Colloquium Schedule Spring 2026

- STAT 5924, CRN 20257
- Fridays
- 2:30 pm to 3:30 pm
- 300 Seitz

January 23, 2026: Shounak Chattopadhyay, University of Virginia

January 29, 2026, Thursday, 3:30 - 4:30 pm: Donald Richards, Penn State

Canceled: February 6, 2026: Kaoru Irie, University of Tokyo

February 12, 2026, Thursday, 3:30 - 4:30 pm: Vasileios Maroulas, University of Tennessee

February 19, 2026, Thursday, 3:30 - 4:30 pm: Yuzhou Chen, UC Riverside

February 27, 2026: Piotr Kokoszka, Colorado State University

March 6, 2026: Anh Bui, Virginia Commonwealth University

March 17, 2026, Tuesday 3:30 - 5 pm: Martha Gardner, GE Aviation, Douglas C. Montgomery Distinguished Lecture, Holtzman Alumni Center Auditorium

March 20, 2026: Mental Well-being with Hokie Wellness

March 27, 2026: Jordan Awan, University of Pittsburgh

April 3, 2026:

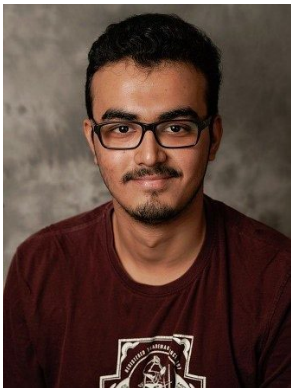
April 10, 2026: Toryn Schafer, Texas A&M

April 17, 2026: Anne van Delft, Columbia University

April 24, 2026: Mirjeta Pasha, Department of Mathematics, Virginia Tech

May 1, 2026: Yu Ding, Georgia Tech

January 23, 2026: Shounak Chattopadhyay, University of Virginia, Location: 300 Seitz



Bio: An assistant Professor in the Department of Statistics, University of Virginia. Prior to joining UVA, I was a Postdoctoral Scholar at the University of California, Los Angeles, working with Dr. Marc A. Suchard. I completed my Ph. D. at the Department of Statistical Science, Duke University, under the supervision of Dr. David B. Dunson. Before Duke, I completed my Bachelors and Masters degrees in Statistics from the Indian Statistical Institute, Kolkata.
https://shounakch.github.io/

Title: Blessing of dimension in Bayesian inference on covariance matrices

Abstract: Bayesian factor analysis is routinely used for dimensionality reduction in modeling of high-dimensional covariance matrices. Factor analytic decompositions express the covariance as a sum of a low rank and diagonal matrix. In practice, Gibbs sampling algorithms are typically used for posterior computation, alternating between updating the latent factors, loadings, and residual variances. In this article, we exploit a blessing of dimensionality to develop a provably accurate posterior approximation for the covariance matrix that bypasses the need for Gibbs or other variants of Markov chain Monte Carlo sampling. Our proposed Factor Analysis with BLEssing of dimensionality (FABLE) approach relies on a first-stage singular value decomposition (SVD) to estimate the latent factors, and then defines a jointly conjugate prior for the loadings and residual variances. The accuracy of the resulting posterior approximation for the covariance improves with increasing samples as well as increasing dimensionality. We show that FABLE has excellent performance in high-dimensional covariance matrix estimation, including producing well-calibrated credible intervals, both theoretically and through simulation experiments. We also demonstrate the strength of our approach in terms of accurate inference and computational efficiency by applying it to a gene expression dataset.

January 29, 2026, Thursday, 3:30 - 4:30 pm: Donald Richards, Penn State, Location: 300 Seitz

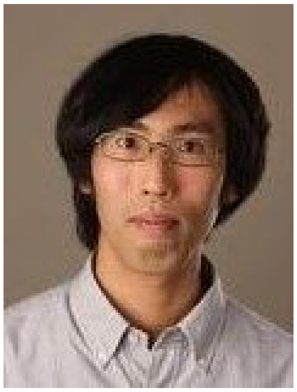
Bio: Donald Richards is a Distinguished Professor Emeritus of Statistics at Penn State University. Upon receiving his Ph.D. from the University of the West Indies (UWI), Dr. Richards held academic positions at the University of North Carolina at Chapel Hill, the University of Virginia where he served as Department Head, and Penn State University. Dr. Richards has been serving as Board Member of the Institute for Mathematics and its Applications at the University of Minnesota, School of Mathematics, Institute for Advanced Study at Princeton University, National Research Council (NRC), and most recently, the Statistics Review Committee of Centers for Disease Control and Prevention (CDC), just to name a few. Richards' research interests include algebraic techniques in multivariate statistical analysis, combinatorics, probability inequalities, value investing, and financial derivatives. He applies these methods to topics in astronomy and astrophysics, finance, actuarial science, and medical imaging. Richards is a Fellow of the Institute of Mathematical Statistics, Fellow of the American Mathematical Society, and a Member of the International Statistical Institute.

https://science.psu.edu/stat/people/dsr11

Title: From Divination to Gambling to Annuities

Abstract: This talk traces the history of the development of annuities, starting from humans' earliest efforts at divination, proceeding to the study of gambling, and thence to annuities. Along the way, we will see how divination has occurred on every populated continent, and similarly for gambling. Finally, we will describe how annuities were developed by some of the earliest practitioners of mathematical probability.

Canceled: February 6, 2026: Kaoru Irie, University of Tokyo, Location: 300 Seitz



Bio: Associate Professor of Economics
Visiting scientist (Since December 2024)
Statistical Mathematics Collaboration Unit
Center for Brain Science
RIKEN

Research Interests
Bayesian Statistics and Econometrics
State space models; Sequential Monte Carlo; Stochastic volatility;
Shrinkage priors; Modeling of integer-valued time series;
Outlier-robust posterior inference.

Title: Outlier-robust posterior inference for linear models (Kaoru Irie, Associate Professor, Faculty of Economics, the University of Tokyo)

Abstract: Robust statistics, or outlier-robust inference, has been a focus of statistical research for many years and has produced abundant research results. In the Bayesian framework, however, research on outlier-robust "posterior" inference has advanced dramatically in the last 10 years. The key concept in modeling outlier-contaminated observations is the use of a super-heavy-tailed distribution, such as the log-Pareto distribution, as the error distribution. It has also been shown across multiple settings that the Student's t-distribution is not sufficiently heavy-tailed to accommodate outliers without affecting the posterior distribution.

In this talk, we review the recent literature on so-called posterior robustness, then showcase a series of our research, focusing on three classes of models: hierarchical linear models, Poisson/negative binomial linear models, and multivariate extensions. First, we construct a new super-heavy-tailed error distribution for hierarchical linear models using the variance mixture of normals, which is conditionally conjugate and allows a Gibbs sampler, while theoretically guaranteeing robustness of the posterior distribution.

Second, we consider the generalized linear models, specifically the Poisson/negative binomial sampling distributions for count-valued responses. In this research, we treat inflated zeros and small counts as outliers, alongside extremely large counts, within a robust statistics framework. Third, we extend the use of super-heavy-tailed distributions to multivariate linear models, including graphical models. To protect the correlation parameter against the effects of outliers, we propose introducing a real-valued latent variable and multiplying it by the standard deviation parameter based on the variance-correlation decomposition. Other covariance models are shown to lack posterior robustness or to require stronger assumptions for robustness.

This is joint work with Yasuyuki Hamura (Kyoto U) and Shonosuke Sugasawa (Keio U).

February 12, 2026, Thursday, 3:30 - 4:30 pm: Vasileios Maroulas, University of Tennessee, Location: 300 Seitz

Bio: Dr. Vasileios Maroulas is the Associate Vice Chancellor for Research, Innovation, and Economic Development and Director of AI Tennessee at the University of Tennessee, Knoxville. He is a Professor of Mathematics with joint appointments in Business Analytics and Statistics and Data Science and Engineering. A nationally recognized leader in artificial intelligence, Dr. Maroulas leads AI Tennessee, a statewide initiative advancing AI research, workforce development, and economic competitiveness. His work spans machine learning, statistics, and data science, with applications across industry, national security, healthcare, and science.

Title: Random Persistence Diagram Generator

Abstract: Topological data analysis (TDA) is a way of understanding the shape of data, how patterns, holes, and structures emerge across different scales. One of its most widely used tools, persistent homology, captures these structures and represents them in persistence diagrams, which summarize how features appear and disappear as we vary the scale of analysis.

In this talk, we introduce a Random Persistence Diagram Generator (RPDG), a new method that generates realistic sequences of persistence diagrams directly from those produced by data. RPDG is grounded in a principled probabilistic model based on interacting point processes and is implemented using a reversible-jump Markov chain Monte Carlo approach.

We demonstrate the power of RPDG in two settings. First, using synthetic data, we show that RPDG effectively captures the variability of persistence diagrams and compares favorably with existing sampling methods. Second, we apply RPDG to a real materials science problem, illustrating how it enables meaningful inference even when only a small amount of data is available.

February 19, 2026, Thursday, 3:30 - 4:30 pm: Yuzhou Chen, UC Riverside, Location: 300 Seitz

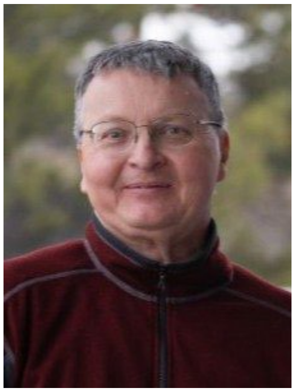
Bio: Dr. Yuzhou Chen is an Assistant Professor in the Department of Statistics at UC Riverside. He is also an affiliated faculty member in the Department of Electrical and Computer Engineering at UC Riverside and a Visiting Research Collaborator in the Department of Electrical and Computer Engineering at Princeton University. Before that, Dr. Chen was a postdoctoral scholar in the Department of Electrical and Computer Engineering at Princeton University and received his Ph.D. in Statistics from Southern Methodist University. His research agenda is at the intersection of statistical learning
of graphs and spatio-temporal processes, geometric deep learning, and topological data analysis, with applications to power systems, biosurveillance, finance, and environmental data analytics. His research has appeared in the top ML/AI and statistical venues, including ICML, ICLR, NeurIPS, KDD, AAAI, JRSS, and PNAS, and received multiple best paper awards from the American Statistical Association. He is the recipient of the Chateaubriand Fellowship from the Embassy of France in the United States.

UCR Regents Faculty Fellowship, and Leonard Transportation Center Research Faculty Fellowship.

Title: Frontiers of Topology-guided Machine Learning Models for Spatio-temporal Data Learning

Abstract: Spatio-temporal event data play a central role in many modern scientific and societal applications, including urban mobility, public health, financial systems, and blockchain ecosystems. Such data are inherently relational, dynamically evolving, and non-Euclidean in nature, which poses fundamental challenges for classical statistical models and standard deep learning approaches. In this talk, I will introduce novel topology-guided machine learning frameworks that integrate spatio-temporal modeling, graph representation learning, and topological data analysis to address these challenges in a principled manner. First, I introduce a topology-enhanced diffusion framework for spatio-temporal point processes, which captures complex event dependencies through joint spatio-temporal graph construction and topological representations and leads to improved predictive performance and interpretability. Second, I discuss a multilayer topology-aware graph contrastive learning framework for fraud detection in time-evolving blockchain transaction networks, where persistent homology is employed to encode higher-order structural patterns beyond pairwise interactions. Together, these works demonstrate how topological structure can serve as a unifying analytical lens for learning from complex spatio-temporal data, and offer both practical performance gains and deeper statistical insight into the organization and dynamics of real-world complex systems.

February 27, 2026: Piotr Kokoszka, Colorado State University, Location: 300 Seitz



Bio: Professor Piotr Kokoszka received PhD in 1993 at Boston University. He has made contributions to the theory of stochastic processes, time series analysis, spatial statistics, functional data analysis and applications to geophysics, finance and power systems. He has coauthored almost 200 papers in these fields, as well as a monograph and a textbook on functional data analysis. He has served on many editorial boards, including the Journal of the American Statistical Association and the Journal of the Royal Statistical Society (B).

Title: Deep learning estimation of the spectral density of functional time series on large domains

Abstract: The talk is concerned with the estimation of the spectral density of a time series of surfaces or images defined on large grids of hundreds of thousands of points. Such sequences occur in biomedical and climate studies. The estimator is the output of a multilayer perceptron neural network. Existing estimators use sample autocovariance kernels represented as high-dimensional matrices whose manipulation is computationally demanding or even infeasible. We use the theory of spectral functional principal components to derive our deep learning estimator and prove that it is a universal approximator to the spectral density under general assumptions. It can be trained without computing the autocovariance kernels and can be parallelized to provide the estimates much faster than existing approaches. We validate its performance by simulations and an application to fMRI images.

March 6, 2026: Anh Bui, Virginia Commonwealth University, Location: 300 Seitz

Bio: Anh T. Bui received the Ph.D. degree in industrial engineering & management sciences from Northwestern University. He is currently an assistant professor in the Department of Statistical Sciences & Operations Research at Virginia Commonwealth University. His broad research interests lie in statistics and machine learning, for improving the understanding and quality of manufacturing, healthcare, and other complex systems. His main research theme has revolved around data reduction related problems, such as representation learning, dimension reduction, and data sampling. Dr. Bui is a recipient of the Lloyd S. Nelson Award and the Frank Wilcoxon Award from the American Society for Quality.

Title: Conditional Dimension Reduction for Dissimilarity Data

Abstract: Dimension reduction aims to discover underlying features that govern data variation. However, certain underlying features are often known in most applications. Conventional dimension reduction methods do not utilize information of these known features. This leads to two potential problems: (i) failing to utilize available information generally leads to poorer solutions, and (ii) the known features may mask the unknown features in the reduced-dimension space. In this talk, I will discuss how conditional dimension reduction can address these problems to improve statistical estimation and simplify knowledge discovery tasks. The emphasis is on applications that involve dissimilarity data. In particular, methods for conditional multidimensional scaling (MDS) will be presented. This includes an iterative/majorization optimization approach to minimize a conditional stress objective function and algorithms for handling incomplete conditioning data. Theoretical properties of these techniques are studied under mild assumptions. The benefits of these methods are demonstrated with both simulated and real examples. Computer codes for conditional MDS are maintained in the cml R package on CRAN

March 17, 2026, Tuesday, 3:30 - 5 pm: Martha Gardner, GE Aviation, Douglas C. Montgomery Distinguished Lecture, Holtzman Alumni Center Auditorium, Location: Holtzman Alumni Center Auditorium

Bio: Dr. Martha M. Gardner is a Technology Executive focused on Quality, Data Science and Applied Statistics at GE Aerospace. She began her career as a Statistician in the Applied Statistics Laboratory at GE Research and spent 20 years there applying statistics, quality, and innovation methods for developing new products and improving manufacturing processes. She joined GE Aerospace as a Quality Leader and then led the Edison Engineering Development Program prior to her current role. Martha is a Fellow of the American Statistical Association and the 2020 recipient of the Gerald J. Hahn Quality and Productivity Achievement Award.

Title: Statistical Engineering in the Wild: Choose Your Own Adventure

Abstract: Statistics is often introduced as a collection of formulas, but in practice it is a way of thinking – one that shows up everywhere from industrial applications, public policy, to sports, medicine, and every day decision making. A career as an industrial statistician is shaped less by specific techniques and more by having problem solving focus, strong collaborations with engineers and scientists, and comfort with handling and communicating uncertainty; traits which align with what I have learned from Professor Montgomery's books and presentations over the years. In this talk, I'll share a career-long tour of industrial statistical engineering in the wild. We'll explore a few engineering applications, talk about how my experiences along the way influenced my career directions, moving between statistics and engineering roles, and what future possibilities might be in our expanded world of AI.

March 20, 2026: Mental Well-being with Hokie Wellness, Location: 300 Seitz

Hokie Wellness
Our team offers a variety of educational services, programming, and resources to enhance the health and well-being of all Hokies.
https://hokiewellness.vt.edu/

Title: Mental Well-being with Hokie Wellness

Join Hokie Wellness in an interactive session to discuss true self-care and ways to support the mental well-being of students, faculty, and staff. We will discuss and demonstrate various concrete skills, help you develop at least one attainable self-care goal, and provide a list of easily accessible mental wellness resources for yourself and others. All are welcome!

March 27, 2026: Jordan Awan, University of Pittsburgh, Location: 300 Seitz

Bio: I am an Assistant Professor of Statistics at the University of Pittsburgh. Before joining Pitt, I was an Assistant Professor of Statistics at Purdue University from 2020-2025. I completed my Ph.D. in Statistics at Penn State University in 2020, under the advisement of Dr. Aleksandra Slavkovic and Dr. Matthew Reimherr. My primary research interest is in data privacy, where the goal is to publish meaningful statistical results on sensitive datasets without compromising the privacy of the participants in the dataset.

Title: Beyond Data Splitting: Full-Data Conformal Prediction by Differential Privacy

Abstract: Privacy protection and uncertainty quantification are increasingly important in data-driven decision making. Conformal prediction provides finite-sample marginal coverage, but existing private approaches often rely on data splitting, reducing the effective sample size. We propose a full-data privacy-preserving conformal prediction framework that avoids splitting. Our framework leverages stability induced by differential privacy to control the gap between in-sample and out-of-sample conformal scores, and pairs this with a conservative private quantile routine designed to prevent under-coverage. We show that a generic differential privacy guarantee yields a universal coverage floor, yet cannot generally recover the nominal $1 - \alpha$ level. We then provide a refined, mechanism-specific stability analysis that yields asymptotic recovery of the nominal level. Experiments demonstrate sharper prediction sets than the split-based private baseline.

April 3, 2026:

April 10, 2026: Toryn Schafer, Texas A&M, Location: 300 Seitz

Bio: Toryn Schafer is an Assistant Professor in the Department of Statistics at Texas A&M University. She was previously a postdoctoral associate in the Department of Statistics and Data Science at Cornell University and earned her PhD in Statistics from the University of Missouri in 2020. Her research interests include Bayesian spatio-temporal modeling, statistical computing, and applied work in ecological and environmental sciences.

Title: Scalable Bayesian Spatial Modeling and Agent-Based Simulation for Ecological Decision Support

Abstract: Ecological management problems often require linking uncertain data to scenario-based predictions about how human activity may affect wildlife. This is challenging when observations are sparse, collected across multiple observation windows, and the quantities needed for decisions are not direct outputs of standard statistical models. In this talk I will describe a modeling framework that connects Bayesian spatial inference with simulation to support decision-making under uncertainty

The first component is a computationally efficient, recursive Bayesian approach for fitting spatial point process models to point pattern data derived from long-term monitoring and aerial imagery with non-overlapping sampling frames. This provides flexible estimates of spatial structure and yields posterior predictive spatial fields that can be carried forward into downstream analyses. The second component uses these posterior predictions to initialize a custom agent-based model that represents interactions among environmental conditions, human activity, and animal responses, with outputs summarized through disturbance metrics. The agent-based model is built as a modular system so that alternative submodels can be incorporated, including demographic components such as matrix population models and decision-oriented components such as Markov decision processes.

April 17, 2026: Anne van Delft, Columbia University, Location: 300 Seitz

Bio: Anne van Delft joined Columbia University in Fall 2020 as a Tenure Track Assistant Professor of Statistics. Her research focuses on stochastic processes that take values in function spaces, and in particular on the development of theory and methodology for function-valued time series with nonstationary characteristics. This applies to the analysis of sequential collections of data points that themselves come in the form of complex mathematical structures, such as curves, surfaces or manifolds. Examples can be found in (neuro-)imaging, climatology, genomics, and econometrics. Inference techniques to analyze such data not only require a mathematically rigorous and quantitative study of the geometric features and temporal-spatial dependence structure, but moreover must translate into computationally efficient methods.

Anne obtained her PhD at Maastricht University, the Netherlands, in December 2016. Before joining the Department of Statistics at Columbia University, she held a postdoctoral fellowship in mathematical statistics at the Ruhr University in Bochum, Germany. She was the recipient of the Itô prize in 2021. Anne is currently Associate Editor for Journal of the American Statistical Association Reviews

Title: A statistical framework for analyzing shape in a time series of random geometric objects

Abstract: We introduce a new framework to analyze shape descriptors that capture the geometric features of an ensemble of point clouds. At the core of our approach is the point of view that the data arises as sampled recordings from a metric space-valued stochastic process, possibly of nonstationary nature, thereby integrating geometric data analysis into the realm of functional time series analysis. Our framework allows for natural incorporation of spatial-temporal dynamics, heterogeneous sampling, and the study of convergence rates. Further, we derive complete invariants for classes of metric space-valued stochastic processes in the spirit of Gromov, and relate these invariants to so-called ball volume processes. Under mild dependence conditions, a weak invariance principle in $D([0,1]\times [0,\mathscr{R}])$ is established for sequential empirical versions of the latter, assuming the probabilistic structure possibly changes over time. Finally, we use this result to introduce novel test statistics for topological change, which are distribution-free in the limit under the hypothesis of stationarity. We explore these test statistics on time series of single-cell mRNA expression data, using shape descriptors coming from topological data analysis.

April 24, 2026: Mirjeta Pasha, Department of Mathematics, Virginia Tech, Location: 300 Seitz

Bio: Mirjeta Pasha is an Assistant Professor of Mathematics at Virginia Tech. Her research focuses on developing scalable computational methods for large-scale inverse problems and high-dimensional data analysis, with a particular emphasis on tensor-based models and uncertainty quantification. She was previously a postdoctoral researcher at MIT, Tufts University, and Arizona State University, where she worked on tensor methods, uncertainty quantification, and inverse problems. Her work aims to enable efficient and reliable inference in complex, data-rich applications arising in science and engineering.

Title: From Deterministic Modeling to Bayesian Inference: A Computational Journey through Large-Scale Inverse Problems

Abstract: Rapidly-growing fields such as data science, uncertainty quantification, and machine learning rely on fast and accurate methods for inverse problems. Three emerging challenges on obtaining relevant solutions to large-scale and data-intensive inverse problems are ill-posedness of the problem, large dimensionality of the parameters, and the complexity of the model constraints. Tackling the immediate challenges that arise from growing model complexities (spatiotemporal measurements) and data-intensive studies (large-scale and high-dimensional measurements collected as time-series), state-of-the-art methods can easily exceed their limits of applicability. In this talk we discuss efficient methods for computing solutions to dynamic inverse problems, where both the quantities of interest and the forward operator may change at different time instances. We consider large-scale ill-posed problems that are made more challenging by their dynamic nature and, possibly, by the limited amount of available data per measurement step. In the first part of the talk, to remedy these difficulties, we apply efficient regularization methods that enforce simultaneous regularization in space and time (such as edge enhancement at each time instant and proximity at consecutive time instants) and achieve this with low computational cost and enhanced accuracy. In the remainder of the talk, we focus on designing spatio-temporal Bayesian Besov priors for computing the MAP estimate in large-scale and dynamic inverse problems. Numerical examples from a wide range of applications, such as biomedical applications, tomographic reconstruction, image deblurring, and multichannel dynamic tomography are used to illustrate the effectiveness of the described approaches.

May 1, 2026: Yu Ding, Georgia Tech, Location: 300 Seitz

Bio: Dr. Yu Ding is the Anderson-Interface Chair and Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. Prior to joining Georgia Tech, he was the Mike and Sugar Barnes Professor of Industrial and Systems Engineering at Texas A&M University. Dr. Ding's research focuses on engineering data science. He is the author of the CRC Press book, Data Science for Wind Energy and a co-author of the Springer Nature book, Data Science for Nano Image Analysis. His research work is recognized by the 2019 IISE's Technical Innovation Award and 2022 INFORMS' Impact Prize. Dr. Ding is a former Editor-in-Chief of IISE Transactions, and the current Editor-in-Chief of INFORMS Journal on Data Science. He is a Fellow of INFORMS, IISE, and ASME.

Title: Gaussian Process Model for Handling Autocorrelated Wind Data

Abstract: The speaker published a book in 2019, "Data Science for Wind Energy," in which a set of wind operational problems is framed through a general formulation of , where is the response on a turbine and is the wind and environmental input. When is the wind power output, the conditional density, , is related to the power curve model of wind turbines and can be used to characterize a wind turbine's performance and production efficiency. When is the mechanical load, the conditional density, , is related to the load estimation and reliability management for the wind turbines. Modeling and understanding the dynamics in itself is also important, because a good model of leads to better wind forecasting. The book is therefore structured in three parts: (1) wind field modeling for forecasting, (2) power curve modeling for performance analysis, and (3) turbine reliability and fault detection. In this talk, the speaker will discuss a problem related to the second part of the book, which is how to handle autocorrelated wind data while building a power curve for wind turbines.