

Jun Yan

Department of Statistics
University of Connecticut
215 Glenbrook Rd. U-4120
Storrs, CT 06269

Phone: (860) 486-3416
Fax: (860) 486-4113
Email: jun.yan@uconn.edu
Web: <https://statcomp.org/>

Degrees

Ph.D., 2003, Statistics, University of Wisconsin, Madison, WI
M.A., 1998, Economics, University of Miami, Miami, FL
M.Econ., 1996, Statistics, Renmin University of China, Beijing, China
B.Econ., 1993, Statistics, Renmin University of China, Beijing, China

Research Interests

Social Network, Spatial Extremes, Dynamic Survival Models, Multivariate Dependence, Clustered Data Analysis
Statistical Computing, Biostatistics, Environmental Applications, Public Health Applications, Sports Analytics

Positions

Professor, Department of Statistics, University of Connecticut, 2015–Present.
Affiliated Faculty, The Connecticut Institute for the Brain and Cognitive Sciences, 2023–Present.
Research Fellow, Center for Population Health, University of Connecticut Health Center, 2007–Present.
Associate Professor, Department of Statistics, University of Connecticut, 2010–2015.
Affiliated Faculty, Center for Environmental Sciences and Engineering, University of Connecticut, 2007–Present.
Assistant Professor, Department of Statistics, University of Connecticut, 2007–2010.
Assistant Professor, Department of Statistics and Actuarial Science, The University of Iowa, 2003–2007.
Statistical Consultant, CALS Statistical Consulting Lab, University of Wisconsin–Madison, 2001–2003.

Honors and Awards

Elected Member, 2023, Connecticut Academy of Science and Engineering
Fellow, 2022, Institute of Mathematical Statistics
Fellow, 2017, American Statistical Association
Elected Member, 2014, International Statistical Institute
Outstanding Service Award, 2014, International Chinese Statistical Association

Professional Memberships

American Association for the Advancement of Science (AAAS)
American Geophysical Union (AGU)
American Statistical Association (ASA)
International Chinese Statistical Association (ICSA)
Institute of Mathematical Statistics (IMS)

Services and Outreaches

Editorial

- 2020–present: Editor-in-Chief *Journal of Data Science*.
- 2020–2021: Associate editor, *Brazilian Journal of Probability and Statistics*.

3. 2020: Co-Guest Editor, Special Issue on “Data Science in Action in Response to the Outbreak of COVID-19”, *Journal of Data Science*.
4. 2014–2020: Associate editor, *Ecological and Environmental Statistics*.
5. 2014: Associate editor, *STEM Forums in Statistics*.

Professional Society Services

1. 2022–2023: ASA Committee on Data Science and AI
2. 2021–2023: Chair-Elect/Chair/Past-Chair of the ASA Section on Statistical Computing.
3. 2020: Co-Chair, ASA–Journal of Data Science webinar series on “Data Science in Action in Response to the Outbreak of COVID-19”, April 17 — July 24, 2020.
4. 2018–2021: IMS Committee to Select Editors.
5. 2017–2020: Awards Chair, ASA Section on Statistical Computing and Section on Statistical Graphics.
6. 2016–2020, Editor, LIDA Newsletter, The official newsletter of the Lifetime Data Analysis Section, American Statistical Association.
7. 2015–2016: Chair, IT Committee, International Chinese Statistical Association.
8. 2012–2014: Editor-in-Chief, International Chinese Statistical Association Bulletin.

Conference Organization

1. 2023: Lead Organizer, Banff International Research Station Workshop 24w5284: Statistical Aspects of Trustworthy Machine Learning (with Stephanie Hicks, Keegan Korthauer, Xiaotong Shen, and Helen Zhang)
2. 2019–present: Chair, Organizing Committee, Annual UConn Sports Analytics Symposium.
3. 2022: Chair, Organizing Committee, Excellence in Statistical Science: Celebrating the 60th Anniversary of UConn Statistics.
4. 2022: Chair, Steering Committee, ASA Section on Statistical Computing Mini-Symposium: Statistical Computing in Action.
5. 2022: Program Committee, 2023 ICSA Applied Statistics Symposium.
6. 2020: Chair, Organizing Committee, The 3rd International Forum on Statistical Science and Big Data, June 13–15, 2020, School of Statistics, Shanxi University of Finance and Economics.
7. 2019: Chair, Organizing Committee, 33rd New England Statistics Symposium: Statistical Data Science in Action.
8. 2018: Organizing Committee, The 8th International Statistics Forum, Renmin University of China.
9. 2018: Co-Chair, Short-course Committee, 2018 ICSA Applied Statistics Symposium.
10. 2017: Co-Chair, Local Organizing Committee, Conference on Lifetime Data Science: Data Science, Precision Medicine and Risk Analysis with Lifetime Data at University of Connecticut, May 25–27, 2017, Sponsored by the Lifetime Data Analysis Interest Group, American Statistical Association.
11. 2015: Chair, Poster Committee, 2015 ICSA Applied Statistics Symposium

Grant Reviews

1. NSF DMS CDSE-MSS panel review, 2016.
2. NSF DMS Statistics panel review, 2013.

Journal Reviews

1. *Annals of Applied Statistics*; *Annals of Statistics*; *Applied Stochastic Models in Business and Industry*; *Bernoulli*; *Biometrics*; *Biometrika*; *Bioscience*; *Canadian Journal of Statistics*; *Computational Statistics*; *Computational Statistics and Data Analysis*; *Econometrics and Statistics*; *Environmental and Ecological Statistics*; *IEEE Transactions on Neural Networks*; *International Journal of Forecasting*; *Journal of the American Statistical Association*; *Journal of Data Science*; *Journal of Multivariate Statistics*; *Journal of Royal Statistical Society Series B*; *Journal of Royal Statistical Society Series C*; *Journal of Statistical Software*; *R Journal*; *Statistics and Probability Letters*; *Statistica Sinica*; *Statistics in Medicine*; *Stochastic Environmental Research and Risk Assessment*; among others.
2. Statistical Reviewer: *Journal of Applied Physiology*; *Advances in Physiology Education*.

Departmental Service

1. 2023–2024: Promotion, Tenure, and Reappointment Committee (chair); X + 3 + 1 Admission (chair); Search Committee for Assistant Professors; Computer Committee; Graduate Examination Committee.
2. 2022–2023: Promotion, Tenure, and Reappointment Committee (chair); 10-Year Strategic Plan Committee (chair); Organizing Committee of the 60th Anniversary Celebration (chair); X + 3 + 1 Admission (chair); Search Committee for Assistant Professors; Computer Committee; Graduate Examination Committee; SET+ Faculty Teaching Evaluations.
3. 2021–2022: Promotion, Tenure, and Reappointment Committee (chair); 10-Year Strategic Plan Committee (chair); Organizing Committee of the 60th Anniversary Celebration (chair); X + 3 + 1 Admission (chair); Search Committee for Assistant Professor in Data Science; Computer Committee; Graduate Examination Committee.
4. 2020–2021: Data Science Program (chair), Computer Committee (chair), Graduate Examination, X + 3 + 1 Admission (Chair), Promotion, Tenure, and Reappointment Committee.
5. 2019–2020: Data Science Program (chair), Computer Committee (chair), Graduate Examination, X + 3 + 1 Admission, Promotion, Tenure, and Reappointment Committee.
6. 2018–2019: Search Committee (chair), Data Science Program (chair), Computer Committee (chair), Graduate Examination, Distinguished Statistician Series, 3 + 1 Admission, Promotion, Tenure, and Reappointment Committee.
7. 2017–2018: Data Science Program (chair), Computer Committee (chair), Graduate Examination, Distinguished Statistician Series, 3 + 1 Admission, Graduate Admission Committee, Promotion, Tenure, and Reappointment Committee.
8. 2016–2017: Computer Committee (chair), Program Review Committee, Graduate Students and Distinguished Alumni Awards Committee, X + 3 + 1 Admission, New England Statistics Symposium (chair), New England Statistics Society, Faculty Search Committee, Graduate Admission Committee, Promotion, Tenure, and Reappointment Committee.
9. 2015–2016: Computer Committee (chair), Admission Committee, Department Head Search Committee, Promotion, Tenure, and Reappointment Committee.
10. 2014–2015: Computer Committee (chair), Admission Committee, New England Statistics Symposium Committee.
11. 2013–2014: Computer Committee (chair), Admission Committee.
12. 2012–2013: Computer Committee (chair), Associate/Full professor search committee, Department's 50th Year Celebration committee, New England Statistics Symposium Committee, Admission Committee.
13. 2011–2012: Computer Committee (chair), Admissions Committee, Biostatistics Search Committee, Department's 50th Year Celebration committee, Biostatistics Program Development Committee.
14. 2010–2011: Colloquium Committee (chair), Library/Tech Reports Committee, Search Committee, Biostatistics Program Development Committee, New England Statistics Symposium Committee.
15. 2008–2010: Library/Tech Reports Committee, Biostatistics Program Development Committee, Social Committee.
16. 2006–2007: Colloquium Committee (chair), M.S. Exam – Minor Committee (chair), Computer Committee.
17. 2005–2006: Colloquium Committee, M.S. Exam – Minor Committee, Search Committee, and Social Committee.
18. 2004–2005: M.S. Exam – Minor Committee, Search Committee, and Social Committee (chair).
19. 2003–2004: Computer Committee, M.S. Exam – Minor Committee, and Social Committee.

Media Coverage

1. UConn Today (2023): UConn Students to Showcase Work at NYC Open Data Week.
2. ASA Member News (2023): Data is My Job.
3. AMStat News (2023): UConn Department of Statistics Celebrates 60th Anniversary.
4. AMStat News (2023): UConn Sports Analytics Symposium Boasts New Features.
5. UConn Today (2022): From Wyoming Mountains to Connecticut Forests, Tracking Feline Apex Predators.
6. Methods Blog (Methods in Ecology and Evolution) (2022): Revealing the hidden lives of cryptic mountain lions using GPS data and a Moving-Resting Motion model.
7. AMStat News (2022): UConn Sports Analytics Symposium a Home Run.
8. UConn Today (2020): UConn Holds Second Annual Sports Analytics Conference.

9. UConn Today (2020): CLAS Faculty and Students Shifting Work to COVID-19.
10. UConn Today (2019): UConn Hosts New Sports Analytics Symposium.

Teaching Experience

University of Connecticut

1. Data Science: One-week intensive Pre-College Summer course for high schoolers; Summer 2021; Summer 2022; Summer 2023.
2. Special topics: Advanced Data Manipulation and Analysis with Python (4185), 1-credit undergraduate level: Fall 2022; Spring 2023; Fall 2023.
3. Undergraduate Seminar/Investigation of Special Topics: Statistical Writing (3494W/5095): 1-credit graduate/undergraduate statistical literacy requirement; Fall 2022.
4. Introduction to Data Science (5255/3255): 3-credit graduate/undergraduate level; Fall 2021; Spring 2022; Spring 2023.
5. Biostatistics (5625/4625), 3-credit graduate/undergraduate level; Spring 2021.
6. Data Science in Action (6494), 3-credit seminar; Spring 2018 (co-taught with Kun Chen and Elizabeth Schifano); Spring 2019.
7. Statistical Computing (5361), 3-credit graduate level; Spring 2018; Fall 2018; Fall 2020.
8. Applied Statistics II (5605), 3-credit graduate level (required for qualifying exam); Spring 2017.
9. Multivariate Modeling with Copulas (6494), 3-credit graduate level; Fall 2016.
10. Advanced Statistical Computing (6494), 3-credit graduate level; Spring 2014/2015/2016.
11. Applied Longitudinal Data Analysis (6494), 3-credit graduate level; Fall 2014.
12. Mathematical Statistics I/II (5585/5685), 3-credit graduate level sequence (required for qualifying exam); Fall 2011 — Spring 2013, Fall 2007 — Spring 2010, Fall 2015, Fall 2017.
13. Data Analysis Using R (6494), 3-credit graduate level; Spring 2011.
14. Environmental Statistics (6494), 3-credit graduate level; Fall 2010.
15. Statistical Methods (220), 3-credit undergraduate level; Spring 2008.

University of Iowa

1. Probability and Stochastic Processes I (22s:195), Fall 2006.
2. Mathematical Statistics II (22s:154), Spring 2007, Spring 2006, Spring 2004.
3. Mathematical Statistics I (22s:153), Fall 2005, Fall 2004, Fall 2003.
4. Applied Time Series Analysis (22s:156), Spring 2007, Spring 2005.

Students

Ph.D. Thesis Advisees

1. Shiyang Xiao, Ph.D. expected 2027: Network analytics.
2. Zefang Min, Ph.D. expected 2026: Time series and causal inference.
3. Xiaomin Lu, Ph.D. expected 2025: Survival analysis and artificial intelligence.
4. Sydney Louit, Ph.D. expected 2025: Social network clustering.
5. Zhenyu Xu, Ph.D. expected 2025: Generalized estimating equations.
6. Yingfa Xie, Ph.D. expected 2024: Recurrent event modeling.
7. Jun Jin, Ph.D. expected 2024 (joint with Kun Chen): Machine learning and health analytics.
8. Surya Eada, Ph.D. expected 2024 (joint with Vladimir Pozdnyakov): Stochastic processes governed by telegraph processes.
9. Lucas Godoy, Ph.D. expected 2024: Spatial regression with errors-in-variables.

10. Yelie Yuan, Ph.D. 2023: On assortativity of weighted directed networks. Placement: Consumer and Community Banking Risk Program Associate, JP Morgan.
11. Zehan Yang, Ph.D. 2023 (joint with HaiYing Wang): Optimal subsampling methods for massive survival data using accelerated failure time models. Placement: Mathematical Statistician, US Food and Drug Administration.
12. Jackson Lautier, Ph.D. 2023 (joint with Vladimir Pozdnyakov): Essays on discrete-time survival analysis with applications to securitization and consumer finance. Placement: Assistant Professor, Bentley University.
13. Sai Ma, Ph.D. 2022: Optimal fingerprinting with estimating equations. Placement: Statistician, Vertex Pharmaceuticals.
14. Abby Lau, Ph.D. 2022: Extreme value modeling with errors-in-variables in detection and attribution of changes in climate extremes. Placement: Postdoc researcher, University of Pennsylvania.
15. Yan Li, Ph.D. 2021 (joint with Kun Chen): Amalgamation-based statistical learning for compositional data. Placement: Postdoc researcher, University of Michigan.
16. Jieying Jiao, Ph.D. 2020: On Bayesian methods for spatial point processes. Placement: Senior Consultant, Travelers Insurance.
17. Chaoran Hu, Ph.D. 2020 (joint with Vladimir Pozdnyakov): On Brownian motion governed by telegraph process. Placement: Senior Statistician, Ely Lilly.
18. Wenjie Wang, Ph.D. 2019 (joint with Kun Chen): Integrated survival analysis with application to suicide risk. Placement: Senior Statistician, Eli Lilly.
19. Yishu Xue, Ph.D. 2019 (joint with Elizabeth Schifano): Diagnostic methods for big survival data. Placement: Senior Consultant, Travelers Insurance.
20. Greg Vaughan, Ph.D. 2017 (joint with Kun Chen): Stagewise generalized estimating equations. Placement: Assistant Professor, Bentley University.
21. Yujing Jiang, Ph.D. 2017: Marginal score equations for spatial extremes with latent signals and applications to fingerprinting changes in climate extremes. Placement: Postdoc researcher, Colorado State University.
22. Brian Bader, Ph.D. 2016: Automated, efficient, and practical extreme value analysis with environmental applications. Placement: Statistician, KPMG.
23. Chun Wang, Ph.D. 2016 (joint with Elizabeth Schifano): Online updating methods for big data streams. Placement: Senior Analyst, Liberty Mutual Insurance.
24. Zhuo Wang, Ph.D. 2015: Estimating equations for spatial extremes with application to detection and attribution analysis of changes in climate extremes. Placement: Assistant Professor, Shenzhen University, China.
25. Hongwei Shang: Ph.D. 2013: A Two-step estimation procedure and a goodness-of-fit test for spatial extremes models. Placement: Statistician, HP Analytic Lab.
26. Sy Han (Steven) Chiou, Ph.D. 2013 (joint with Sangwook Kang): Statistical Methods and Computing for Semiparametric Accelerated Failure Time Model with Induced Smoothing. Placement: Assistant Professor, Department of Mathematics and Statistics, University of Minnesota Duluth.
27. Xiaojing Wang: Ph.D. 2011: Statistical inferences for interval censored data. Placement: Quantitative Analyst, Google New York.
28. Marcos Prates, Ph.D. 2011 (joint with Dipak Dey): Link specification and spatial dependence for generalized linear mixed models. Placement: Assistant Professor, Departamento de Estatística, Universidade Federal de Minas Gerais, Brazil.

Undergraduate Thesis Advisees

1. Shannon Yeung (2023): Varying effects of short term interest rates. Placement: GMI Analyst, BNY Mellon.
2. Owen Fiore (2023): Was Devon Allen unjustly disqualified at the 2022 World Track and Field Championships? Placement: MS Student in Data Science, University of Connecticut.
3. Pranav Tavildar (2023): Sentiment analysis of twitter in relation to fossil fuel stock Prices. Placement: MS Student in Data Analytics, Georgia Institute of Technology.
4. Samuel Hughes (2022): Statistical evaluation of field hockey penalty corners. Placement: MS Student in Data Science, Northeastern University.

5. Anthony Zeimbekakis (2022): On misuses of the Kolmogorov—Smirnov test for one-sample goodness-of-fit. Placement: Analyst Development Program, Electric Insurance Company.
6. Brian Krikorian (2022): Points above replacement: A new NBA metric to evaluate player performance. Placement: Success Metrics Intern, UMass Chan Medical School.
7. Justin Franklin (2021): Comparison of fraud detection methods: A case study with insurance. Placement: Software Engineer, Travelers.
8. Andrew Tammaro (2021): NFL front office analytics with R. Placement: Corporate Strategy and Data Analytics Intern, 1BusinessWorld.
9. Michael Price (2021): The effects of the NBA COVID bubble on the NBA playoffs: A case study for home-court advantage. Placement: MS program in Applied Statistics, University of Delaware.
10. Dylan Barrett (2020): When is the best time to steal bases? Placement: Customer Operations Agent, FanDuel.
11. Taaj Cheema (2020): An analysis of Oliver's four factors in the golden age of NBA offense. Placement: Data Scientist, IBM.
12. Jack Schooley (2020): Predicting the outcomes of soccer games. Placement: MS Student in Data Science, MIT.
13. Thomas Kennon (2018): Finding an ultimate limit for an NBA player's shooting percentage. Placement: Data Engineer, The Hartford.
14. Junghi Kim (2010, Joint with Evarist Giné and Nalini Ravishanker): A close look at marginal Cox model and conditional Cox model with application of recurrent gap times. Placement: Ph.D. Student in Biostatistics, University of Minnesota.

Teaching Accomplishments

Student Awards

1. Lucus Godoy (2024): Honorable Mention, Student Paper Award, Section on Statistics and the Environment, American Statistical Association.
2. Yingfa Xie (2024): Student Paper Award, Section on Lifetime Data Science, American Statistical Association.
3. Zehan Yang (2023): Honorable Mention, Student Paper Award, 2023 Applied Statistics Symposium, International Chinese Statistical Association.
4. Yelie Yuan (2023): Honorable Mention, John M. Chambers Statistical Software Award, American Statistical Association.
5. Jackson Lautier (2022): Student Paper Award, Special Conference Celebrating the 60th Anniversary of UConn Department of Statistics.
6. Zehan Yang (2022): Student Paper Award, Section on Lifetime Data Science, American Statistical Association.
7. Jackson Lautier (2022): Student Paper Award, Risk Analysis Section, American Statistical Association.
8. Jun Jin (2022): Honorable Mention, Student Paper Award, Risk Analysis Section, American Statistical Association.
9. Chaoran Hu (2020): Student Paper Award, Statistical Computing and Statistical Graphics Sections, American Statistical Association.
10. Yan Li (2020): ENAR Distinguished Student Paper Award, International Biometric Society.
11. Yan Li (2019): IBM Student Paper Award, 33rd New England Statistics Symposium.
12. Yishu Xue (2019): ENAR Distinguished Student Paper Award, International Biometric Society.
13. Wenjie Wang (2017): IBM Student Paper Award, 31st New England Statistics Symposium.
14. Gregory Vaughan (2017): Student Paper Award, Mental Health Section, American Statistical Association.
15. Steven Chiou (2012): Student Paper Award, Applied Statistics Symposium, International Chinese Statistical Association.
16. Yung-wei Chen (2012): Student Paper Award, Social Statistics, Government Statistics, and Survey Research Methods Sections, American Statistical Association.
17. Xiaojing Wang (2010): IBM Student Paper Award, 24th New England Statistics Symposium.

Teaching Highlights

1. Undergraduate advisees have published journal articles since 2022.
2. Advised the Undergraduate Data Science Club as faculty advisor since 2019.
3. Co-taught “Data Science” to pre-college students in the UConn Pre-College Summer Program since 2020.
4. Developed new courses for the data science program: Statistical Data Science in Action; Introduction to Data Science; Spatiotemporal Statistics.
5. Advised undergraduates in data science in UConn Individualized Major Program.
6. Mentored statistics students to develop software packages in R.
7. Taught students in ecology and evolutionary biology how to do data analysis with R.
8. Taught the first-year graduate required sequence of mathematical statistics effectively.

Grants

External

1. NSF DMS2219336, 09/01/2022 — 08/31/2025: Conference: UConn Sports Analytics Symposium: Engaging Students into Data Science. \$49,986. PI: Jun Yan; Co-PIs: Laura Burton, Kun Chen, Robert Huggins, Elizabeth Schifano.
2. NSF DMS2210735: 08/01/2022 — 07/31/2025: Models and Inferences for Heterogeneous Interaction Patterns in Social Networks. \$360,000. PI: Jun Yan; Co-PI: Xianyang Zhang.
3. NSF CC19325716, 08/01/2019 — 07/31/2021: CC* Compute: Shared Computing Infrastructure for Large-scale Science Problems. \$400,000. PI: Richard T. Jones, Co-PIs: Vernon Cormier, Kyungseon Joo, Cara D. Battersby, and Jun Yan
4. NSF DMS1521730, 2015/09/01 — 2018/8/31: Fingerprint Methods for Detection and Attribution of Changes in Climate Extremes with Spatial Estimating Equations. \$100,000. PI: Jun Yan.
5. NSF DMS1209022, 08/15/2012 — 07/31/2015: Statistical Inferences, Computing, and Applications for Semiparametric Accelerated Failure Time Models. \$130,000. PI: Jun Yan. Co-PI: Sangwook Kang.
6. University of Wisconsin (NIH RO1 subcontract, PI: Hui-Chuan Lai), 09/01/2011 — 08/31/2016: Newborn Screening, Malnutrition and Lung Disease in Children with Cystic Fibrosis. \$37,284. PI: Jun Yan.
7. NASA NNX10AG77G, 06/01/2010 — 05/31/2013: Testing the Suitability of Satellite Precipitation Products for Hydrological Modeling at Multiple Scales across the Blue Nile Basin. \$308,845. PI: Mekonnen Gebremichael; Co-PI Jun Yan.
8. NOAA NA10NWS4680004, 05/01/2010 — 04/30/2012: National Weather Service (NWS) Collaborative Science Technology, and Applied Research (CSTAR) — A New Statistical Model of Streamflow Forecast Error. \$148,574. PI: Mekonnen Gebremichael; Co-PI Jun Yan.
9. NSF DMS0805965, 07/01/2008 — 06/30/2011: Unified Dynamic Modeling of Event Times with Semiparametric Profile Estimating Functions: Theory, Computing, and Applications. \$150,000. PI: Jun Yan.
10. NSF DMS0618883, 07/15/2006 — 07/14/2007: Statistical Computing Research Environments (SCREMS). \$95,000. PI: Mary Kathryn Cowles; Co-PIs: John Geweke, Jian Huang, Luke Tierney, and Jun Yan.

Internal

1. January 2017 – December 2019, Innovative Education in Science, College of Liberal Arts and Sciences, University of Connecticut — Data Science Lab: Real World Data Science Problems Meet Future Data Scientists (with Kun Chen and Elizabeth Schifano).
2. April 2015 – March 2016, Research Excellence Program, University of Connecticut — Statistical Methods and Computing for Detection and Attribution of Changes in Climate Extremes.
3. January 2011 – May 2011, Multidisciplinary Environmental Research Award, Center for Environmental Sciences and Engineering, University of Connecticut — A Constrained Stochastic Model for Animal Movement Data with Application to Deer Home Range (with Thomas Meyer).
4. January 2010 – December 2010, Faculty Large Grant, University of Connecticut — Semiparametric Methods for Spatial Extremes with Application to Extremal Peak Flow in Connecticut.

5. January 2008 – December 2008, Faculty Large Grant, University of Connecticut — Partly Functional Temporal Process Regression with Semiparametric Profile Estimating Functions: Theory and Application.
6. January 2008 – May 2008, Multidisciplinary Environmental Research Award, Center for Environmental Sciences and Engineering, University of Connecticut — A Hierarchical Spatio-Temporal Model for Terrestrial Snails Abundances in a Tropical Forest (with Michael Willig).
7. July 2006 – June 2007, Mathematical & Physical Sciences Funding Program (MPSFP), University of Iowa — Partly Functional Temporal Process Regression.
8. January 2004 – December 2004, Mathematical & Physical Sciences Funding Program (MPSFP), University of Iowa — Nonparametric Inference for Nonstationary Stochastic Processes.

Publications (*Student or postdoc supervisee)

Books and Edited Volumes

1. Hofert, M., Kojadinovic, I., Mächler, M., and Yan, J. (2018): *Elements of Copula Modeling with R*. Springer.
2. Dey, D. K. and Yan, J. (eds.) (2015): *Extreme Value Modeling and Risk Analysis: Methods and Applications*. Chapman & Hall/CRC.

Book Chapters

1. Yin, F., *Jiao, J., Yan, J., and Hu, G.. (2022): Bayesian nonparametric estimation for point processes with spatial homogeneity: A spatial analysis of NBA shot locations. *Proceedings of the 39th International Conference on Machine Learning*. 162: 25523–25551.
2. *Vaughan, G., Aseltine, R., Chiou, S., and Yan, J. (2016): An alarm system for flu outbreaks using Google Flu Trend data. In J. Lin, B. Wang, X. Hu, K. Chen, and R. Liu (eds.) *Statistical Applications from Clinical Trials and Personalized Medicine to Finance and Business Analytics*, pp.293–304, Springer.
3. *Chiou, S., Kang, S., and Yan, J. (2015): Change point analysis of top baseball batting average. In D. K. Dey and J. Yan (eds.) *Extreme Value Modeling and Risk Analysis: Methods and Applications*, pp.493–504, Chapman & Hall/CRC.
4. Dey, D. K., Roy, D., and Yan, J. (2015): Univariate extreme value analysis. In D. K. Dey and J. Yan (eds.), *Extreme Value Modeling and Risk Analysis: Methods and Applications*, pp.1–22, Chapman & Hall/CRC.
5. *Jiang, Y., Dey, D. K., and Yan, J. (2015): Multivariate extreme value analysis. In D. K. Dey and J. Yan (eds.), *Extreme Value Modeling and Risk Analysis: Methods and Applications*, pp.23–39, Chapman & Hall/CRC.
6. *Wang, X., Sinha, A., Yan, J., and Chen, M.-H. (2012): Bayesian inference of interval-censored survival data. In D.-G. Chen, J. Sun, and K. E. Peace (eds.), *Interval-Censored Time-to-Event Data: Methods and Applications*, pp.167–196, Chapman & Hall/CRC.
7. Yan, J. (2006): Multivariate modeling with copulas and engineering applications. In H. Pham (ed.), *Handbook of Engineering Statistics*, pp. 973–990, Springer.

Refereed Journal Articles

1. *Yuan, Y., Yan, J., and Zhang, P. (2024+): A strength and sparsity preserving algorithm for generating weighted, directed networks with predetermined assortativity. *Physica A: Statistical Mechanics and Its Applications*. Forthcoming.
2. *Chandy, M., Schifano, E., and Yan, J. (2024+): On sample size needed for block bootstrap confidence intervals to have desired coverage rates. *American Journal of Undergraduate Research*. Forthcoming.
3. Choi, D., Bae, W., Yan, J., and Kang, S. (2024+): A general model-checking procedure for semiparametric accelerated failure time models. *Statistics and Computing*. Forthcoming.
4. *Yang, Z., Wang, H., and Yan, J. (2024+): Optimal subsampling for semi-parametric accelerated failure time models with massive survival data using the least-squares approach. *Statistics and Computing*. Forthcoming.
5. *Jin, J., Yan, J., Aseltine, R. H., and Chen, K. (2024+) Transfer learning with quantile regression. *Technometrics*. Forthcoming.
6. *Lautier, J., Pozdnyakov, V., and Yan, J. (2024+): On the maximum likelihood estimation of a discrete, finite support distribution under left-truncation and competing risks. *Statistics and Probability Letters*. Forthcoming.

7. *Lautier, J., Pozdnyakov, V., and Yan, J. (2023+): Estimating a discrete distribution subject to random left-truncation with an application to structured finance. *Econometrics and Statistics*. Forthcoming.
8. *Wang, W., Luo, C., Aseltine, R. H., Wang, F., Yan, J., and Chen, K. (2023+): Survival modeling of suicide risk with rare and uncertain diagnoses. *Statistics in Biosciences*. Forthcoming.
9. Bar, H. and Yan, J. (2023): Legendary career and colorful life: A conversation with Dr. Bob Riffenburgh. *Journal of Data Science*. 21(4): 818–837.
10. Chiou, S., Xu, G., Yan, J., and Huang, C.-Y. (2023): Regression modeling for recurrent events possibly with an informative terminal event using R package reReg. *Journal of Statistical Software*. 105(1): 1–34.
11. *Lau, A. Y. Z., Wang, T., Yan, J., and Zhang, X. (2023): Extreme value modeling with errors-in-variables in detection and attribution of changes in climate extremes. *Statistics and Computing*. 33(6): 125.
12. *Lautier, J. P., Pozdnyakov, P. and Yan, J. (2023): Pricing time-to-event contingent cash flows: A discrete-time survival analysis approach. *Insurance: Mathematics and Economics*. 110: 53–71.
13. *Li, Y., Chen, K., Yan, J., and Zhang, X. (2023): Regularized fingerprinting in detection and attribution of climate change with weight matrix optimizing the efficiency in scaling factor estimation. *Annals of Applied Statistics*. 17(1): 225–239.
14. *Ma, S., Wang, T., Yan, J., and Zhang, X. (2023): Optimal fingerprinting with estimating equations. *Journal of Climate*. 36(20): 7109–7122.
15. *Yuan, Y., Wang, T., Yan, J., and Zhang, P. (2023): Generating general preferential attachment networks with R package wdnet. *Journal of Data Science*. 21(3): 538–556.
16. *Wang, F., Wang, H., and Yan, J. (2023): Diagnostic tests for the necessity of weight in regression with survey data. *International Statistical Review*. 91(1): 55–71.
17. *Hu, C., Pozdnyakov, V., and Yan, J. (2022): On occupation time for on-off processes with multiple off-states. *Modern Stochastics: Theory and Applications*. 9(4): 413–430.
18. *Jiao, J., Tang, Z., Yue, M., Zhang, P., and Yan, J. (2022): Cyberattack-resilient load forecasting with adaptive robust regression. *International Journal of Forecasting*. 38(3): 910–919.
19. *Lau, A. Y. Z and Yan, J. (2022): Bias analysis of generalized estimating equations under measurement error and practical bias correction. *Stat*. 11(1): e418.
20. *Price, M. and Yan, J. (2022): The effects of the NBA COVID Bubble on the NBA playoffs: A case study for home-court advantage. *American Journal of Undergraduate Research*. 18(4): 3–15.
21. Sun, Q., Zwiers, F., Zhang, X., and Yan, J. (2022): Quantifying the human influence on the intensity of extreme 1- and 5-day precipitation amounts at global, continental, and regional scales. *Journal of Climate*. 35(1): 195–210.
22. Wang, T., Yan, J., *Yuan, Y., and Zhang, P. (2022): Generating directed networks with predetermined assortativity measures. *Statistics and Computing*. 32: 91.
23. *Xiao, S., Yan, J., and Zhang, P. (2022): Incorporating auxiliary information in betweenness measure for input-output networks. *Physica A: Statistical Mechanics and Its Applications*. 607: 128200.
24. *Yang, Z., Wang, H., and Yan, J. (2022): Optimal subsampling for parametric accelerated failure time models with massive survival data. *Statistics in Medicine*. 41(27): 5241–5431.
25. Zhang, P., Wang, T., and Yan, J. (2022): PageRank centrality and algorithms for weighted, directed networks. *Physica A: Statistical Mechanics and Its Applications*. 586: 126438.
26. Chang, S.-Y., Jin, J., Yan, J., Dong, X., Chaudhuri, B., Nagapudi, K., and Ma, A. W. K. (2021): Development of a pilot-scale HuskyJet binder jet 3D printer for additive manufacturing of pharmaceutical tablets. *International Journal of Pharmaceutics*. 605: 120791.
27. Feng Chang, C., Garcia, V., Tang, C., Vlahos, P., Wanik, D., Yan, J., Bash, J. O., and Astitha, M. (2021): Linking multi-media modeling with machine learning to assess and predict lake chlorophyll- α concentrations. *Journal of Great Lakes Research*. 47(6): 1656–1670.
28. *Hu, C., Elbroach, M., Meyer, T., Pozdnyakov, V., and Yan, J. (2021): Moving-resting process with measurement error in animal movement modeling. *Methods in Ecology and Evolution*. 12(11): 2221–2233.
29. *Jiao, J., Hu, G., and Yan, J. (2021): A Bayesian marked spatial point processes for basketball shot chart. *Journal of Quantitative Analysis in Sports*. 17(2): 77–90.

30. *Jiao, J., Hu, G., and Yan, J. (2021): Heterogeneity pursuit for spatial point pattern with application to tree locations: A Bayesian semiparametric recourse. *Environmetrics*. 32(7): e2694.
31. *Li, Y., Chen, K., Yan, J., and Zhang, X. (2021): Uncertainty in optimal fingerprinting is underestimated. *Environmental Research Letters*. 16(8): 084043.
32. *Wang, T., Xiao, S., Yan, J., and Zhang, P. (2021): Regional and sectoral structures of the Chinese economy: A network perspective from multi-regional input-output tables. *Physica A: Statistical Mechanics and Its Applications*. 581: 126196.
33. *Wang, W. and Yan, J. (2021): Shape-restricted regression splines with R packae splines2. *Journal of Data Science*. 19(3): 498–517.
34. *Wang, Z., Jiang, Y., Wan, H., Yan, J., and Zhang, X. (2021): Towards optimal fingerprinting in detection and attribution of changes in climate extremes. *Journal of the American Statistical Association*. 116(553): 1–13.
35. Wang, T., and Yan, J. (2021): Discussion of “On studying extreme values and systematic risks with nonlinear time series models and tail dependence measures”. *Statistical Theory and Related Fields*. 5(1): 38–40.
36. *Wu, J., Chen, M.-H., Schifano, E. D., and Yan, J. (2021): Online updating of survival analysis. *Journal of Computational and Graphical Statistics*. 30(4): 1209–1223.
37. *Xue, Y., Yan, J., and Schifano, E. (2021): Simultaneous monitoring for regression coefficients and baseline hazard profile in Cox modeling of time-to-event data. *Biostatistics*. 22(4): 756–771.
38. *Yuan, Y., Yan, J., Zhang, P. (2021): Assortativity measures for weighted and directed networks. *Journal of Complex Networks*. 9(2): cnab017.
39. *Doshi, R., Yan, J., and Aseltine, R. (2020): Age differences in racial/ethnic disparities in preventable hospitalizations for heart failure in Connecticut, 2009-2015: A population-based longitudinal study. *Public Health Reports* 135(1): 56–65.
40. *Hu, C., Pozdnyakov, V., and Yan, J. (2020): Density and distribution evaluation for convolution of independent gamma variables. *Computational Statistics*. 35(1): 327–342.
41. *Jiang, Y., He, X., Lee, M.-L. T., Rosner, B., and Yan, J. (2020): Rank-based tests for clustered data with R package clusrank. *Journal of Statistical Software*. 96(6): 1–26.
42. *Li, Y., Li, Y., Qin, Y., and Yan, J. (2020): Copula modeling for data with ties. *Statistics and Its Interfaces*. 13(1): 103–117.
43. Pozdnyakov, V., Elbroach, L. M., Hu, C., Meyer, T., and Yan, J. (2020): On estimation for Brownian motion governed by telegraph process with multiple off states. *Methodology and Computing in Applied Probability* 22: 1275–1291.
44. *Vaughan, G., Aseltine, R., Chen, K., and Yan, J. (2020): Efficient interaction selection for clustered data via stage-wise generalized estimating equations. *Statistics in Medicine*. 39(22): 2855–2868.
45. *Wang, W., Aseltine, R., Chen, K., and Yan, J. (2020): Integrative survival analysis with uncertain event times in application to a suicide risk study. *Annals of Applied Statistics*. 14(1): 51-73.
46. *Wang, C., Schifano, E., and Yan, J. (2020): Geographic ratings with spatial random effects in a two-part model. *Variance*. 13(1): 141–160.
47. Xu, G., Chiou, S., Yan, J., Marr, K., and Huang, C.-Y. (2020): Generalized scale-change models for recurrent event processes under informative censoring. *Statistica Sinica*. 30(4): 1773–1795.
48. *Xue, Y., Wang, H., Yan, J., and Schifano, E. (2020), An online updating approach for testing the proportional hazards assumption with streams survival data. *Biometrics*. 76(1): 171–182.
49. Aseltine, R., *Wang, W., Benthien, R., Katz, M., Wagner, C., Yan, J., and Lewis, C. (2019): Reductions in race and ethnic disparities in hospital readmissions following total joint arthroplasty from 2005–2015. *Journal of Bone and Joint Surgery*. 101(22) 2044–2050.
50. Caplan, D. J., *Li, Y., *Wang, W., Kang, S., Marchini, L., Cowen, H. J., and Yan, J. (2019): Dental restoration longevity among geriatric and special needs patients. *JDR Clinical & Translational Research*. 4(1): 41–48.
51. Chiou, S., Huang, C., Xu, G., and Yan, J. (2019): Semiparametric regression analysis of panel count data: A practical review. *International Statistical Review*. 87(1), 24–43.
52. Pozdnyakov, V., Elbroch, M., Labarga, J. A., Meyer, T., and Yan, J. (2019): Discretely observed Brownian motion governed by telegraph process: Estimation. *Methodology & Computing in Applied Probability* 21(3): 907–920.

53. *Bader, B., Yan, J., and Zhang, X. (2018): Automated threshold selection for extreme value analysis via goodness-of-fit tests with application to return level mapping. *Annals of Applied Statistics* 12(1): 310–329.
54. Chiou, S., Xu, G., Yan, J., and Huang, C. (2018): Semiparametric estimation of the accelerated mean model for panel count data with informative examination times. *Biometrics*, 74(3): 944–953.
55. *Wang, C., Chen, M.-H., Wu, J., Yan, J., Zhang, Y., and Schifano, E. (2018): Online updating method with new variables for big data streams. *Canadian Journal of Statistics* 46(1): 123–146.
56. *Bader, B., Yan, J., and Zhang, X. (2017): Automated selection of r for the r largest order statistics approach with adjustment for sequential testing. *Statistics and Computing* 27(6): 1435–1451.
57. *Vaughan, G., Aseltine, R. H., Chen, K., and Yan, J. (2017): Stagewise estimating equations with grouped variables. *Biometrics* 73(4): 1332–1342.
58. *Wang, Z., Jiang, Y., Wan, H., Yan, J., and Zhang, X. (2017): Detection and attribution of changes in extreme temperatures at regional level. *Journal of Climate* 30(17): 7035–7047.
59. Xu, G., Chiou, S., Huang, C.-Y., Wang, M.-C., and Yan, J. (2017): Joint scale-change models for recurrent events and failure time. *Journal of the American Statistical Association* 112: 794–805.
60. Olayivola, J. N., Adherson, D. R., Jepeal, N., Aseltine, R. H., Pickett, C., Yan, J., and Zlateva, I. (2016): Electronic consultations to improve the primary care-specialty care interface for cardiology in the medically underserved: A cluster-randomized controlled trial. *Annals of Family Medicine* 14(2): 133–140.
61. Schifano, E. D., Wu, J., Wang, C., Yan, J. and Chen, M.-H. (2016): Online updating of statistical inference in the big data setting. *Technometrics* 58(3): 393–403.
62. *Wang, C., Chen, M.-H., Schifano, E., Wu, J., and Yan, J. (2016): Statistical methods and computing for big data. *Statistics and Its Interfaces* 9(4): 399–414.
63. *Wang, W., Chen, M.-H., Chiou, S., Lai, H.-C., Wang, X., Yan, J., and Zhang, Z. (2016): Onset of persistent *Pseudomonas Aeruginosa* infection in children with cystic fibrosis with interval censored data. *BMC Medical Research Methodology* 16(122): 1–10.
64. Aseltine, R. H., Yan, J., Fleischman, S., Katz, M., and DeFrancesco M. (2015): Race and ethnic disparities in hospital readmissions following vaginal and cesarean delivery. *Obstetrics & Gynecology* 126(5): 1040–1047.
65. Aseltine, R. H., Yan, J., Gruss, C. B., Wagner, C., and Katz, M. (2015): Connecticut hospital readmissions related to chest pain and heart failure: Differences by race, ethnic, and payer. *Connecticut Medicine* 79(2): 69–76.
66. Chi, Z., Pozdnyakov, V., and Yan, J. (2015): On occupation time of Brownian motion. *Statistics and Probability Letters* 97: 83–87.
67. *Chiou, S., Kang, S., and Yan, J. (2015): Semiparametric accelerated failure time modeling for clustered failure times from stratified sampling. *Journal of the American Statistical Association* 110: 621–629.
68. *Chiou, S., Kang, S., and Yan, J. (2015): Rank-based estimating equations with general weight for accelerated failure time models: An induced smoothing approach. *Statistics in Medicine* 34(9): 1495–1510.
69. Kojadinovic, I., Shang, H., and Yan, J. (2015): A class of goodness-of-fit tests for spatial extremes models based on max-stable processes. *Statistics and Its Interfaces* 8(1): 45–62.
70. *Prates, M. O., Dey, D. K., Willig, M. R., and Yan, J. (2015): Transformed Gaussian Markov random fields and spatial modeling. *Spatial Statistics* 14(C), 382–399.
71. *Shang, H., Yan, J., and Zhang, X. (2015): A two-step approach to model precipitation extremes in California based on max-stable and marginal point processes. *The Annals of Applied Statistics* 9(1): 452–473.
72. *Chiou, S., Kang, S., Kim, J., and Yan, J. (2014): Marginal semiparametric multivariate accelerated failure time model with generalized estimating equations. *Lifetime Data Analysis* 20(4): 599–618.
73. *Chiou, S., Kang, S., and Yan, J. (2014): Fitting accelerated failure time models in routine survival analysis. *Journal of Statistical Software* 61(11): 1–23.
74. *Chiou, S., Kang, S., and Yan, J. (2014): Fast accelerated failure time modeling for case-cohort data. *Statistics and Computing* 24(4): 559–568.
75. Pozdnyakov, V., Meyer, T., Wang, Y., and Yan, J. (2014): On modeling animal movement using Brownian motion with measurement error. *Ecology* 95(2): 247–253.

76. *Wang, Z., Yan, J., and Zhang, X. (2014): Incorporating spatial dependence in regional frequency analysis. *Water Resources Research* 50(12): 9570–9585.
77. Yan, J., Chen, Y., Lawrence-Apfel, K., Ortega, I. M., Pozdnyakov, V., Williams, S., and Meyer, T. (2014): A moving-resting process with an embedded Brownian motion for animal movements. *Population Ecology* 56(2): 401–415.
78. Yan, J., Guo, C., Paarlberg, L. E. (2014): Are antipoverty nonprofit organizations located where they are needed? Spatial analysis of the Greater Hartford region. *American Statistician* 68(4): 243–252.
79. *Prates, M. O., Aseltine, R. H., Dey, D. K., and Yan, J. (2013): Assessing intervention efficacy on high risk drinkers using generalized linear mixed models with a new class of link functions. *Biometrical Journal* 55(6): 912–924.
80. *Wang, X., Chen, M.-H., and Yan, J. (2013): Bayesian dynamic regression model for interval censored data. *Lifetime Data Analysis* 19(3): 297–316.
81. *Wang, X., Ma, S., and Yan, J. (2013): Augmented estimating equations for semiparametric panel count regression with informative observation times and censoring time. *Statistica Sinica* 23(1): 359–381.
82. *Wang, X. and Yan, J. (2013): Practical notes on multivariate modeling based on elliptical copulas. *Journal of the French Statistical Society* 154(1): 102–115.
83. Yan, J., Aseltine, R., and Harel, O. (2013): Comparing regression coefficients between nested models for clustered data with generalized estimating equations. *Journal of Educational and Behavioral Statistics* 38(2): 172–189.
84. Cavallo, A., Rosenthal, B., Wang, X., and Yan, J. (2012): Treatment of the data collection threshold in operational risk: A case study using the lognormal distribution. *Journal of Operational Risk* 7(1): 3–38.
85. Chen, D. C. R., Kirshenbaum, D. S., Yan, J., Kirshenbaum, E., and Aseltine, R. H. (2012): Characterizing changes in student empathy throughout medical school. *Medical Teacher* 34(4): 305–311.
86. Havens, E. K., Martin, K. S., Yan, J., Dauser-Forrest, D., and Ferris, A. M. (2012): Federal nutrition program changes and healthy food availability. *American Journal of Preventive Medicine* 43(4): 419–422.
87. Kojadinovic, I. and Yan, J. (2012): A nonparametric test of exchangeability for bivariate extreme-value copulas. *Scandinavian Journal of Statistics* 39(3): 480–496.
88. Kojadinovic, I. and Yan, J. (2012): Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap. *Canadian Journal of Statistics* 40(3): 480–500.
89. Yan, J. and Huang, J. (2012): Model selection for time-varying coefficient Cox models. *Biometrics* 68(2): 419–428.
90. Yan, J., Liao, G.-Y., Gebremichael, M., Shedd, R., and Vallee, D. (2012): Characterizing the Uncertainty in River Stage Forecasts Conditional on Point Forecast Values. *Water Resources Research* 48: W12509.
91. Allignol, A., Latouche, A., Yan, J., and Fine, J. P. (2011): A regression model for the conditional probability of a competing event: Application to monoclonal gammopathy of unknown significance. *Journal of Royal Statistical Society, Series C: Applied Statistics* 60(1): 135–142.
92. Gebremichael, M., Liao, G.-Y., and Yan, J. (2011): Non-parametric error model for high resolution satellite rainfall products. *Water Resources Research* 47: W07504–W07512.
93. Genest C., Kojadinovic I., Nešlehová J., and Yan J. (2011): A goodness-of-fit test for bivariate extreme value copulas. *Bernoulli* 17(1): 253–275.
94. Guan, Y., Yan, J., and Sinha, R. (2011): Variance estimation for statistics computed from single recurrent event processes. *Biometrics* 17(3): 711–718.
95. Harel, O., Mukhopadhyay, N., and Yan, J. (2011): On a sequential probability ratio test subject to incomplete data. *Sequential Analysis* 30: 441–456.
96. Kojadinovic, I., Segers, J., and Yan, J. (2011): Large-sample tests of extreme-value dependence for multivariate copulas. *Canadian Journal of Statistics* 39(4): 703–720.
97. Kojadinovic, I. and Yan, J. (2011): Tests of serial independence for multivariate time series based on a Möbius decomposition of the independence empirical copula process. *Annals of the Institute of Statistical Mathematics* 63(2): 347–373.
98. Kojadinovic, I. and Yan, J. (2011): A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing* 21(1): 17–30.
99. Kojadinovic, I., Yan, J., and Holmes, M. (2011): Fast large-sample goodness-of-fit for copulas. *Statistica Sinica* 21(2): 841–871.

100. *Prates, M. O., Dey, D. K., Willig, M. R. and Yan, J. (2011): Intervention analysis of the hurricane impact on snail abundance in a tropical forest with spatio-temporal data. *Journal of Agricultural, Biological, and Ecological Statistics* 16(1): 142–156.
101. *Shang, H., Yan, J., Gebremichael, M., and Ayalew, S. M. (2011): Trend analysis of extreme precipitation in the northwestern highlands of Ethiopia with a case study of Debre Markos. *Hydrology and Earth System Sciences* 15(3): 1937–1944.
102. *Shang, H., Yan, J., and Zhang, X. (2011): ENSO influence on winter maximum daily precipitation in California in a spatial model. *Water Resources Research* 47: W11507–W11515.
103. *Wang, X. and Yan, J. (2011): Fitting semiparametric regressions for panel count survival data with an R package *spef*. *Computer Methods and Programs in Biomedicine* 104(2): 278–285.
104. Kojadinovic, I. and Yan, J. (2010): Nonparametric rank-based tests of bivariate extreme-value dependence. *Journal of Multivariate Analysis* 101(9): 2234–2249.
105. Kojadinovic, I. and Yan, J. (2010): Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics* 47(1): 52–63.
106. Kojadinovic, I. and Yan, J. (2010): Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software* 34(9): 1–20.
107. *Wang, X., Gebremichael, M., and Yan, J. (2010): Weighted likelihood copula modeling of extreme rainfall events in Connecticut. *Journal of Hydrology* 390(1–2): 108–115.
108. Yan, J. and the Academic ED SBIRT Research Collaborative. (2010): The impact of screening, brief intervention and referral for treatment in emergency department patients' alcohol use: A 3-, 6- and 12-month follow-up. *Alcohol & Alcoholism* 45(6): 514–519.
109. Yan, J., Cheng, Y., Fine, J. P., and Lai, H.-C. (2010): Uncovering symptom progression history from disease registry data with application to young cystic fibrosis patients. *Biometrics* 66(2): 594–602.
110. Cowles, M. K., Yan, J., and Smith, B. J. (2009): Reparameterized and marginalized posterior and predictive sampling for complex Bayesian geostatistical models. *Journal of Computational and Graphical Statistics* 18(2): 262–282.
111. Yan, J. and Gebremichael, M. (2009): Estimating actual rainfall from satellite rainfall products. *Atmospheric Research* 92(4): 481–488.
112. Yan, J. and Huang, J. (2009): Partly functional temporal process regression with semiparametric profile estimating functions. *Biometrics* 65(2): 431–440.
113. Smith, B. J., Yan, J., and Cowles, M. K. (2008): Unified geostatistical modeling for data fusion and spatial heteroskedasticity with R package *ramps*. *Journal of Statistical Software* 25(10): 1–21.
114. Yan, J. and Fine, J. P. (2008): Analysis of episodic data with application to recurrent pulmonary exacerbations in cystic fibrosis patients. *Journal of the American Statistical Association* 103: 498–510.
115. Stramer, O. and Yan, J. (2007): Asymptotics of an efficient Monte Carlo estimation for the transition density of diffusion processes. *Methodology & Computing in Applied Probability* 9(4): 483–496.
116. Stramer, O. and Yan, J. (2007): On simulated likelihood of discretely observed diffusion processes and comparison to closed-form approximation. *Journal of Computational and Graphical Statistics* 16(3): 672–691.
117. Yan, J. (2007): Enjoy the joy of copulas. *Journal of Statistical Software* 21(4): 1–21.
118. Yan, J. (2007): Spatial stochastic volatility for lattice Data. *Journal of Agricultural, Biological, and Environmental Statistics* 12(1): 25–40.
119. Yan, J., Cowles, M. K., Wang, S., and Armstrong, M. P. (2007): Parallelizing MCMC for Bayesian spatiotemporal geostatistical models. *Statistics and Computing* 17(4): 323–335.
120. Yan, J. and Tamboli, C. P. (2007): Testing concordance of clinical characteristics in familial studies with application to inflammatory bowel diseases. *Biometrical Journal* 49(6): 840–853.
121. Halekoh, U., Højsgaard, S., and Yan, J. (2006): The R package *geepack* for generalized estimating equations. *Journal of Statistical Software* 15(2): 1–11.
122. Yan, J. and Fine, J. P. (2005): Functional association models for multivariate survival processes. *Journal of the American Statistical Association* 100(469): 184–196.
123. Fine, J. P., Yan, J., and Kosorok, M. R. (2004): Temporal process regression. *Biometrika* 91(3): 683–703.

124. Yan, J. and Fine, J. P. (2004): Estimating equations for association structures (Pkg: P859-880). *Statistics in Medicine* 23(6): 859–874.
125. Yan, J. and Fine, J. P. (2004): Reply to comment on “Estimating equations for association structures” (Pkg: 859–880). *Statistics in Medicine* 23(6): 879–880.

Papers Submitted to Refereed Journals

1. *Hughes, S., Matthews, G., and Yan, J. (2024): Statistical Evaluation of Outdoor Field Hockey Penalty Corners.
2. *Yang, Z., Wang, H., and Yan, J. (2023): Optimal subsampling for semi-parametric accelerated failure time models with massive survival data using a rank-based approach.
3. *Fiore, O., Schifano, E., and Yan, J. (2023): On Devon Allen’s disqualification at the 2022 World Track and Field Championships.
4. *Zeimbekakis, A., Schifano, E., and Yan, J. (2023): On misuses of the Kolmogorov–Smirnov test for one-sample goodness-of-fit.
5. *Wang, T., *Xiao, S., and Yan, J. (2023): Comparison of sectoral structures between China and Japan: A network perspective.
6. Chiou, S., Aseltine, R., Schilling, E., Lutz, K., and Yan, J. (2022): A bivariate two-part model for censored durations of depression and relational stressor in young adults.
7. *Lautier, J., Pozdnyakov, V., and Yan, J. (2022): On the convergence of credit risk in current consumer automobile loans.
8. *Xie, Y., Fu, H., Huang, Y., Pozdnyakov, V., and Yan, J. (2022): Recurrent events modeling based on a reflected Brownian motion with application to hypoglycemia.
9. *Godoy, L., Prates, M., and Yan, J. (2022): Model-based Voronoi linkage between point-referenced data and areal data in spatial analysis with application to Brazilian election 2018.
10. *Jiao, J., Song, W., Xue, Y., and Yan, J. (2022): Heteroscedastic growth curve modeling with shape-restricted splines.
11. *Ma, S., Yan, J., and Zhang, X. (2020): Extreme value modeling with generalized Pareto distributions for rounded data.

Software

1. *Bader, B. and Yan, J.: R package `eva` on CRAN, extreme value analysis.
2. *Chiou, S., Kang, S., and Yan, J.: R package `aftgee` on CRAN, multivariate accelerated failure time modeling with generalized estimating equations.
3. *Chiou, S., Wang, X. and Yan, J.: R package `spef` on CRAN, semiparametric estimating functions.
4. Hofert, M., Kojadinovic, I., Mächler, M. and Yan, J.: R package `copula` on CRAN, multivariate dependence with copula.
5. *Hu, C., Pozdnyakov, V., and Yan, J.: R package `coga` on CRAN, convolution of gamma variables.
6. *Hu, C., Yan, J., and Pozdnyakov, V.: R package `smam` on CRAN, statistical modeling of animal movement.
7. *Jiang, Y., Lee, M.-L. T., and Yan, J.: R package `clusrank` on CRAN, rank-based tests for clustered data.
8. Kojadinovic, I. and Yan, J.: R package `fgof` on CRAN, fast goodness-of-fit test.
9. *Li, Y., Chen, K., and Yan, J.: R package `tls` on CRAN, total least squares.
10. *Li, Y., Chen, K., and Yan, J.: R package `dacc` on CRAN, detection and attribution of climate change.
11. *Li, Y., Wang, W., and Yan, J.: R package `touch` on CRAN, tools of utilization and cost in healthcare.
12. *Prates, M. O., Wang, W., and Yan, J.: R package `rbugs` on CRAN, fusing R with OpenBugs.
13. Smith, B. P., Yan, J., and Cowles, M. K.: R package `ramps` on CRAN, reparametrized and marginalized posterior sampling.
14. *Vaughan, G., Chen, K., and Yan, J.: R package `sgee` on CRAN, stagewise generalized estimating equations.
15. *Wang, W., Chen, K., and Yan, J.: R package `intsurv` on CRAN, integrative survival analysis.
16. *Wang, W., Fu, H., and Yan, J.: R package `reda` on CRAN, recurrent event data analysis.

17. *Wang, W. and Yan, J.: R package `splines2` on CRAN, regression spline functions and classes too.
18. *Wang, X., Chen, M.-H., and Yan, J.: R package `dynsurv` on CRAN, dynamic survival modeling.
19. Yan, J.: R package `tpr` on CRAN, temporal process regression.
20. Yan, J.: R package `som` on CRAN, self-organizing map with application to gene clustering.
21. Yan, J., Højsgaard, S., and Halekoh, U.: R package `geepack` on CRAN, generalized estimating equation package.
22. Yan, J.: R package `KMsurv` on CRAN, datasets and functions for Klein and Moeschberger (1997), “Survival Analysis, Techniques for Censored and Truncated Data”, Springer.
23. *Yuan, Y., Wang, T., Yan, J., and Zhang, P.: R package `wdnet` on CRAN, weighted directed networks.

Non-Refereed Publications

1. Yan, J. (2020): A reformed Journal of Data Science for the era of data science. *Journal of Data Science*, 18(3): 405–406.
2. Follman, D., Song, P. X.-K., Wang, H., and Yan, J. (2020): Data science in action in response to the outbreak of COVID-19. *Journal of Data Science* 18(3): 407–408.
3. Yan, J. (2004): Fusing R and BUGS through Wine. *R News* 4(2): 19–21.
4. Yan, J. and Rossini, A. (2003): Building Microsoft Windows versions of R and R packages under Intel Linux. *R News* 3(1): 15–17.
5. Yan, J. (2002): geepack: Yet another package for generalized estimating equations. *R News* 2(3): 12–14.

Book Reviews

1. Yan, J. (2006): Gaussian Markov random fields: Theory and applications. Harvard Rue and Leonhard Held. *Journal of the American Statistical Association* 101(473): 388–389.
2. Yan, J. (2005): Analysis of multivariate survival data. Phillip Hougaard. *Journal of the American Statistical Association* 100(469): 355–356.
3. Yan, J. (2004): Bayesian survival analysis. Joseph G. Ibrahim, Ming-Hui Chen, and Debajyoti Sinha. *Journal of the American Statistical Association* 99(468): 1202–1203.
4. Yan, J. (2004): Survival analysis: Techniques for censored and truncated data (2nd ed.). John P. Klein and Melvin L. Moeschberger. *Journal of the American Statistical Association* 99(467): 900–901.

Invited Talks

- Optimal Fingerprinting in Climate Change Detection and Attribution with Estimating Equations, 11/20/2023, Department of Statistics, Oregon State University (virtual).
- Recurrent Events Modeling Based on a Reflected Brownian Motion with Application to Hypoglycemia, 03/30/2023, Department of Statistics, Kansas State University (virtual).
- Optimal Fingerprinting with Estimating-Equations, 08/22/2023, Learning the Earth with Artificial Intelligence and Physics Center, Columbia University.
- Introductory Overview Lecture: Sports Analytics Beyond Performance Evaluation, 08/08/2022, Joint Statistical Meetings.
- Optimal Fingerprinting with Bias-Corrected Estimating Equations, 06/11/2022, Special Conference Celebrating the 80th Anniversary of Renmin, Institute of Statistics and Big Data, Renmin University of China (virtual).
- Recurrent Events Modeling Based on a Reflected Brownian Motion with Application to Hypoglycemia, 05/18/2022, Department of Applied Mathematics, The Hong Kong Polytechnic University (virtual).
- Recurrent Events Modeling Based on a Reflected Brownian Motion with Application to Hypoglycemia, 03/29/2022, ENAR Spring Meeting (virtual).
- Optimal Fingerprinting with Bias-Corrected Estimating Equations, 01/31/2022, International Detection and Attribution Group (virtual).
- Brownian Motion Governed by Telegraph Process in Modeling High-Frequency Financial Series, 08/27/2021, Department of Statistics, Universidade Federal de Minas Gerais (virtual).
- Correct Working Correlation for Generalized Estimating Equations May Lead to More Bias Under Measurement Error than Working Independence, 06/09/2021, Department of Applied Statistics, Yonsei University (virtual).

- Moving-Resting Process with Measurement Error in Animal Movement Modeling, 10/26/2020, Center for Statistical Science, Tsinghua University (virtual).
- An Applied Statistician's Adventure: Climate Change, Animal Movement, Sports Analytics, and Beyond, 05/09/2020, Clubear Lecture (virtual).
- An Online Updating Approach for Testing the Proportional Hazards Assumption with Streams of Survival Data, 11/29/2019, The 6th Workshop on Survival Analysis and Applications, University of San Paulo, Brazil.
- Integrative Survival Analysis with Uncertain Event Times in Application to a Suicide Risk Study, 07/05/2019, School of Statistics, Shanxi University of Finance and Economics.
- Acrobatic Regression in Detection and Attribution of Climate Change, 07/03/2019, Guanghua School of Management, Peking University.
- An Online Updating Approach for Testing the Proportional Hazards Assumption with Streams of Survival Data, 06/30/2019, School of Mathematics, Jilin University
- Fingerprinting Changes in Climate Extremes with Joint Modeling of Observations and Climate Model Simulation, 07/30/2018, Vancouver, BC, Canada, Joint Statistical Meetings.
- Generalized Scale-Change Models for Recurrent Event Processes under Informative Censoring, 07/03/2018, 2018 ICSA China Conference with the Focus on Data Science, Qingdao, China.
- Growth Curve Analysis with Shape-restricted Splines, 07/01/2018, International Statistics Forum, Renmin University of China, Beijing, China.
- Generalized Scale-Change Models for Recurrent Event Processes under Informative Censoring, 06/30/2018, School of Mathematics, Jilin University.
- Optimal Fingerprinting in Detection and Attribution of Changes in Climate Extremes, 06/26/2018, Center for Statistical Science, Peking University.
- Statistical Methods for Big Stream Data, 06/19/2018, Shanxi University of Finance and Economics
- Things about Being a Professor, 12/20/2017, School of Statistics, Renmin University of China.
- Online Updating Method with New Variables for Big Data Streams, 12/17/2017, School of Statistics, Shanxi University of Finance and Economics.
- Balancing the Bias-Variance Tradeoff in Extreme Value Analysis, 08/31/2017, Center for Mathematical Research, University of Montréal.
- Stagewise Generalized Estimating Equations with Grouped Variables, 03/02/2017, Department of Mathematics and Statistics, Boston University.
- Online Updating Method with New Variables for Big Data Streams, 10/12/2016, Department of Statistics and Biostatistics, Rutgers University.
- Semiparametric Accelerated Failure Time Modeling for Clustered Failure Times From Stratified Sampling, 07/02/2016, IBS-China 4th International Biostatistics Symposium, Shanghai, China.
- Spatial estimating equations for detection and attribution of changes in climate extremes, 05/29/2016, China R Conference, Beijing, China.
- Spatial estimating equations for detection and attribution of changes in climate extremes, 02/02/2016, International Detection and Attribution Group Meeting, Boulder, CO.
- Spatial estimating equations with application to changes in climate extremes, 10/26/2015, Department of Mathematical Sciences, Worcester Polytechnic Institute.
- Optimal fingerprinting in detection and attribution of changes in climate extremes with combined score equations, 06/20/2015, Alumni Symposium, School of Statistics, Renmin University of China, Beijing, China.
- Onset time of chronic pseudomonas aeruginosa infection of cystic fibrosis patients with interval censored data, 06/16/2015, Applied Statistics Symposium, International Chinese Statistical Association, Fort Collins, CO.
- A bivariate two-part model to assess the effect of coping strategy on stressor and depression, 06/06/2015, Frontiers in Applied and Computational Mathematics 2015, New Jersey Institute of Technology (NJIT) in Newark, New Jersey.
- Incorporating spatial dependence in regional frequency analysis, 05/25/2014, International Statistics Forum, Renmin University of China, Beijing China.

- A partial review of software for big data statistics, 02/12/2014, Statistical and Computational Theory and Methodology for Big Data Analysis, Banff International Research Station, Banff, Alberta, Canada.
- Statistics methods and computing for semiparametric accelerated failure time models with induced smoothing, 05/13/2013, Department of Biostatistics, Brown University.
- Transformed Gaussian Markov random fields and Spatial Modeling, 11/23/2012, Universidade Federal de Minas Gerais, Brazil.
- Transformed Gaussian Markov random fields and Spatial Modeling, 10/05/2012, Department of Biostatistics, University of Massachusetts—Amherst.
- Fast accelerated failure time model for case-cohort data, 06/24/2012, Boston, MA, International Chinese Statistical Association Applied Statistics Symposium.
- Model selection for Cox models with time-varying coefficients, 04/21/2012, Boston University, New England Statistics Symposium.
- Max-Stable processes for spatial extremes modeling: A review and some ongoing research, 03/01/2012, Climate Research Division, Environmental Canada.
- Multivariate accelerated failure time models with generalized estimating equations, 10/04/2011, Biostatistics Research Branch, National Institute of Allergy and Infectious Disease.
- Augmented estimating equations for semiparametric panel count regression with informative censoring, 03/23/2011, ENAR Spring Meeting.
- Nonparametric rank-based tests of bivariate extreme-value dependence, 10/27/2010, UMass-UConn Joint Statistics Symposium.
- Augmented estimating equations for semiparametric panel count regression with informative censoring, 06/30/2010, Yunnan University, International Conference on Statistical Analysis of Complex Data.
- Nonparametric rank-based tests of bivariate extreme-value dependence, 04/17/2010, Harvard University, 2010 New England Statistics Symposium.
- Combining data for efficient prediction of the spatial distribution of Iowa residential radon levels, 08/02/2009, Washington, DC, Joint Statistical Meetings.
- Fast large sample goodness-of-fit test for copulas, 10/02/2008, Département de mathématiques et de statistique, Université Laval.
- Tests of serial independence for multivariate time series based on a Möbius decomposition of the independence empirical copula process, 06/21/2008, Renmin University of China, International Statistics Forum 2008.
- Partly functional temporal process regression with semiparametric profile estimating functions, 01/29/2008, Division of Biostatistics, Yale University.
- Spatial stochastic volatility, 07/12/2006, Beijing, China, Far Eastern Meeting of the Econometrics Society (FEMES) 2006.
- Partly functional temporal process regression, 06/28/2006, Hong Kong, China, INFORMS International Conference 2006.
- Spatial stochastic volatility, 04/21/2006, Department of Economics, University of Illinois – Urbana-Champaign.
- Temporal process regression, 08/10/2004, Toronto, Canada, Joint Statistical Meeting.

Invited Workshops/Shortcourses

- June 2024, Climate Change Detection and Attribution with Estimating Equations (2-hour short course), 15th International Meeting on Statistical Climatology, Toulouse, France.
- June 2023, Applied Event Time Data Analysis with R (1-day course with Steven Chiou). 2023 ICASA Applied Statistics Symposium, Ann Arbor, Michigan.
- May 2022, Applied Event Time Data Analysis with R (1-day course with Steven Chiou). 2022 New England Statistics Symposium, Storrs, CT.
- July 2016. Advanced Statistical Computing (20-hour course). Shanghai University of Finance and Economics, Shanghai, China.
- June 2015. Advanced Statistical Computing (20-hour workshop). Shanghai University of Finance and Economics, Shanghai, China.

- April 2015. Modern Multivariate Statistical Learning: Methods and Applications (1-day short course with Kun Chen). The 29th NESS at University of Connecticut, Storrs, CT.
- May 2014. Advanced Statistical Computing (32-hour course). Renmin University of China, Beijing China.
- April 2013. Statistical Analysis of Spatial Data and Visualization with Google Map (1-day short course with Marcos Prates). The 27th NESS at University of Connecticut, Storrs, CT.
- November 2012. Introduction to the Theory and Practice of Copulas (1-week short course). Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.
- April 2011. Introduction to the Theory and Practice of Copulas (1-day short course with Ivan Kojadinovic). The 25th NESS at University of Connecticut. Storrs, CT.
- May 2008. Introduction to the Theory and Practice of Copulas (32-hour course). Renmin University of China, Beijing, China.