

Anna Liu

Department of Mathematics and Statistics
University of Massachusetts, Amherst
Amherst, MA 01003

Phone: (413) 545-6021

Fax: (413) 545-1801

Email: anna@math.umass.edu

Web: <http://www.math.umass.edu/~anna>

DEGREES

Ph.D., Statistics, University of California at Santa Barbara, 2004

M.S., Statistics, Institute of Systems Science, Chinese Academy of Science, 2000

B.S., Statistics, Peking University, China, 1997

RESEARCH INTERESTS

Design of experiments

Nonparametric and semiparametric regression and density estimation

Biostatistical modeling

Mixed effects models and longitudinal data analysis

Batch correction for Single-cell RNA sequencing data

Machine learning and neural networks

Interactive data exploration and active learning

PROFESSIONAL EXPERIENCE

Associate head, Department of Mathematics and Statistics, University of Massachusetts, Amherst, 2020 – present

Professor, Department of Mathematics and Statistics, University of Massachusetts, Amherst, 2019 – present

Associate professor, Department of Mathematics and Statistics, University of Massachusetts, Amherst, 2011 – 2019

Visiting assistant professor, Department of Biostatistics and Computational Biology, University of Rochester, 2011 – 2012

Assistant professor, Department of Mathematics and Statistics, University of Massachusetts, Amherst, 2004 – 2011

GRANTS

Characterization of Microbiota-derived Polymethoxyflavone Metabolites and their Anti-inflammatory Actions in the Colon.

PI: Hang Xiao, Food Science, **Co-PI**: Anna Liu

Funding: NIH 1R01AT010229-O1, 2018-2022, \$1,922,500.00

Overcoming Learned Non-Use In Chronic Aphasia: Behavioral, fMRI and QOL Outcomes.

PI: Jacquie Kurland, Communication Disorders, **Statistics Consultant:** Anna Liu
Funding: NIH NIDCD, 2016-2018, \$1,969,399.00

Collaborative research: Smoothing spline semiparametric density models

PI: Anna Liu

Collaborated with Yuedong Wang at University of California, Santa Barbara
Funding: NSF DMS, 2015-2019, \$129,265.00

Vitamins B1 and K Degradation in Spaceflight Foods: Establishment of Prediction Models and Prevention Strategies

PI: Hang Xiao, Food Science, **Co-PI:** Anna Liu

Funding: NASA NNX14AP32G, 2014-2018, \$982,687.00

III-COR-small: High-Performance Complex Processing of Continuous Uncertain Data

PI: Yanlei Diao, Computer Science, **Co-PI:** Anna Liu

Funding: NSF Information Integration and Informatics, 2012-2017, \$511,961.00

Collaborative Research: Modeling DOC Dynamics from Landscapes to Coasts: Hydrological Connectivity and Estuary Processes

PI: Qian Yu, Geosciences, **Co-PI:** Anna Liu

Collaborated with Yong Tian and Robert Chen at UMass-Boston

Funding: NSF Collaboration in Mathematical Geosciences, 2010-2013, \$517,987 (\$329,346 to Amherst Campus)

The UMass Foundation-Mass Mutual Group

PI: Susan Worgaftik, Development, Michael Sullivan, Math and Stat, **Co-PI:** Anna Liu

Funding: Mass Mutual Group, 2010-2011, \$22,436

Generalized Spline Mixed-Effects Models: Implementation, Testing, and Applications

PI: Anna Liu

Funding: NSA, 2009-2011, \$29,997

III-COR-Small: Capturing Data Uncertainty in High-Volume Stream Processing

PI: Yanlei Diao, Computer Science, **Co-PI:** Anna Liu

Funding: NSF Information Integration and Informatics, 2008-2011, \$441,594

The UMass Foundation-Mass Mutual Group

PI: James Mallet, Development, Michael Sullivan, Math and Stat, **Co-PI:** Anna Liu

Funding: Mass Mutual Group, 2009-2010, \$21,374

Genome-Based Models to Optimize In Situ Bioremediation of Uranium and Harvesting Electrical Energy from Waste Organic Matter

PI: Derek Lovley, Department of Microbiology, UMass-Amherst, with nine Co-PIs

Statistical Consultant: Anna Liu

Funding: U.S. Department of Energy, 2005-2010, \$21.8 million

Lack-of-fit Tests for Mixed Effects Models

PI: Anna Liu

Funding: UMASS Faculty Research Grant, 2005-2006, \$12,559

MANUSCRIPTS and PUBLICATIONS in REFEREED JOURNALS

1. J. J. Yu, J. Shi, **A. Liu**, and Y. Wang (2021). Smoothing Spline Semiparametric Density Models. *Journal of the American Statistical Association*. DOI: 10.1080/01621459.2020.1769636.
2. C. Chiu, **A. Liu**, and Y. Wang (2020). Smoothing spline mixed-effects density models for clustered data. *Statistica Sinica*. 30, 397-416. ss.202017.0211.
3. J. Shi, **A. Liu**, and Y. Wang (2019). Spline density estimation and inference with model-based penalties. *Journal of Nonparametric Statistics*. 31, 596-611.
4. E. Huang, L. Peng, L. Di Palma, A. Abdelkafi, **A. Liu**, and Y. Diao (2018). Optimization for Active Learning-based Interactive Database Exploration. *The Proceedings of the VLDB (Very Large Data Bases) Endowment (PVLDB)*. 12(1):71-84. DOI: <https://doi.org/10.14778/3275536.3275542>.
5. J. Kurland, **A. Liu**, and P. Stokes (2018). Effects of a Tablet-based Home Practice Program with Telepractice on Treatment Outcomes in Chronic Aphasia. *Journal of Speech, Language, and Hearing Research*. 61, 1140-1156.
6. J. Kurland, **A. Liu**, and P. Stokes (2018). Practice Effects in Healthy Older Adults: Implications for Treatment-Induced Neuroplasticity in Aphasia. *Neuropsychologia*. 109, 116-125.
7. W. Liu, Y. Diao, and **A. Liu** (2018). An analysis of query-agnostic sampling for interactive data exploration. *Communication in Statistics - Theory and Methods*. 47(16), 3820-3837.
8. H. Zhang, J. Zhang, **A. Liu**, H. Xiao, and L. He (2016). Label-free Imaging and Characterization of Cancer Cell Responses to Polymethoxyflavones Using Raman Microscopy. *J. Agric. Food Chem.* 64(51), 9708-9713.
9. X. Zhang, H. Liang, **A. Liu**, D. Ruppert, and G. Zou, (2016) Selection Strategy for Covariance Structure of Random Effects in Linear Mixed-effects Models. *Scandinavian Journal of Statistics*. 43, 275-291.
10. J. Gagnon, H. Liang, and **A. Liu** (2016). Spherical radial approximation for nested mixed effects models. *Statistics and Computing*. 26, 119-130.
11. T. Tran, Y. Diao, C. Sutton, and **A. Liu** (2013). Supporting User-Defined Functions on Uncertain Data. *PVLDB*. 6(6), 469-480.
12. M. Aklujkar, M.V. Coppi, C. Leang, B.C. Kim, M.A. Chavan, L.A. Perpetua, L. Giloteaux, **A. Liu**, and D.E. Holmes (2013). Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by *Geobacter sulfurreducens* and *Geobacter uraniireducens*. *Microbiology*. 159, 515-535.
13. T. Tran, L. Peng, Y. Diao, A. McGregor, and **A. Liu** (2012). CLARO: Modeling and Processing Uncertain Data Streams. *The VLDB Journal*. 21(5), 651-676.

14. **A. Liu**, R. Kronmal, X. Zhou, and S. Ma (2011). Determination of proportionality in two-part models and analysis of Multi-Ethnic Study of Atherosclerosis (MESA). *Statistics and Its Interface*. 4, 457-488.
15. L. Peng, Y. Diao, and **A. Liu** (2011). Optimizing Probabilistic Query Processing on Continuous Uncertain Data. *PVLDB*, 4(11), 1169-1180.
16. S.M. Strycharz, R.H. Glavev, M.V. Coppi, S.M. Gannon, L.A. Perpetua, **A. Liu**, K.P. Nevin, and D.R. Lovley (2011). Gene Expression and Deletion Analysis of Mechanisms for Electron Transfer from Electrodes to *Geobacter Sulfurreducens*. *Bioelectrochemistry*. 80(2), 142-150.
17. S. Ma, **A. Liu**, J. Carr, W. Post, and R. Kronmal (2010). Statistical Modeling of Agatston Score in Multi-Ethnic Study of Atherosclerosis (MESA). *PLoS ONE*. 5(8): e12036.
18. **A. Liu**, L. Qin, and J. Staudenmayer (2010). M-Type Smoothing Spline ANOVA for Correlated Data. *Journal of Multivariate Analysis*. 101, 2282-2296.
19. Q. Yu, Y.Q. Tian, R. Chen, **A. Liu**, B. Gardner, and W. Zhu (2010). Functional Linear Analysis for Estimating Riverine CDOM in Coastal Environment Using In Situ Hyperspectral Data. *Photogrammetric Engineering and Remote Sensing*. 76(10), 1147-1158.
20. T. Tran, A. McGregor, Y. Diao, L. Peng, and **A. Liu** (2010). Conditioning and Aggregating Uncertain Data Streams: Going Beyond Expectations. *PVLDB*, 3, 1302-1313.
21. D.E. Holmes, R.A. O'Neil, M.A. Chavan, L.A. N'Guessan, H.A. Vrionis, L.A. Perpetua, M.J. Larrahondo, R. DiDonato, **A. Liu**, and D.R. Lovley (2009). Transcriptome of *Geobacter Uranireducens* Growing in Uranium-Contaminated Subsurface Sediments. *ISME J*. 3(2):216-230.
22. K.P. Nevin, B.C. Kim, R.H. Glaven, J.P. Johnson, T.L. Woodard, B.A. Methe, R.J. DiDonato, S.F. Covalla, A.E. Franks, **A. Liu**, and D.R. Lovley (2009). Anode Biofilm Transcriptomics Reveals Outer Surface Components Essential for High Density Current Production in *Geobacter Sulfurreducens* Fuel Cells. *PLoS ONE*. 4(5):e5628.
23. P.J. Mouser, D.E. Holmes, L.A. Perpetua, R. DiDonato, B. Postier, **A. Liu**, and D.R. Lovley (2009). Quantifying Expression of *Geobacter* Spp. Oxidative Stress Genes in Pure Culture and during In Situ Uranium Bioremediation. *ISME J*. 3(4):454-65.
24. S.A. Haveman, R.J. DiDonato, L. Villanueva, E.S. Shelobolina, B.L. Postier, B. Xu, **A. Liu**, and D.R. Lovley (2008). Genome-Wide Gene Expression Patterns and Growth Requirements Suggest that *Pelobacter Carbinolicus* Reduces Fe(III) Indirectly via Sulfide Production. *Appl Environ Microbiol*. 74(14):4277-4284.
25. T. Tong, **A. Liu**, and Y. Wang (2008). Relative Errors of Difference-Based Variance Estimators in Nonparametric Regression. *Communications in Statistics - Theory and Methods*. 18, 2890-2902.
26. B. Postier, R. DiDonato, K.P. Nevin, **A. Liu**, B. Frank, D. R. Lovley, and B.A. Methe (2008). Benefits of In-Situ Synthesized Microarrays for Analysis of Gene

- Expression in Understudied Microorganisms. *J Microbiol Methods*, 74(1):26-32.
27. M.V. Coppi, R.A. O'Neil, C. Leang, F. Kaufmann, B.A. Methe, K.P. Nevin, T.L. Woodard, **A. Liu**, and D.R. Lovley (2007). Involvement of *Geobacter Sulfurreducens* SfrAB in Acetate Oxidation Rather than Intracellular, Respiration-Linked Fe(III) Citrate Reduction. *Microbiology*. 153, 3572-3585.
 28. E.M. Conlon, J.J. Song, and **A. Liu** (2007). Bayesian Meta-Analysis Models for Microarray Data: a Comparative Study. *BMC Bioinformatics*. 8, 80.
 29. **A. Liu**, T. Tong, and Y. Wang (2007). Smoothing Spline Estimation of Variance Functions. *Journal of Computational and Graphical Statistics*. 16, 312-329.
 30. **A. Liu** and Y. Wang (2006). Modeling of Hormone Secretion-Generating Mechanisms with Splines: A Pseudo-Likelihood Approach. *Biometrics*. 63, 201-208.
 31. D.E. Holmes, S.K. Chaudhuri, K.P. Nevin, T. Mehta, B.A. Methe, **A. Liu**, J.E. Ward, T.L. Woodard, J. Webster, and D.R. Lovley (2006). Microarray and Genetic Analysis of Electron Transfer Electrodes in *Geobacter Sulfurreducens*. *Environmental Microbiology*. 8, 1805-1815.
 32. Y. Yang, **A. Liu**, and Y. Wang (2006). Detecting Pulsatile Hormone Secretions Using Nonlinear Mixed Effects Partial Spline Models. *Biometrics*. 62, 230-238.
 33. **A. Liu**, W. Meiring, and Y. Wang (2005). Testing Generalized Linear Models Using Smoothing Spline Methods. *Statistica Sinica*. 15, 235-256.
 34. **A. Liu** and Y. Wang (2004). Hypothesis Testing in Smoothing Spline Models. *Journal of Statistical Computation and Simulation*. 74, 581-597.
 35. J. Zhang and **A. Liu** (2003). Local Polynomial Fitting Based on Empirical Likelihood. *Bernoulli*. 9, 579-605.

PUBLICATIONS in REFEREED CONFERENCES

- E Huang, L Palma, L Cetinsoy, Y Diao, **A Liu** (2019). AIDEme: An active learning based system for interactive exploration of large datasets. *NeurIPS 2019*
- L. Di Palma, Y. Diao, and **A Liu** (2019). A Factorized Version Space Algorithm for 'Human-In-the-Loop' Data Exploration. *ICDM - 19th IEEE International Conference in Data Mining*
- T. Tran, L. Peng, B. Li, Y. Diao, and **A. Liu** (2010). PODS: A New Model and Processing Algorithms for Uncertain Data Streams. *Proceedings of ACM SIGMOD Conference on Management of Data (SIGMOD 2010)*(acceptance rate: 20.8%)
- Y. Diao, B. Li, **A. Liu**, L. Peng, C. Sutton, T. Tran, and M. Zink (2009). Capturing Uncertainty in High-Volume Stream Processing. *Proceedings of the 4th Biennial Conference on Innovative Data Systems Research (CIDR 2009)*(acceptance rate: 27.1%)

INVITED PRESENTATIONS

- Smoothing Spline Mixed-Effects Density Models for Clustered Data. The 32nd New England Statistics Symposium (NESS), April 2018.
- Spline Density Estimation and Inference with Model-Based Penalties. 2017 ICSA China Conference, Beihau University, Jilin, China. July 2017.
- Variable Selection in Single-index Varying Coefficient Models. The 4th Institute of Mathematical Statistics, Asia Pacific Rim Meeting 2016 (IMS-APRM 2016), The Chinese University of Hong Kong. June 2016.
- My research on variable selection, density estimation and interactive data exploration. Workshop on joint MassMutual/UMass projects, MassMutual Data Labs. December 2015.
- Variable Selection in Single-index Varying Coefficient Models. The 29th New England Statistical Symposium, University of Connecticut. April 2015.
- Two-Part Models with Proportionality Constraint: Analysis of The Multi-Ethnic Study of Atherosclerosis (MESA). 2014 ICSA (International Chinese Statistical Association) China Statistics Conference, Shanghai China. July 2014.
- Determination of Proportionality in Two-part Models and Analysis of Multi-Ethnic Study of Atherosclerosis (MESA). The 27th New England Statistics Symposium, April 2013
- M-Type Smoothing Spline ANOVA for Correlated Data, Department of Biostatistics, Harvard University, Boston, MA, October 2010
- M-Type Smoothing Spline ANOVA for Correlated Data, WNAR, Seattle, WA, June 2010
- Generalized Spline Mixed-effects Models with Applications in AIDS Clinical Trials, International Indian Statistical Association (IISA) Conference "Frontiers of Probability and Statistical Science", Uconn, Storrs, CT, May 2008
- Modeling of Hormone Secretion-Generating Mechanisms with Splines: A Pseudo-Likelihood Approach, Department of Mathematics and Statistics, Boston University, Boston, MA, December 2007
- Evaluating Gene Expression of Environmentally Relevant Microorganisms Using Electrochemically Synthesized Oligonucleotide Microarrays, Norwegian Food Research Institute - MATFORSK, Ås, Norway, March 2007
- Microarray Data Analysis and Testing Non-Difference of Curves, Department of Food Science, UMASS, Amherst, MA, October 2006
- Smoothing Spline Models, Diagnostics and Applications, Department of Mathematics and Statistics, UMASS, Amherst, MA, November 2006

Modeling of Hormone Secretion-Generating Mechanisms with Splines: A Pseudo-Likelihood Approach, Fred Hutchinson Cancer Research Center, Seattle, WA, July 2006

Modeling of hormone secretion-generating mechanisms with Splines: A Pseudo-Likelihood Approach, Joint Statistical Meetings, Minneapolis, MN, August 2005

Hypothesis Testing in Smoothing Spline Models, the Eighth North American Meeting of New Researchers in Statistics and Applied Probability, Minneapolis, MN, August 2005

Hypothesis Testing in Smoothing Spline Models, Department of Statistics at University of Connecticut, Storrs, CT, March 2005

Hypothesis Testing in Smoothing Spline Models, Department of Mathematics and Statistics at University of Massachusetts, Amherst, MA, February 2004

Introduction to S-Plus. Quantitative Methods in the Social Sciences (QMSS) Colloquium, University of California, Santa Barbara, CA, March 2003

PROFESSIONAL SOCIETIES

The American Statistical Association

The New England Statistical Society

The International Statistical Institute

The International Chinese Statistical Association

HONORS AND AWARDS

IMS travel grant The 8th New Researcher's Conference in Statistics and Applied Probability
Minneapolis, MN, 2005

Ruth and Joe Gani Prize for Best Ph.D. Dissertation in Statistics, University of California
at Santa Barbara, 2004

University of California Regents Fellowship 2000, 2001

Wald Memorial Prize for Best Performance in Statistics Qualifying Examinations 2001

STUDENT ADVISING

Dissertation Committee Chair:

Jacob Gagnon, Math and Stat, Ph.D., September 2010

Peng Wang, Math and Stat, Ph.D., September 2015

Joy Yu, Math and Stat, Ph.D., September 2019

Weilong Hu, Math and Stat, Ph.D., December 2019

Meilan Chen, Math and Stat, Ph.D., In progress

Undergraduate advising:

Amy Pinkos, Internship with New York Life Insurance, Spring 2006

Gabriel Clara, Jing Feng, Summer REU 2016

COURSES TAUGHT at UMASS

STAT 501 – Methods of Applied Statistics
 STAT 515 – Statistics I
 STAT 516 – Statistics II
 STAT 525 – Regression Analysis
 STAT 535 – Statistics Computing
 STAT 597TS– Time Series
 STAT 598C – Statistical Consulting Practicum
 STAT 607 – Mathematical Statistics I
 STAT 608 – Mathematical Statistics II
 STAT 697L– Categorical Data Analysis
 STAT 697D– Applied Data Analysis
 STAT 697F– Special Topics in Regression - Statistical Learning

PROFESSIONAL SERVICES

Journal Review Journal of Computational and Graphical Statistics, Journal of the American Statistical Association, Canadian Journal of Statistics, Annals of Statistics, Journal of Nonparametric Statistics, Biometrics, BMC Bioinformatics, Biometrical Journal
Reviewer for a tenure promotion case 2015
Grant Panel Statistics panel, NSF, 2009, 2016
Program Committee ENAR 2011
Organizer, co-chair The 32th New England Statistics Symposium, 2018

DEPARTMENTAL and CAMPUS-WIDE SERVICES

Actuarial Committee: developed the undergraduate actuarial concentration with Michael Sullivan 2010
Interdisciplinary Faculty Search Committee: the computational social sciences cluster in 2010, and the large data analytics position with IALS 2018
Faculty Search Committee in Statistics and predictive modeling: 2005-2008, 2013-2017
Faculty Search Committee Chair in Statistics: 2017
Personnel Committee: 2012
Statistics Coordinator and Statistics advising: 2014-2018. Graduate student advising in statistics; assist administration in terms teaching assignments in statistics and changes, coordinate with biostatistics in terms of course offerings, revised statistics MS requirements and obtained approval from faculty senate; revised statistics PhD requirements and obtained approval from GAC; in the process of revising our undergraduate concentration in statistics; modify graduate handbook and webpage according to these changes.
Five college statistics liaison: 2016. Served as the math&stat representative to participate the five college statistics group meetings. During the meetings, we discussed issues on how to involve with the new data science center at UMASS, cooperation among campuses in terms of courses, inviting visitors and organizing competition such as dataFest, and nominating a junior or senior to be awarded the Five College Statistics Prize.
Graduate Admission Committee: 2009
Undergraduate Affairs Committee Member: 2008-2010, 2013
Graduate Affairs Committee Member: 2014-2017

