry and Technology (as co-Chair of Joint Research Conference on Statistics in Quality, Industry, and Technology), Knoxville, TN, 2006

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.

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

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

Panelist: Interagency Working Group on High End Computing (HEC IWG), Washington, DC, June 16-18, 2003.

Planning Committee: American Statistical Association Planning Meeting on Statistics and National Defense and Security, Washington, DC, 2002

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

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

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

Associate Editor: Technometrics, 1995-2002.

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

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

Grants and Awards

PI: “Harnessing Scalable Libraries for Statistical Computing on Modern Architectures and Bringing Statistics to Large Scale Computing,” National Science Foundation, Division of Mathematical Sciences Grant, 2014-2017, \$600,000

Data Analysis Services Lead: “NICS Remote Data Analysis and Visualization Center,” National Science Foundation, Office of Cyber Infrastructure Grant, 2009-2013, \$10,000,000.

Co-PI: “Petascale Enabled Discovery,” Oak Ridge National Laboratory Computational Science Initiative, 2008-2009, \$62,000.

Senior Personnel: “Visualization and Analytics Center for Enabling Technologies (VACET)” Department of Energy, Office of Science Award, 2007-2011, \$11,000,000 (\$2,000,000 for ORNL)

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

Senior Personnel: “Scientific Data Management Integrated Software Infrastructure Center” Department of Energy, Office of Science Award, 2002-2004, \$1,660,000.

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

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

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

PI: “Dose Estimation from Daily and Weekly Dosimetry Data,” National Institutes of Health Grant, 1996-1997, \$150,000.

Mentoring Activities

Linda Keleher, B.S., Ohio Northern University, Oak Ridge Science and Engineering Research Semester 1989

Eric Sedlacek, B.S., Ohio Northern University, Oak Ridge Science and Engineering Research Semester 1990

John Pospisil, B.S., Nebraska Wesleyan University, Oak Ridge Science and Engineering Research Semester 1991

Asim YarKhan, M.S., University of Tennessee, Graduate Research Assistant 1993-1994

DeMarkus V. Webb, B.S., University of Tennessee, Research Alliance for Minorities Summer, 2000

Yongming Qu, Ph.D., Iowa State University, DOE Higher Education Research Experience, 2001

Ade Ola, Ph.D., Virginia State University, Minority Educational Institution Summer Faculty Research Program 2002

Bryan Hathorn, Ph.D., California Institute of Technology, Postdoctoral Appointment, 2001-2002

David A. Bauer, B.Sc., Georgia Institute of Technology, Energy Research Undergraduate Laboratory Fellowship 2002

Byung Hoon Park, Ph.D. University of Maryland, Postdoctoral Appointment, 2002-2004

Jennifer Golek, Ph.D., University of Tennessee, DOE Higher Education Research Experience 2002-2003

Ian Watkins, M.S., University of Tennessee, DOE Higher Education Research Experience 2002-2003

Alan Parks, Ph.D., Lawrence University, Oak Ridge Science Semester 2003

Rajesh V. Munavalli, M.S., DOE Higher Education Research Experience 2003-2004

Aruna Buddana, M.S., University of Tennessee, DOE Higher Education Research Experience, 2004

Abdelhamid Meziani, Ph.D., Florida International University, Minority Educational Institution Summer Faculty Research Program, 2004

Lionel Lovett, B.Sc., Jackson State University, ORNL Research Alliance in Math and Science 2005

Houssain Kettani, Ph.D., Jackson State University, DOE Higher Education Research Experience, 2005, 2006

James R. Wilcox, Tennessee Governors Academy, 2008-2009

Fernando E. Fuentes, B.Sc., Polytechnic University of Puerto Rico, ORNL Research Alliance in Math and Science 2009

Yael M Camacho-Bonaparte, B.Sc., Polytechnic University of Puerto Rico, FaST, 2009, 2011

Shamir J Quinones Dueno, B.Sc., Polytechnic University of Puerto Rico, FaST, 2009

Emmanuel Aviles Saez, B.Sc., Polytechnic University of Puerto Rico, FaST, 2009

Javier Colon, B.Sc., Polytechnic University of Puerto Rico, FaST, 2011

Yanran Lu, B.Sc., Princeton University, NSF, 2011

UMBC REU Program University of Maryland, Baltimore County, National Science Foundation, Research Experiences for Undergraduates, 2011

Wei-Chen Chen, Ph.D., Iowa State University, Postdoctoral Appointment, 2011-2013

Hilde Oliver, B.Sc., University of South Carolina, DOE Higher Education Research Experience, 2012

UMBC REU Program University of Maryland, Baltimore County, National Science Foundation, Research Experiences for Undergraduates, 2013

Denver Coker, B.Sc. University of North Georgia, Joint Institute of Computational Sciences Summer Internship, 2013

Reid Vincent Paris, B.Sc. University of Maine, DOE Higher Education Research
Experience, 2015

Drew Schmidt, M.Sc. University of Tennessee, Graduate Research Assistant, 2015-2017

Publications

Refereed and Solicited Publications

- [1] George Ostrouchov. “Symbolic Givens reduction and row-ordering in large sparse least squares problems”. In: *SIAM J. Scientific and Statistical Computation* 8 (1987), pp. 248–264.
- [2] George Ostrouchov and W. Q. Meeker, Jr. “Accuracy of approximate confidence bounds computed from interval censored Weibull and log-normal data”. In: *J. Statistical Computation and Simulation* 29 (1988), pp. 43–76.
- [3] George Ostrouchov. “ANOVA model fitting via sparse matrix computations: a fast direct method”. In: *SIAM J. Scientific and Statistical Computation* 10 (1989), pp. 58–71.
- [4] George Ostrouchov. “Computer communication: electronic bulletin boards”. In: *Statistical Computing & Graphics* 1 (1 1990), pp. 14–15.
- [5] George Ostrouchov. “Computer communication: software distribution libraries”. In: *Statistical Computing & Graphics* 1 (2 1990), pp. 17–18.
- [6] George Ostrouchov. “Computer communication: anonymous ftp”. In: *Statistical Computing & Graphics* 2 (1 1991), pp. 15–16.
- [7] George Ostrouchov. “Computer communication: what’s your e-mail address?”. In: *Statistical Computing & Graphics* 2 (2 1991), pp. 22–23.
- [8] George Ostrouchov. “Computer communication: alternate connections and white pages”. In: *Statistical Computing & Graphics* 3 (2 1992), pp. 22–23.
- [9] George Ostrouchov. “Computer communication: resource discovery”. In: *Statistical Computing & Graphics* 3 (1 1992), pp. 21–22.
- [10] George Ostrouchov. “Gopher and other resource discovery tools”. In: *Statistical Computing & Graphics* 4(1) (1 1993), pp. 16–17.
- [11] George Ostrouchov and Edward L. Frome. “A model search procedure for hierarchical models”. In: *Computational Statistics & Data Analysis* 15 (1993), pp. 285–296.
- [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”. In: *Radiation Research* 147 (1997), pp. 195–207.
- [13] Jingqian Jiang, Michael W. Berry, June M. Donato, George Ostrouchov, and Nancy W. Grady. “Mining consumer product data via latent semantic indexing”. In: *Intelligent Data Analysis* 3 (1999), pp. 377–398.
- [14] George Ostrouchov. “Accounting for bias and measurement error in occupational studies”. In: *Radiation Research* 151 (1999), pp. 107–108.
- [15] Darryl J. Downing, Valerii V. Fedorov, William F. Lawkins, Max D. Morris, and George Ostrouchov. “Large Data Series: Modeling the Usual to Identify the Unusual”. In: *Computational Statistics & Data Analysis* 32 (2000), pp. 245–258.

- [16] 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*. ACTA Press, 2002, pp. 174–178.
- [17] Yong Ming Qu, George Ostrouchov, Nagiza F. Samatova, and G. A. Geist III. “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*. 2002, pp. 4–9.
- [18] 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)*. 2002, pp. 8–17.
- [19] Nagiza F. Samatova, George Ostrouchov, G. Al Geist, and Anatoli Melechko. “RACHET: An Efficient Cover-Based Merging of Clustering Hierarchies from Distributed Datasets”. In: *Distributed and Parallel Databases* 11 (2002), pp. 157–180.
- [20] Byung-Hoon Park, Nagiza Samatova, George Ostrouchov, and G. A. Geist III. “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*. 2003, pp. 1–6.
- [21] Gong-Xin Yu, George Ostrouchov, Al Geist, and Nagiza F. Samatova. “An SVM-based Algorithm for Identification of Photosynthesis-specific Genome Features”. In: *Computational Systems Bioinformatics Conference, International IEEE Computer Society* (2003), p. 235. DOI: <http://doi.ieeecomputersociety.org/10.1109/CSB.2003.1227323>.
- [22] George Ostrouchov and Nagiza F. Samatova. “Embedding methods and robust statistics for dimension reduction”. In: *COMPSTAT2004*. Ed. by Jaromir Antoch. Physica-Verlag, 2004, pp. 359–370.
- [23] Byung-Hoon Park, George Ostrouchov, and Nagiza F. Samatova. “Reservoir-based random sampling with replacement from a data stream”. In: *Proceedings of the 2004 SIAM International Conference on Data Mining*. 2004, pp. 492–496.
- [24] George Ostrouchov and Nagiza F. Samatova. “On FastMap and the Convex Hull of Multivariate Data: Toward Fast and Robust Dimension Reduction”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (2005), pp. 1340–1343.
- [25] 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”. In: *Journal of Physics: Conference Series* 46 (2006), pp. 550–555. URL: <http://stacks.iop.org/1742-6596/46/550>.
- [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”. In: *Journal of Physics: Conference Series* 46 (2006), pp. 561–569. URL: <http://stacks.iop.org/1742-6596/46/561>.

- [27] 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”. In: *Geophysical Research Letters* 33 (2006). L24402, doi:10.1029/2006GL027941.
- [28] E. W. Bethel, C. Johnson, C. Aragon, Prabhat, O. Rübel, 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. “DOE’s SciDAC Visualization and Analytics Center for Enabling Technologies - Strategy for Petascale Visual Data Analysis Success”. In: *CTWatch Quarterly* 3.4 (Nov. 2007). URL: <http://www.ctwatch.org/quarterly/articles/2007/11/does-scidac-visualization-and-analytics-center-for-enabling-technologies-strategy-for-petascale-visual-data-analysis-success/>.
- [29] 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”. In: *Journal of Physics: Conference Series* 78 (2007), 012032 (5pp). URL: <http://stacks.iop.org/1742-6596/78/012032>.
- [30] Kenneth I Joy, Mark Miller, Hank Childs, E Wes Bethel, John Clyne, George Ostrouchov, and Sean Ahern. “Frameworks for visualization at the extreme scale”. In: *Journal of Physics: Conference Series* 78 (2007), 012035 (10pp). URL: <http://stacks.iop.org/1742-6596/78/012035>.
- [31] 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”. In: *Physical Review E* 76 (2007), pp. 1–15.
- [32] 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”. In: *Water Resources Research* 43 (2007). W11424, doi:10.1029/2006WR005384.
- [33] Byung-Hoon Park, George Ostrouchov, and Nagiza F. Samatova. “Sampling Streaming Data with Replacement”. In: *Computational Statistics & Data Analysis* 52 (2007), pp. 750–762.
- [34] N. Taerat, N. Naksinehaboon, C. Chandler, J. Elliott, C. Leangsuksun, G. Ostrouchov, and S. L. Scott. “Using Log Information to Perform Statistical Analysis on Failures Encountered by Large-Scale HPC Deployments”. In: *High Availability and Performance Computing Workshop (HAPCW 2008)*. 2008, (6pp).
- [35] E W Bethel, C Johnson, S Ahern, J Bell, P-T Bremer, H Childs, E Cormier-Michel, M Day, E Deines, T Fogal, C Garth, C G R Geddes, H Hagen, B Hamann, C Hansen, J Jacobsen, K Joy, J Kruger, J Meredith, P Messmer, G Ostrouchov, V Pascucci, K Potter, Prabhat, D Pugmire, O Rubel, A Sanderson, C Silva, D Ushizima, G Weber, B Whitlock, and K Wu. “Occam’s razor and petascale visual data analysis”. In: *Journal of Physics: Conference Series* 180 (2009), 012084 (18pp). URL: <http://stacks.iop.org/1742-6596/180/012084>.

- [36] G. Ostrouchov, T. Naughton, C. Engelmann, G. Vallée, and S. L. Scott. “Nonparametric Multivariate Anomaly Analysis in Support of HPC Resilience”. In: *Proceedings of the 5th IEEE International Conference on E-Science Workshops*. Dec. 2009, pp. 80–85. DOI: 10.1109/ESCIW.2009.5407992.
- [37] George Ostrouchov. “A Matrix Computation View of FastMap and RobustMap Dimension Reduction Algorithms”. In: *SIAM Journal on Matrix Analysis and Applications* 31.3 (2009), pp. 1351–1360. DOI: 10.1137/070710767.
- [38] George Ostrouchov, William E. Doll, Les P. Beard, Max D. Morris, and Dennis A. Wolf. “Multiscale Structure of UXO Site Characterization: Spatial Estimation and Uncertainty Quantification”. In: *Stochastic Environmental Research and Risk Assessment* 23.2 (2009), pp. 215–225.
- [39] N. Taerat, N. Naksinehaboon, C. Chandler, J. Elliott, C. Leangsuksun, G. Ostrouchov, S. L. Scott, and C. Engelmann. “Blue Gene/L Log Analysis and Time to Interrupt Estimation”. In: *Availability, Reliability and Security, International Conference on*. Los Alamitos, CA, USA: IEEE Computer Society, 2009, pp. 173–180. ISBN: 978-0-7695-3564-7. DOI: <http://doi.ieeecomputersociety.org/10.1109/ARES.2009.105>.
- [40] F. Fuentes, H. Kettani, G. Ostrouchov, M. Stoitsov, and H.A. Nam. “Exploration of High-Dimensional Nuclei Data”. In: *Communication Software and Networks, 2010. ICCSN '10. Second International Conference on*. Feb. 2010, pp. 521–524. DOI: 10.1109/ICCSN.2010.105.
- [41] Robert Sisneros, Jian Huang, George Ostrouchov, and Forrest M. Hoffman. “Visualizing Life Zone Boundary Sensitivities Across Climate Models and Temporal Spans”. In: *Procedia CS* 4 (2011), pp. 1582–1591.
- [42] Lei Jiang, Pragneshkumar B. Patel, George Ostrouchov, and Ferdinand Jamitzky. “OpenMP-style parallelism in data-centered multicore computing with R”. In: *SIGPLAN Not.* 47.8 (Feb. 2012), pp. 335–336. ISSN: 0362-1340. DOI: 10.1145/2370036.2145882. URL: <http://doi.acm.org/10.1145/2370036.2145882>.
- [43] Jeremy Logan, Scott Klasky, Hasan Abbasi, Qing Liu, George Ostrouchov, Manish Parashar, Norbert Podhorszki, Yuan Tian, and Matthew Wolf. “Understanding I/O Performance Using I/O Skeletal Applications”. In: *Euro-Par 2012 Parallel Processing*. Ed. by Christos Kaklamanis, Theodore Papatheodorou, and PaulG. Spirakis. Vol. 7484. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 77–88. ISBN: 978-3-642-32819-0. DOI: 10.1007/978-3-642-32820-6_10. URL: http://dx.doi.org/10.1007/978-3-642-32820-6_10.
- [44] Drew Schmidt, George Ostrouchov, Wei-Chen Chen, and Pragneshkumar Patel. “Tight Coupling of R and Distributed Linear Algebra for High-Level Programming with Big Data”. In: *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion: 2012*, pp. 811–815. DOI: 10.1109/SC.Companion.2012.113.

- [45] Wei-Chen Chen, George Ostrouchov, David Pugmire, Prabhat, and Michael Wehner. “A Parallel EM Algorithm for Model-Based Clustering Applied to the Exploration of Large Spatio-Temporal Data”. In: *Technometrics* 55.4 (2013), pp. 513–523. DOI: 10.1080/00401706.2013.826146.
- [46] George Ostrouchov, Drew Schmidt, Wei-Chen Chen, and Pragneshkumar Patel. “Combining R with Scalable Libraries to Get the Best of Both for Big Data”. In: *IASC Satellite Conference for the 59th ISI WSC & the 8th Conference of IASC-ARS*. 2013, pp. 85–90.
- [47] Robert Sisneros, Jian Huang, George Ostrouchov, Sean Ahern, and B. David Semeraro. “Contrasting Climate Ensembles: A Model-Based Visualization Approach for Analyzing Extreme Events”. In: *Procedia Computer Science* 18 (2013). 2013 International Conference on Computational Science, pp. 2347–2356. DOI: <http://dx.doi.org/10.1016/j.procs.2013.05.406>.
- [48] George Ostrouchov, Joshua New, Jibonananda Sanyal, and Pragneshkumar Patel. “Uncertainty Analysis of a Heavily Instrumented Building at Different Scales of Simulation”. In: *3rd International High Performance Buildings Conference*. 3561 (10pp). Purdue University, Lafayette, IN, July 2014.
- [49] Drew Schmidt, Wei-Chen Chen, Michael Matheson, and George Ostrouchov. “Programming with BIG Data in R: Scaling Analytics from One to Thousands of Nodes”. In: *Big Data Research* (2016). DOI: <http://dx.doi.org/10.1016/j.bdr.2016.10.002>.

Technical Reports and Non-Refereed Proceedings Publications

- [1] George Ostrouchov. “Large sparse least squares computations”. PhD thesis. Iowa State University, 1984.
- [2] George Ostrouchov. “Parallel computing on a hypercube: an overview of the architecture and some applications”. In: *Proceedings of the 19th Symposium on the Interface of Computer Science and Statistics*. Ed. by Richard M. Heiberger. American Statistical Association, 1987, pp. 27–32.
- [3] George Ostrouchov. “Statistical computing on a hypercube”. In: *20th Symposium on the Interface: Computing Science and Statistics, Final Report*. Ed. by Edward J. Wegman. American Statistical Association, 1988, p. 120.
- [4] George Ostrouchov and Edward L. Frome. “A model search procedure for hierarchical log-linear models”. In: *1989 Proceedings of the Statistical Computing Section*. American Statistical Association, 1989, pp. 277–282.
- [5] George Ostrouchov. *Minimum-size samples in model-based sampling*. Tech. rep. Oak Ridge, TN 37830: Oak Ridge National Laboratory, 1991.
- [6] 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*. Tech. rep. K/DSRD-733, limited distribution. Oak Ridge, TN 37831: Martin Marietta Energy Systems, 1991.

- [7] George Ostrouchov. “HModel: An X tool for global model search”. In: *Computational Statistics, Volume 1*. Ed. by Yadolah Dodge and Joe Whittaker. Physica-Verlag, 1992, pp. 269–274.
- [8] Darryl J. Downing, William F. Lawkins, Max D. Morris, and George Ostrouchov. *A Method for Detecting Changes in Long Time Series*. Tech. rep. ORNL/TM-12879. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 1995.
- [9] Sallie Keller-McNulty and George Ostrouchov. *Error-Free Least Squares Based on LU Factorization Applicable to Sparse Problems*. Tech. rep. I-95-1. Department of Statistics, Kansas State University, 1995.
- [10] Darryl J. Downing, Valerii V. Fedorov, William F. Lawkins, Max D. Morris, and George Ostrouchov. *Large Datasets: Segmentation, Feature Extraction, and Compression*. Tech. rep. ORNL/TM-13114. Oak Ridge National Laboratory, 1996.
- [11] Darryl J. Downing, Valerii V. Fedorov, William F. Lawkins, Max D. Morris, and George Ostrouchov. *Analysing Perturbations and Nonstationarity in Time Series using Techniques Motivated by the Theory of Chaotic Nonlinear Dynamical Systems*. Tech. rep. ORNL/TM-13115. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 1996.
- [12] Darryl J. Downing, Valerii V. Fedorov, William F. Lawkins, Max D. Morris, and George Ostrouchov. “Large Data Series: Modeling the Usual to Identify the Unusual”. In: *Computing Science and Statistics 29.2* (1997), pp. 8–14.
- [13] Richard L. Schmoyer and George Ostrouchov. “Multinomial Likelihood Ratio Test for Trend in Count Data”. Unpublished report, Computer Science and Mathematics Division, Oak Ridge National Laboratory. 1997.
- [14] 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”. In: *Short Book Reviews* 18.2 (Aug. 1998), pp. 26–27.
- [15] 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*. Tech. rep. ORNL/TM-13588. Oak Ridge, TN 37830: Oak Ridge National Laboratory, 1998.
- [16] George Ostrouchov, Edward L. Frome, and George D. Kerr. *Dose estimation from daily and weekly dosimetry data*. Tech. rep. ORNL/TM-1999/282. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 1999.
- [17] 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*. 2000, J8–J11.
- [18] Thomas H. Dunigan and George Ostrouchov. *Flow Characterization for Intrusion Detection*. Tech. rep. ORNL/TM-2001/115. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 2001.
- [19] 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*. 2001, pp. 16–24.

- [20] 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*. Tech. rep. ORNL/TM-TBD. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 2002.
- [21] 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*. Tech. rep. ORNL/TM-2003/140. Oak Ridge, TN 37831: Oak Ridge National Laboratory, 2003.
- [22] 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*. Tech. rep. Arlington, Virginia: Genomes to Life Contractor-Grantee Workshop I, Feb. 9–12, 2003, pp. 16–17.
- [23] George Ostrouchov, William E. Doll, Dennis A. Wolf, Max 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*. Tech. rep. Project CU-1201. SERDP, 2003.
- [24] George Ostrouchov and Nagiza F. Samatova. *High end computing for full-context analysis and visualization: when the experiment is done*. White paper accepted by the High End Computing Revitalization Task Force (HECRTF) Washington, DC. June 2003.
- [25] Byung-Hoon Park, George Ostrouchov, Gong-Xin Yu, G. A. Geist III, Andrey Gorin, and Nagiza F. Samatova. “Inference of protein-protein Interactions by unlikely profile pair”. In: *Third IEEE International Conference on Data Mining*. 2003, pp. 735–738.
- [26] 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.
- [27] 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 (KDD-05)*. 2005, pp. 25–26.
- [28] George Ostrouchov. *Inner Layer Threat with Two Layer Protection Systems*. Tech. rep. Oak Ridge, TN, 37831: OUO report to DOE/SC Office of Safety, Security and Infrastructure (SC-31), 2006.
- [29] George Ostrouchov. “Odds of more math articles increase”. In: *ORNL Reporter* (2006). (A note on probability of football score matching winning lottery numbers), p. 3.
- [30] M. Anitescu, G. Ostrouchov, and L. Pytak-Nolte. “Priority Research Direction: Improved Media Parameterization and Reconstruction”. In: *Report of the Computational Subsurface Sciences Workshop*. DOE Office of Science, 2007, pp. 242–250.

- [31] George Ostrouchov. *Analysis by Threat Decomposition*. Tech. rep. Oak Ridge, TN, 37831: OUO report to DOE/SC Office of Safety, Security and Infrastructure (SC-31), 2007.
- [32] S. L. Scott, C. Engelmann, H. H. Ong, G. R. Vallée, T. Naughton, A. Tikotekar, G. Ostrouchov, C. (Box) Leangsuksun, N. Naksinehaboon, R. Nassar, M. Paun, F. Mueller, C. Wang, A. B. Nagarajan, J. Varma, X. (Ben) He, L. Ou, and X. Chen. *Resiliency for High-Performance Computing Systems*. Poster at the <http://www.hpcsw.org> 1st High-Performance Computer Science Week (HPCSW) 2008, Denver, CO, USA. Mar. 30–Apr. 5, 2008.
- [33] George Ostrouchov. “Call for SPES Sponsored Contributed Posters and Papers at JSM 2010”. In: *SPES & QP Newsletter* 17.2 (Dec. 2009), p. 6.
- [34] George Ostrouchov. “Invited Sessions for SPES Program at JSM 2010”. In: *SPES & QP Newsletter* 17.2 (Dec. 2009), pp. 5–6.
- [35] George Ostrouchov. “JSM 2009 Roundtables Roundup”. In: *SPES & QP Newsletter* 17.2 (Dec. 2009), p. 5.
- [36] George Ostrouchov. *Mid-Life Failure Rate Estimation in Complex Repairable Systems*. Tech. rep. Oak Ridge, TN, 37831: A report to USEC review board, 2009.
- [37] Stephen L. Scott, Christian Engelmann, Geoffroy R. Vallée, Thomas Naughton, Anand Tikotekar, George Ostrouchov, Chokchai Leangsuksun, Nichamon Naksinehaboon, Raja Nassar, Mihaela Paun, Frank Mueller, Chao Wang, Arun B. Nagarajan, and Jyothish Varma. “A tunable holistic resiliency approach for high-performance computing systems”. In: *PPoPP ’09: Proceedings of the 14th ACM SIGPLAN symposium on Principles and practice of parallel programming*. Raleigh, NC, USA: ACM, New York, NY, USA, 2009, pp. 305–306. ISBN: 978-1-60558-397-6. DOI: <http://doi.acm.org/10.1145/1504176.1504227>.
- [38] George Ostrouchov. “Physical and Engineering Sciences JSM Invited Program Outlined”. In: *Amstat News* (June 2010), p. 53.
- [39] S Ahern, A Shoshani, K-L Ma, A Choudhary, T Critchlow, S Klasky, V Pascucci, J Ahrens, EW Bethel, H Childs, J Huang, K Joy, Q Koziol, G Lofstead, J Meredith, K Moreland, G Ostrouchov, M Papka, V Vishwanath, M Wolf, N Wright, and K Wu. “Scientific discovery at the exascale. Report from the DOE ASCR 2011 Workshop on Exascale Data Management”. In: *Analysis, and Visualization* (Feb. 2011).
- [40] Jeremy Bejarano, Koushiki Bose, Tyler Brannan, Anita Thomas, Kofi Adragani, Nagaraj K. Neerchal, and George Ostrouchov. *Sampling within k-Means Algorithm to Cluster Large Datasets*. Tech. rep. HPCF–2011–12. (HPCF machines used: tara.) UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2011. URL: <http://www.umbc.edu/hpcreu/2011/projects/team2.html>.
- [41] Kary Myers, George Ostrouchov, and Stephanie Pickle. “Section Highlights JSM 2010: Looks Ahead”. In: *Amstat News* (Nov. 2011), p. 24. URL: <http://magazine.amstat.org/blog/2010/11/01/spes/>.
- [42] George Ostrouchov. “Highlights from Vancouver: SPES Invited and Contributed Program at JSM 2010”. In: *SPES & QP Newsletter* 18.2 (Apr. 2011), p. 9.

- [43] M. Prabhat, S. Byna, C. Paciorek, G. Weber, K. Wu, T. Yopes, M. F. Wehner, G. Ostrouchov, D. Pugmire, R. Strelitz, W. Collins, and W. Bethel. “Pattern Detection and Extreme Value Analysis on Large Climate Data”. In: *AGU Fall Meeting* (Dec. 2011), Poster, C3.
- [44] Matthew G. Bachmann, Ashley D. Dyas, Shelby C. Kilmer, Julian Sass, Andrew Raim, Nagaraj K. Neerchal, Kofi P. Adragani, George Ostrouchov, and Ian F. Thorpe. *Block Cyclic Distribution of Data in pbdR and its Effects on Computational Efficiency*. Tech. rep. HPCF-2013-11. UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2013.
- [45] George Ostrouchov and. “Computation and Volume Rendering of Large-Scale EOF Coherent Modes in Rotating Turbulent Flow Data”. In: *AGU Fall Meeting*. Poster. San Francisco, CA, 2014.
- [46] Pragneshkumar Patel, Drew Schmidt, Wei-Chen Chen, and George Ostrouchov. “High-Level Analytics with R and pbdR on Cray Systems”. In: *Cray User Group Conference Proceedings*. Lugano, Switzerland, 2014.

Software and Guides

- [1] Wei-Chen Chen, George Ostrouchov, Drew Schmidt, Pragneshkumar Patel, and Hao Yu. “A Quick Guide for the pbdMPI Package”. In: *CRAN Vignette* (2012). (Updated January 24, 2016 to version 0.4-3.) URL: <http://cran.r-project.org/web/packages/pbdMPI/vignettes/pbdMPI-guide.pdf>.
- [2] Wei-Chen Chen, George Ostrouchov, Drew Schmidt, Pragneshkumar Patel, and Hao Yu. *pbdMPI: Programming with Big Data-Interface to MPI*. (Current version 0.3-1, published January 21, 2016.) 2012. URL: <http://cran.r-project.org/package=pbdMPI>.
- [3] Wei-Chen Chen, Drew Schmidt, George Ostrouchov, and Pragneshkumar Patel. “A Quick Guide for the pbdSLAP Package”. In: *CRAN Vignette* (2012). (Updated January 22, 2016 to version 0.2-1.) URL: <http://cran.r-project.org/web/packages/pbdSLAP/vignettes/pbdSLAP-guide.pdf>.
- [4] Wei-Chen Chen, Drew Schmidt, George Ostrouchov, and Pragneshkumar Patel. *pbdSLAP: Programming with Big Data - Scalable Linear Algebra Packages*. (Current version 0.2-1, published January 22, 2016.) 2012. URL: <http://cran.r-project.org/package=pbdSLAP>.
- [5] Pragneshkumar Patel, George Ostrouchov, Wei-Chen Chen, Drew Schmidt, and David Pierce. “A Quick Guide for the pbdNCDF4 Package”. In: *R Vignette* (2012). (Updated June 22, 2014 to version 0.1-4.) URL: <http://cran.r-project.org/web/packages/pbdSLAP/vignettes/pbdNCDF4-guide.pdf>.
- [6] Pragneshkumar Patel, George Ostrouchov, Wei-Chen Chen, Drew Schmidt, and David Pierce. *pbdNCDF4: Programming with Big Data - Interface to Parallel Unidata NetCDF4 Format Data Files*. (Current version 0.1-4, published June 22, 2014.) 2012. URL: <http://cran.r-project.org/package=pbdNCDF4>.
- [7] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. “A Quick Guide for the pbdBASE package”. In: *CRAN Vignette* (2012). (Updated January 24, 2016 to version 0.4-3.) URL: <http://cran.r-project.org/web/packages/pbdBASE/vignettes/pbdBASE-guide.pdf>.

- [8] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. “A Quick Guide for the pbdDMAT Package”. In: *CRAN Vignette* (2012). (Updated January 24, 2016 to version 0.4-0.) URL: <http://cran.r-project.org/web/packages/pbdDMAT/vignettes/pbdDMAT-guide.pdf>.
- [9] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. *pbdBASE: Programming with Big Data – Base Wrappers for Distributed Matrices*. (Current version 0.4-3, published January 24, 2016.) 2012. URL: <http://cran.r-project.org/package=pbdBASE>.
- [10] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. *pbdDMAT Programming with Big Data - Distributed Matrix Methods*. (Current version 0.4-0, published January 24, 2016.) 2012. URL: <http://cran.r-project.org/package=pbdDMAT>.
- [11] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. *pbdDEMO: Programming with Big Data - Demonstrations of pbd Packages*. (Current version 0.3-0, published January 24, 2016.) 2013. URL: <http://cran.r-project.org/package=pbdDEMO>.
- [12] Drew Schmidt, Wei-Chen Chen, George Ostrouchov, and Pragneshkumar Patel. *Speaking Serial R with a Parallel Accent: pbdR Package Examples and Demonstrations*. <http://cran.r-project.org/web/packages/pbdDEMO/vignettes/pbdDEMO-guide.pdf>. (Updated January 24, 2016 to version 0.3-0.) 2014.
- [13] Wei-Chen Chen, Drew Schmidt, Christian Heckendorf, George Ostrouchov, Whit Armstrong, Brian Ripley, and R Core Team. *pbdZMQ: Programming with Big Data – Interface to ZeroMQ*. R Package. 2015. URL: <https://github.com/RBigData/pbdZMQ>.
- [14] Drew Schmidt, George Ostrouchov, Wei-Chen Chen, and Ahmed Moustafa. *pbdML: Programming with Big Data – Machine Learning*. R Package. 2016. URL: <https://github.com/RBigData/pbdML>.

Tutorials

- [1] George Ostrouchov. *R Programming Language Workshop in Three Lectures*. Oak Ridge Leadership Computing Facility Training Event, July 29, August 4-5, 2011, Oak Ridge, TN. <https://www.olcf.ornl.gov/training-event/r-programming-language-workshop/>. 2011.
- [2] George Ostrouchov, Pragnesh Patel, and Drew Schmidt. *R Programming Language Workshop 2012*. Oak Ridge Leadership Computing Facility Training Event, June 28-29, 2012, Oak Ridge, TN. <https://www.olcf.ornl.gov/training-event/r-programming-language-workshop-2012/>. 2012.
- [3] George Ostrouchov. *pbdR: Programming with Big Data in R*. Seoul National University, August 17-18, 2013, Seoul, South Korea. http://www.ism.ac.jp/events/2014/meeting0217_18.html. 2013.
- [4] George Ostrouchov and Drew Schmidt. *Programming with Big Data in R*. useR!, The R User Conference 2013, July 9-12, Albacete, Spain. <http://www3.uclm.es/congresos/useR-2013/Tutorials/Ostrouchov.html>. 2013.

- [5] Drew Schmidt and George Ostrouchov. *Introducing R: from Your Laptop to HPC and Big Data*. SC13 The International Conference for High Performance Computing, Networking, Storage and Analysis, November 16-22, 2013, Denver, CO.
http://sc13.supercomputing.org/schedule/event_detail.php?evid=tut165. 2013.
- [6] George Ostrouchov. *pbdR: Programming with Big Data in R*. Institute of Statistical Mathematics, February 17-18, 2014, Tokyo, Japan.
http://www.ism.ac.jp/events/2014/meeting0217_18.html. 2014.
- [7] George Ostrouchov and Drew Schmidt. *Introducing R: from Your Laptop to HPC and Big Data*. SC14 The International Conference for High Performance Computing, Networking, Storage and Analysis, November 16-21, 2014, New Orleans, LA.
http://sc14.supercomputing.org/schedule/event_detail?evid=tut168. Nov. 2014.
- [8] Drew Schmidt and George Ostrouchov. *Programming with Big Data in R*. useR!, The R User Conference 2013, June 30 - July 3, Los Angeles, CA.
http://user2014.stat.ucla.edu/files/tutorial_Schmidt.pdf. June 2014.
- [9] Russel Zaretzki, Michael Gilchrist, Eric Carr, George Ostrouchov, and Brian O'Meara. *Computing in the Cloud: What Every Computational Life Scientist Should Know*. National Institute for Mathematical and Biological Synthesis (MBioS), April 6-8, 2014, Knoxville, TN.
http://www.nimbios.org/tutorials/TT_cloud. Apr. 2014.
- [10] George Ostrouchov. *Analytics session: Programming with Big Data in R*. PRACE Winter School 2015 - HPC Tools for Data Intensive Processing, January 12-15, Ostrava, Czech Republic. <https://events.prace-ri.eu/event/330/session/5>. 2015.
- [11] George Ostrouchov. *Parallel Programming with Big Data in R: pbdR*. 60th World Statistics Congress, July 23-31, 2015, Rio De Janeiro, Brazil.
<http://www.isi2015.org/index.php/isi2015/scientific-programme/short-courses-list>. July 2015.
- [12] Norbert Podhorszki, Quing Gary Liu, Scott Klasky, George Ostrouchov, Dave Pugmire, and Jong Choi. *Data Management, Analysis and Visualization Tools for Data-Intensive Science*. SC15 The International Conference for High Performance Computing, Networking, Storage and Analysis, November 15-20, 2015, Austin, TX. http://sc15.supercomputing.org/schedule/event_detail?evid=tut148. Nov. 2015.
- [13] George Ostrouchov. *Using R for HPC Data Science*. IT4Innovations Czech Supercomputing Center, October 6-7, Ostrava, Czechia.
<http://prace.it4i.cz/en/RforHPC-10-2016>. 2016.

NobleProg, a worldwide training and consultancy organization on Management, IT, Statistics, Programming and Artificial Intelligence has independently designed and offers a three day class around the pbdR project software: “Programming with Big Data in R Training Course”

<http://www.nobleprog.us/programming-big-data-r/training-course>

Invited Presentations

- [1] George Ostrouchov. “Symbolic Givens Reduction in Large Sparse Least Squares Problems”. SIAM Summer Meeting. Seattle, Washington, July 16–20, 1984.

- [2] George Ostrouchov. “Parallel Computing on a Hypercube: An Overview of the Architecture and Some Applications”. 19th Symposium on the Interface of Computer Science and Statistics. Philadelphia, PA, Mar. 8–11, 1987.
- [3] George Ostrouchov. “Sparse Least Squares Computations in Statistical Applications”. Numerical Linear Algebra Year Lectures on Least Squares Computations. University of Tennessee, May 1988.
- [4] George Ostrouchov. “Sparse Matrix Computations in Analysis of Variance”. Department of Statistics, University of Tennessee, Feb. 25, 1988.
- [5] George Ostrouchov. “Sparse Matrix Computations in Analysis of Variance”. Kansas State University, Manhattan, Kansas, Apr. 7, 1988.
- [6] George Ostrouchov. “Sparse Matrix Computations in Analysis of Variance”. Numerical Linear Algebra Year Lectures on Least Squares Computations. University of Tennessee, Apr. 1988.
- [7] George Ostrouchov. “Sparse Matrix Computations in Analysis of Variance”. Third SIAM Conference on Applied Linear Algebra. Madison, Wisconsin, May 23–26, 1988.
- [8] George Ostrouchov. “Statistical Computing on a Hypercube”. Kansas State University, Manhattan, Kansas, Apr. 5, 1988.
- [9] George Ostrouchov. “Statistical Computing on a Hypercube”. 20th Symposium on the Interface of Computing Science and Statistics. Reston, VA, Apr. 21–23, 1988.
- [10] George Ostrouchov and Sallie Keller-McNulty. “Error-Free Computation in Sparse Least Squares”. Third SIAM Conference on Applied Linear Algebra. Madison, Wisconsin, May 23–26, 1988.
- [11] George Ostrouchov. “Spatial Point Process Models and Geophysics for Accurate Remediation Decisions at UXO Sites”. 5th EPA/COE Conceptual Site Model Meeting. Seattle, Washington, Mar. 6–7, 2002.
- [12] George Ostrouchov and Nagiza F. Samatova. “Can Dimension Reduction be Fast and Robust? FastMap and the Convex Hull of Multivariate Data”. Department of Statistics, University of Tennessee, Dec. 6, 2002.
- [13] George Ostrouchov and Nagiza F. Samatova. “Combining Distributed Local Principal Component Analyses into a Global Analysis”. C. Warren Neel Conference on Statistical Data Mining and Knowledge Discovery. Knoxville, Tennessee, June 22–25, 2002.
- [14] George Ostrouchov and Nagiza F. Samatova. “Multivariate Analysis of Massive Distributed Data Sets”. Spring Research Conference on Statistics. Ann Arbor, Michigan, Mar. 20–22, 2002.
- [15] Nagiza F. Samatova, G. Al Geist, and George Ostrouchov. “RACHET: Petascale Distributed Data Analysis Suite”. SPEEDUP Workshop on Distributed Supercomputing Data Intensive Computing. Leukerbad, Valais, Switzerland, Mar. 4–6, 2002.
- [16] George Ostrouchov and Nagiza F. Samatova. “Analysis and Visualization of Massive Simulation Data Sets at ORNL”. Spring Research Conference on Statistics. Dayton, Ohio, June 4–6, 2003.

- [17] George Ostrouchov. “Data from long-running simulations on high performance computers”. Symposium on Statistical Issues in Data Acquisition, The National Academies Board of Mathematical Sciences and Their Applications, Committee on Applied and Theoretical Statistics. Berkeley, CA, July 16, 2004.
- [18] George Ostrouchov. “Data Intensive Analysis and Visualization Projects at ORNL”. National Institute of Standards. Washington, DC, May 6, 2004.
- [19] George Ostrouchov. “Uncertainty Quantification: Barriers and Challenges for Multiscale Mathematics”. DOE Multiscale Mathematics Workshop. Alexandria, VA, May 3–5, 2004.
- [20] George Ostrouchov and Nagiza F. Samatova. “Toward Fast and Robust Dimension Reduction: FastMap and the Convex Hull of Multivariate Data”. COMPSTAT 2004, 16th Symposium of the International Association of Statistical Computing. Prague, Czech, Aug. 23, 2004–Aug. 27, 2008.
- [21] Nagiza F. Samatova, G.-X. Yu, Byung-Hoon Park, Al Geist, and George Ostrouchov. “From Genomics to Functional Proteomics: In silico Approach”. SIAM Conference on Parallel Processing for Scientific Computing. San Francisco, CA, Feb. 25–27, 2004.
- [22] 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”. Fifth Symposium on Understanding Complex Systems. University of Illinois at Urbana-Champaign, May 2005.
- [23] George Ostrouchov. “From Distance-Based Dimension Reduction to Robust Statistics and Matrix Computation”. Department of Statistics, University of Georgia, Oct. 2005.
- [24] George Ostrouchov. “Data-Parallel Analysis and Graphics with R”. DOE Computer Graphics Forum. Duck, NC, Apr. 28–30, 2008.
- [25] George Ostrouchov. “Stalking the Interactive Terabyte with R: Data-Parallel Statistical Computing”. Department of Statistics, Operations, and Management Science, University of Tennessee, Apr. 29, 2008.
- [26] George Ostrouchov. “Data Analysis for HPC Resilience: A Perspective from Statistics”. National HPC Workshop on Resilience. Washington, DC, Aug. 12–14, 2009.
- [27] George Ostrouchov. “Fast Simultaneous Dimension Reduction and Clustering: Viewing Data from Extremes”. Spring Research Conference on Statistics in Industry and Technology. Vancouver, British Columbia, May 27–29, 2009.
- [28] George Ostrouchov. “Statistics and High Performance Computing: Petabytes of Data and Millions of Processors”. Conference CELEBRATING 75 Years of STATISTICS at Iowa State. Ames, Iowa, June 3–5, 2009.
- [29] George Ostrouchov, Thomas J. Naughton, and Stephen L. Scott. “Reliability in Supercomputing: A Million Processors Cooperating to Solve One Problem”. Joint Statistical Meetings. Washington, DC, Aug. 2–6, 2009.
- [30] George Ostrouchov. “Data, Statistics, and High Performance Computing”. International Conference on Software Technology and Engineering. Keynote. San Juan, Puerto Rico, Oct. 3–5, 2010.

- [31] George Ostrouchov. “Data-Parallel Statistical Computing: a Model Based Clustering Example”. Joint Research Conference on Statistics in Quality, Industry, and Technology. Gaithersburg, MD, May 25–27, 2010.
- [32] George Ostrouchov. “Parallel statistical computing: Are we embracing the scalable concurrency revolution?” Joint Statistical Meetings. Vancouver, Canada, Aug. 1–6, 2010.
- [33] George Ostrouchov. “Statistics and High Performance Computing: Focus on Data”. University of Maryland Baltimore County, Baltimore, MD, Dec. 3, 2010.
- [34] George Ostrouchov. “Data and Statistics in High Performance Computing”. International Conference on Software and Intelligent Information (ICSII 2011). Keynote. San Juan, Puerto Rico, Oct. 22, 2011. URL: <http://www.iacsit.org>.
- [35] George Ostrouchov. “Tools for Sampling Large Distributed Data”. REU 2011. University of Maryland Baltimore County, Baltimore, MD, June 30, 2011.
- [36] George Ostrouchov. “Data Parallel Statistical Computing and R”. Conference on Data Analysis. Santa Fe, NM, Feb. 29–Mar. 2, 2012.
- [37] George Ostrouchov. “Statistical Computing with Big Data”. SAMSI Massive Data Working Group. WebEx, Nov. 13, 2012.
- [38] Sean Ahern, David Pugmire, George Ostrouchov, and Scott Klasky. “Data Analysis and Visualization of Big Data from HPC”. SOS17 Conference. Jekyll Island, GA, Mar. 25–28, 2013.
- [39] Wei-Chen Chen and George Ostrouchov. “Distributed Parallel Clustering in R with Large Data”. 59th World Statistics Congress. Hong Kong, PRC, Aug. 25–30, 2013.
- [40] George Ostrouchov. “Optimizing Data Layout in Distributed Parallel Statistical Computing”. REU Site: Interdisciplinary Program in High Performance Computing. University of Maryland, Baltimore County, MD, June 20, 2013.
- [41] George Ostrouchov. “pbdR: Programming with Big Data in R”. Department of Biostatistics. Virginia Commonwealth University, Richmond, VA, Nov. 15, 2013.
- [42] George Ostrouchov, Wei-Chen Chen, Drew Schmidt, and Pragneshkumar Patel. “Programming with Big Data in R”. Joint Statistical Meetings. Montreal, Canada, Aug. 4–8, 2013.
- [43] George Ostrouchov and Drew Schmidt. “pbdR: Programming with big data in R”. Workshop on Processing and Analysis of Very Large Data Sets. Knoxville, TN, Aug. 6–8, 2013.
- [44] George Ostrouchov, Drew Schmidt, Wei-Chen Chen, and Pragneshkumar Patel. “Bringing Exploratory Analytics to Big Data on Leadership Class HPC Platforms”. SIAM Conference on Computational Science and Engineering. Boston, MA, Feb. 25–Mar. 1, 2013.
- [45] George Ostrouchov, Drew Schmidt, Wei-Chen Chen, and Pragneshkumar Patel. “Combining R with Scalable Libraries to Get the Best of Both for Big Data”. In: *IASC Satellite Conference for the 59th ISI WSC & the 8th Conference of IASC-ARS*. 2013, pp. 85–90.
- [46] George Ostrouchov. “Elevating R to Supercomputers with Scalable Libraries”. Institute of Statistical Mathematics. Tokyo, Japan, Feb. 17, 2014.

- [47] George Ostrouchov. “pbdR: Bringing R Analytics to Large Distributed Architectures”. Swiss Supercomputing Center. Lugano, Switzerland, June 20, 2014.
- [48] George Ostrouchov. “Taking R to Big Platforms and Supercomputers with pbdR”. Bioconductor Developer Day Keynote. Boston, MA, July 30, 2014.
- [49] George Ostrouchov. “pbdR: A Sustainable Path for Scalable Statistical Computing”. Workshop on Distributed Computing in R. Palo Alto, CA, Jan. 26–27, 2015.
- [50] George Ostrouchov. “pbdR: Harnessing HPC Research for Parallel Computing with R”. Institute of Statistical Mathematics High Performance Computing Conference and HPC on R Workshop. Tokyo, Japan, Oct. 9–12, 2015.
- [51] George Ostrouchov. “R and pbdR: An Overview”. Faculty of Mathematics and Physics, Charles University. Praha, Czech Republic, Jan. 16, 2015.
- [52] George Ostrouchov. “Taking R from Your Laptop to a Cluster Computer”. Spring Restarch Conference. May 20–22, 2015.
- [53] George Ostrouchov. “Taking R from Your Laptop to a Cluster Computer”. Department of Mathematical Sciences, University of Memphis. Memphis, Tennessee, Dec. 4, 2015.
- [54] George Ostrouchov. “Taking Statistical Computing to Parallel Platforms with R and pbdR”. Department of Business Analytics and Statistics, University of Tennessee, Feb. 27, 2015.
- [55] George Ostrouchov. “Programming with Big Data in R”. Future Trends in Nuclear Physics Computing. Thomas Jefferson National Accelerator Facility, Newport News, VA, Mar. 16–18, 2016.
- [56] George Ostrouchov. “Statistical Computing with R on Distributed and Multicore Platforms”. 44th Annual Meeting of the Statistical Society of Canada. Brock University, St. Catharines, Ontario, Canada, May 29–June 1, 2016.

Reviewer Activities

Analytical Chemistry

Bit

Communications in Statistics: Simulation and Computation

Computational Statistics and Data Analysis

CRC Press, Taylor & Francis Group, book proposal reviews

Department of Energy, Office of Science, Proposals

Handbook of Parallel Computing and Statistics

High Performance Computing (HPC) Review

IEEE Transactions on Systems Man and Cybernetics

Jack Youden Prize for best expository paper in Technometrics

Journal of Computational and Graphical Statistics

Journal of the American Statistical Association

Journal of Statistical Computation and Simulation

Journal of Statistical Software

National Science Foundation, Proposals

Natural Sciences and Engineering Research Council of Canada, Proposals

SIAM Journal on Scientific and Statistical Computing
SIAM/ASA Journal on Uncertainty Quantification
Statistical Analysis and Data Mining
Stochastic Environmental Research and Risk Assessment
Technometrics
The American Statistician

Languages

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