

# Yuling Yao

162 5th Ave · New York NY 10010 · yyao@yyao.dev · <https://yyao.dev>

## Employment

---

**Flatiron Institute, Center for Computational Mathematics, New York, NY** *July 2021–present*  
Flatiron Research Fellow of statistics and machine learning

## Education

---

**Columbia University, New York, NY** *2015–2021*  
Ph.D. Statistics  
Advisor: Andrew Gelman  
Dissertation: Toward a scalable Bayesian workflow.

**Tsinghua University, Beijing, China** *2011–2015*  
B.S. Mathematics, with distinction  
B.Ec. Economics

## Research keywords

---

Bayesian statistics, computing, applied hierarchical modeling, machine learning, model averaging and stacking, causal inference, variational Bayes, importance sampling, MCMC, cross-validation.

## Published papers

---

1. **Y. Yao**, L. Carvalho, D. Mesquita (2022). "Locking and Quacking: Stacking Bayesian models predictions by log-pooling and superposition." In *NeurIPS Workshop 'Score-Based Methods'*.
2. A. Vehtari, D. Simpson, A. Gelman, **Y. Yao**, J. Gabry. (2022). "Pareto smoothed importance sampling." *Journal of Machine Learning Research*, conditional acceptance.
3. **Y. Yao**, A. Vehtari. (2021). "Make cross-validation Bayes again." In *NeurIPS Workshop 'Your model is wrong'*.
4. **Y. Yao**, A. Vehtari, A. Gelman. (2022). "Stacking for non-mixing Bayesian computations: The curse and blessing of multimodal posteriors." *Journal of Machine Learning Research*.
5. P. Barnwal, **Y. Yao** (co-first author), Y. Wang, N. A. Juy, S. Raihan, M. A. Haque, A van Geen. (2021). "Assessment of excess mortality and household income in rural Bangladesh during the Covid-19 pandemic in 2020." *JAMA Network Open*.
6. Y. Su, D. Lien, **Y. Yao**. (2021). "Economic growth and happiness in China: A Bayesian multilevel age-period-cohort analysis based on the CGSS data 2005-2015.", *International Review of Economics and Finance*.
7. **Y. Yao**, G. Pirš, A. Vehtari, A. Gelman. (2021). "Bayesian hierarchical stacking: Some models are (somewhere) useful." *Bayesian Analysis*. (Invited by the journal Editor to the JSM BA spotlight).
8. A. Gelman, **Y. Yao**. (2020). "Holes in Bayesian statistics." *Journal of Physics G*.
9. A. van Geen, **Y. Yao**, T. Ellis, A. Gelman. (2020). "Fallout of lead over Paris from the 2019 Notre-Dame Cathedral fire." *Geohealth*.
10. **Y. Yao**. (2020). "Bayesian aggregation." *Wiley StatsRef*.

11. A. Vehtari, D. Simpson, Y. Yao, A. Gelman. (2018). "Limitations of 'limitations of Bayesian leave-one-out cross-validation for model selection'." *Computational Brain & Behavior*.
12. Y. Yao, A. Vehtari, D. Simpson, A. Gelman. (2018). "Yes, but did it work?: Evaluating variational inference." In *International Conference on Machine Learning*.
13. Y. Yao, A. Vehtari, D. Simpson, A. Gelman. (2018). "Using stacking to average Bayesian predictive distributions (with discussions and rejoinder)." *Bayesian Analysis*.
14. M. Marsman, D. Schonbrodt, D. Morey, Y. Yao, A. Gelman, E-J. Wagenmakers. (2017). "A Bayesian bird's eye view of 'replications of important results in social psychology'." *Royal Society Open Science*.

## Preprints submitted for publications

---

1. Y. Yao, J. Domke. (2023). "Discriminative calibration." *Under review*.
2. B. Clarke, Y. Yao. (2023). "A cheat sheet for Bayesian prediction." *Under review*.
3. Y. Yao, R. Mozumder, B. Bostick, B. Mailloux, C. Harvey, A. Gelman, A. van Geen. "Making the most of imprecise measurements: Changing patterns of arsenic concentrations in shallow wells of Bangladesh from laboratory and field data." *Under review*.
4. Y. Yao, C. Cademartori, A. Vehtari, A. Gelman. "Adaptive path sampling in metastable posterior distribution." *Under review*.
5. A. Gelman, A. Vehtari, D. Simpson, C. Margossian, B. Carpenter, Y. Yao, P-C. Bürkner, L. Kennedy, J. Gabry, M. Modrák. "Bayesian workflow." *Under review*.

## Open source softwares

---

- loo. (R package, co-author) 2017-present  
Efficient approximate leave-one-out cross-validation using Pareto smoothed importance sampling.  
Two million downloads on CRAN.
- Stan. (core developer team) 2017-present  
A state-of-the-art platform for statistical modeling and high-performance statistical programming.

## Professional service

---

- *Reviewed articles for journals:*  
Journal of the American Statistical Association, Journal of the Royal Statistical Society: Series B & Series C, Journal of Machine Learning Research, Annual of Applied Statistics, Statistics and Computing, Journal of Computational and Graphical Statistics, Journal of Nonparametric Statistics, Science Advance, Neural Processing Letters, Philosophy of Science, Journal of Educational and Behavioral Statistics, Electronic Journals of Statistics.
- *Reviewed articles for conferences:*  
NeurIPS 2018, 2021; ICLR 2019, 2021; ICML 2019, 2020.
- *Reviewed book proposal:*  
CRC press.

## Conference presentation

---

- *Oral presentations* in ICML 2018, JSM 2019, StanCon 2018 & 2019, CMStatistics 2020, Queensland University of Technology 2021, JSM 2022, Bayesian, Fiducial and Frequentist (BFF) 2023, Information and Statistics for Nuclear Experiment and Theory (ISNET) 2023.
- *Poster presentations* in JSM 2018, Neurips 2021, 2022.

## Teaching and training

---

- *Teaching assistant* of statistical inference, probability theory, linear regression, statistical machine learning, nonparametric statistics, stochastic method in finance, survival analysis, sample survey, Bayesian statistics.
- *Recitation instructor* of intro to statistics and probability, statistical machine learning, stochastic method in finance, Bayesian statistics.
- *Mentorship*: Supervised three undergraduate research projects and paid internship in Columbia University and Faltiron Institute, co-supervised a visiting PhD student internship.

## Media appearance

---

1. "Lead fallout from Norte Dame fire was likely overlooked," Interviewed by *State of the Planet*, Columbia University news site, 9 July 2020.
2. "Une étude relance la question de l'exposition au plomb: Des chercheurs de l'université Columbia suggèrent que des Parisiens auraient été exposés à des niveaux très supérieurs aux données officielles," *Le Monde*, 11 July 2020.
3. "Where Covid-19's death grip slipped (briefly)," *State of the Planet*, Columbia University news site, 15 November 2021.

## Blog

---

1. Blog on statistics and machine learning, <https://yulingyao.com/blog>
2. Regular guest posts on "Statistical Modeling, Causal Inference and Social Science" Gelman et al., <https://statmodeling.stat.columbia.edu/author/yulingyao/>