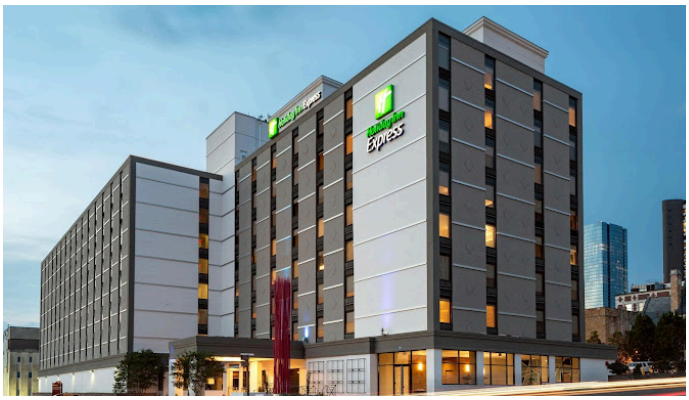
66th ANNUAL FALL TECHNICAL CONFERENCE

Harmonizing Quality, Statistics, and Data Science



Holiday Inn Express
Nashville - Broadway

Nashville, TN

October 8 - 10, 2024

Co-sponsored by:

ASQ
Excellence Through Quality™

Chemical & Process Industries Division
Statistics Division

ASA
AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

Section on Physical & Engineering Sciences
Quality & Productivity Section

Shewell Award scoring is now online!

About the Award

The Shewell Award is awarded to the best contributed paper at FTC, as judged by conference participants as well as a panel of judges after the event. The prize will be awarded at next year's conference and includes a $1,000 monetary award, free registration at next year's conference, and hotel expenses. The award is based on two broad categories of criteria:

1. Oral presentation (60%) as rated by attendee's evaluation as provided on this form
2. Written material (40%), rated by a panel of judges based on quality of subject matter, scientific merit, applicability and originality

Two ways to access

1. **Use your smartphone and the QR code**: scan the QR code below with your smartphone. Most iPhones and Android smartphones can navigate you directly to the survey website just using your camera. Alternatively, you can use a QR scanning app that you have already installed on your phone.

2. **Navigate to the following web link and bookmark it in your smartphone browser.**

https://tinyurl.com/FTCShewellAward

The Shewell Award scoring will not be done using paper forms this year.
Thanks for your understanding!

66th Annual Fall Technical Conference
Harmonizing Quality, Statistics, and Data Science

Dear Attendee,

It is my pleasure to welcome you to Nashville and the 2024 ASQ/ASA Fall Technical Conference! The Fall Technical Conference brings together researchers and practitioners from academia, industry, and government to discuss ways to more effectively use statistical methods for research, innovation, and quality improvement. The conference is co-sponsored by the American Society for Quality (Chemical & Process Industries Division and the Statistics Division) and the American Statistical Association (Section on Physical & Engineering Sciences and the Quality & Productivity Section).

This year's theme, "Harmonizing Quality, Statistics, and Data Science," was motivated by Nashville's musical roots and the need to solve real-world problems with multidisciplinary expertise. The conference begins with short courses on Tuesday that provide training in areas in quality, statistics, and data science. **An opening reception will be held Tuesday evening and I encourage more seasoned attendees to approach new faces to build the FTC fellowship for this year and the future.** Wednesday and Thursday are packed with outstanding sessions organized by the FTC technical program committee, focusing on new methods and applications in design of experiments, quality control, and artificial intelligence. The conference concludes on Thursday afternoon with a SPES reception and panel discussion titled *How to Attract and Prepare Students for Careers in Industrial Statistics*.

The plenary speakers are, as always, high points of the conference. Fred Faltin (Virginia Tech) is this year's Gerald J. Hahn Achievement Award winner and will give the opening plenary about his experiences regarding the relationship between statistics and experimental mathematics. Dr. Adam Pintar will give the WJ Youden Address, in which he will discuss the influence Jack Youden's work has had at the National Institute of Standards and Technology. I am also excited about this year's lunch speakers! On Wednesday, Kevin White from Eastman (Kingsport, TN) will be our "local" speaker, sharing Eastman's history in embracing statistics and quality. Thursday's lunch features the inaugural ASA Presidential Invited Address at FTC, given by Dr. Madhumita (Bonnie) Ghosh-Dastidar. She will provide multiple examples from health policy to highlight both valuable contributions made by statistical scientists and lessons she has learned throughout her career..

Last, but not least, I would like to thank everyone involved in this year's conference for all their hard work. Thanks to the FTC steering committee for their constant support and guidance, thanks to the Technical Program Committee led by Q&P's Ryan Lekivetz, and the Short Course Committee led by Steven Schuelka. I would also like to thank the other conference committee members who helped with so many important details: Sharad Prabhu as FTC Treasurer, John Szarka as Exhibitor & Sponsor Chair, and Adam Pintar as Webmaster. Special thanks also to Emily Griffith as the previous FTC chair.

I hope that you enjoy your time in Nashville and at FTC, especially the great presentations, networking opportunities, luncheons, and evening receptions. Whether this is your first time in Nashville or not, I encourage you to visit Broadway in downtown Nashville (just a few blocks from the conference hotel) to experience the food, live music, and atmosphere that makes Nashville one of my favorite places to visit.

Sincerely,
Jon Stallrich, NC State University
2024 FTC General Conference Chair
2024 ASA SPES Chair
jwstalli@ncsu.edu

Exhibitors & Sponsors

Special thanks to our exhibitors and sponsors for helping to continue FTC's great tradition!

Exhibitors

Exhibitors will showcase their products on Wednesday and Thursday in the Exhibitor Area outside the Arts and Theater Districts. Be sure to stop by to learn more!

EFFEX is a collaborative platform for all your Design of Experiments. Find the best experimental plan for your challenge and analyze your data with the support of unique data analytic tools. Get access to the complete catalog of cost efficient OMARS designs: less physical tests without the loss of valuable insights and knowledge.

The Statistical Sciences Group at Los Alamos National Laboratory is a world leader in advancing statistics for science and national security. We promote statistical expertise to develop and employ breakthrough methods in advanced predictive modeling, artificial intelligence, and computational and industrial statistics to solve novel interdisciplinary problems with high impact.

Sponsors

Diamond Sponsor

Since 1989, JMP software has combined interactive visualization with powerful statistics to help scientists and engineers see their data more clearly. And when data is better understood, insight-driven improvements are sure to follow. It's spelled J-M-P but pronounced "jump," like a leap in insight. JMP is a SAS company.

Silver Sponsor

Shell is a global group of energy and petrochemical companies with 93,000 employees and operations in more than 70 countries. We use advanced technologies and take an innovative approach to help build a sustainable energy future. Visit www.shell.com.

Bronze Sponsor

NC STATE UNIVERSITY

Department of Statistics

The NC State Department of Statistics, founded in 1941 by Gertrude Cox, is one of the oldest, largest, and top-ranked statistics departments in the United States. It is the birthplace of analytics powerhouse SAS Institute.

StatEase®
statistics made easy®

Stat-Ease, Inc. provides software and training for design of experiments (DOE). Industrial researchers worldwide use Design-Expert® software (DX) to make the most from every experiment. For power users, Stat-Ease® 360 software expands the horizon of DX (all features included) to Python scripting and other advanced tools. Get details at www.statease.com.

Congratulations to our student grant award winners!

Simin Zheng, *Virginia Tech**
Kevin Xie, *Virginia Tech**
Januh Heo, *Michigan State University*
William Fisher, *Clemson University*
Ayumi Mutoh, *North Carolina State University*

*We would like to thank Virginia Tech's Department of Statistics for providing the awards to these students.

Acknowledgements

FTC Steering Committee

ASQ CPID: Stephanie DeHart
ASQ STAT: Theresa Utlaut
ASA SPES: Anne Driscoll
ASA Q&P: David Edwards

2024 FTC General Committee

General Chair: Jon Stallrich
Treasurer: Sharad Prabhu
Publicity: Luke Hagar
Exhibitors & Sponsors: John Szarka
Webmaster: Adam Pintar
Program Book Editor: Jon Stallrich

2024 Technical Program Committee

ASQ STAT: Karen Hulting
ASQ CPID: Annie Booth
ASA SPES: Michael Crotty
ASA Q&P: Ryan Lekivetz (Chair)

2024 Short Course Committee

ASQ STAT: Steven Schuelka (Chair)
ASQ CPID: Jennifer Van Mullekom
ASA SPES: Steve Walsh
ASA Q&P: Peng Liu

ASQ Chemical & Process Industries Division

Chair: Katie Brickey
Chair-Elect: Shane Bookhultz
Secretary/Treasurer: Luke Hagar
Shewell Award: Jennifer Kensler
Wilcoxon & Youden Awards: Ying Chen

ASQ Statistics Division

Chair: Harish Jose
Chair-Elect: Jennifer Hellrung
Treasurer: Jennifer Hellrung
Secretary: Michael Mladjenovic

ASA Section on Physical & Engineering Sciences

Chair: Jon Stallrich
Chair-Elect: Arman Sabbaghi
Secretary & Treasurer: Youngdeok Hwang

ASA Quality & Productivity Section

Chair: Maria Weese
Chair-Elect: John Szarka
Treasurer: Sharad Prabu
Secretary: Annie Dudley

Conference Event Times & Locations

All locations are at the Holiday Inn Express Nashville - Broadway

Tuesday, October 8

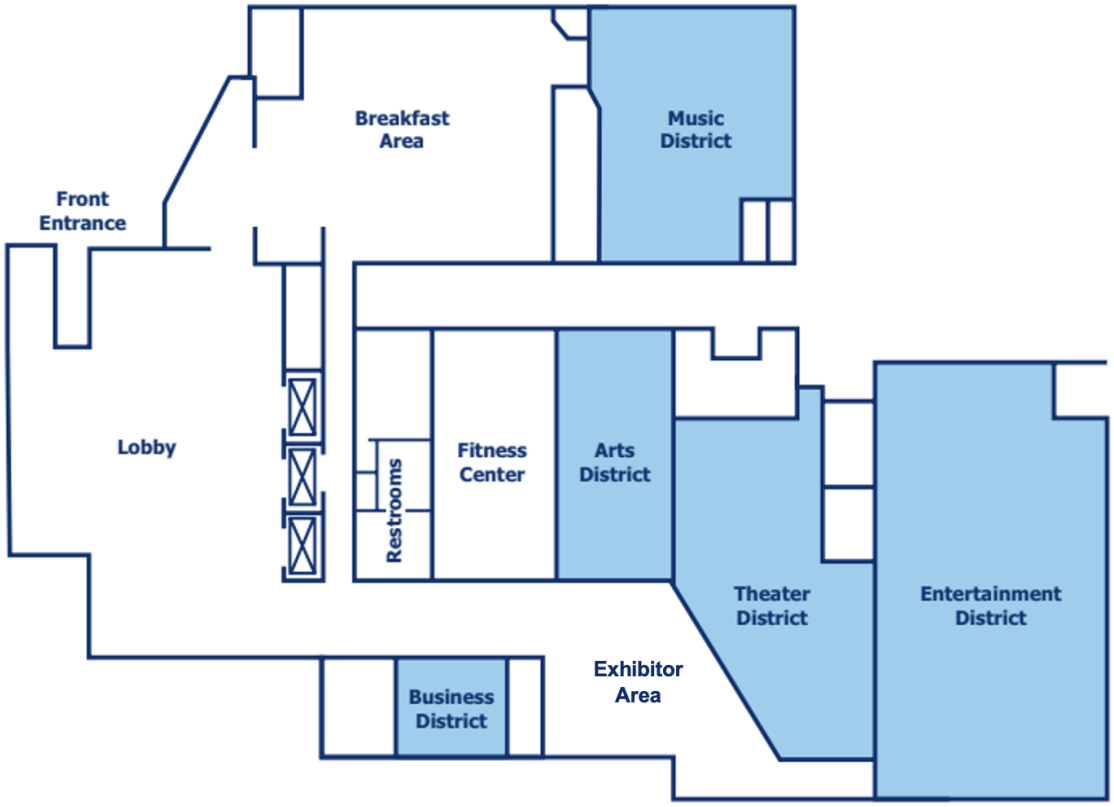
7:30am – 9:00am	Registration	Lobby
7:30am – 9:00am	Breakfast	Breakfast Area
8:30am – 5:30pm	Short Courses	Arts District
		Business District
		Music District
		Theater District
12:00pm – 1:00pm	Lunch	Breakfast Area
3:00pm – 3:15pm	Coffee Break	Exhibitor Area
6:00pm – 7:30pm	ASQ Stat. Division Reception	Breakfast Area

Wednesday, October 9

7:00am – 8:00am	Registration	Lobby
7:00am – 8:00am	Breakfast	Breakfast Area
7:30am – 5:30pm	Exhibits	Outside Arts/Theater District
8:00am – 9:00am	Welcome & Plenary Session	Entertainment District
9:15am – 10:00am	Parallel Sessions 1	A: Arts District
		B: Music District
		C: Theater District
10:00am – 10:30am	Morning Break	Exhibitor Area
10:30am – 12:00pm	Parallel Sessions 2	A: Arts District
		B: Music District
		C: Theater District
12:15pm – 1:45pm	Lunch and Plenary Session	Entertainment District
2:00pm – 3:30pm	Parallel Sessions 3	A: Arts District
		B: Music District
		C: Theater District
3:30pm – 4:00pm	Afternoon Break	Exhibitor Area
4:00pm – 5:00pm	W.J. Youden Address	Entertainment District
7:00pm –	CPID Dinner	By Invitation Only

Thursday, October 10

7:00am – 8:00am	Registration	Lobby
7:00am – 8:00am	Breakfast	Breakfast Area
7:30am – 3:00pm	Exhibits	Outside Arts/Theater District
8:00am – 9:30am	Parallel Sessions 4	A: Arts District
		B: Music District
		C: Theater District
9:30am – 9:45am	Morning Break	Exhibitor Area
10:00am – 11:30am	Parallel Sessions 5	A: Arts District
		B: Music District
		C: Theater District
11:45am – 1:15pm	Lunch and Plenary Session	Entertainment District
1:30pm – 3:00pm	Parallel Sessions 6	A: Arts District
		B: Music District
		C: Theater District
3:15pm – 5:00pm	SPES Reception and Panel Discussion	Entertainment District



Full Day Short Courses: Tuesday, October 8

Full day short courses will be held on Tuesday, October 8 from 8:00am to 5:00pm.

Modern Response Surface Methodology

Steven Gilmour and Olga Egorova, King's College London

Music District

Abstract: This course provides a brief overview of the basics of response surface methodology (RSM) and introduces the most recent work, with a particular focus on the multi-objective design of experiments. We will introduce the methodological background accompanied by the practical tools and examples (in R).

Common practical experimental setups involve limitations of experimental resources and/or restrictions in terms of their allocation, such as hard-to-change factors. At the same time experimenters would like to plan according to various objectives, such as the quality of the model parameters' estimators and overall predictions, being able to protect against and effectively detect a potential lack-of-fit of the fitted model. The course presents the methodological knowledge and tools for such multi-objective planning, focusing on understanding the underlying principles and their implementation in various frameworks. The flexibility of the approach presented allows practitioners to obtain the solutions tailored to their specific experimental aims and circumstances and make informed choices regarding their experimental plan.

Introduction to Python and Statistical Machine Learning

Christian Lucero and Alex Jaimes-Sandoval, Virginia Tech

Arts District

Abstract: Statistical machine learning involves using statistical models and other algorithms that can learn from data and make predictions or decisions. Supervised learning includes regression and classification techniques and will be the main focus of this course. Topics include methods for model building, support vector machines, discriminant analysis methods, and tree-based methods. In order to properly use these methods, data cleaning and manipulation are necessary precursors along with basic exploratory data analysis (i.e. summary statistics and visualizations) which help investigators to form research questions and make decisions about which statistical machine learning methods to use.

In this full-day course, we will focus upon the complete data analysis workflow. During the first half we will introduce the Python programming language including reading in data, cleaning and manipulating data, some exploratory data analysis, and producing visualizations. In the second half of the course, we will have an overview of statistical learning methods and provide some case studies that demonstrate the use of these techniques.

AM Half-Day Short Courses: Tuesday, October 8

Half day short courses will be held on Tuesday, October 8:30AM to 12:30PM

Introduction to Reliability Data Analysis and Reintroduction to the Maximum Likelihood Method

Peng Liu, JMP Statistical Discovery LLC

Theater District

Abstract: Reliability data analysis is an important subject to the reliability engineering discipline. Reliability data analysis is a specialized subject to general statisticians. There have been enormous advances in research and development that have been invested in the area but may not be well-known to the general audience. The methodologies that have been developed and experiences that have been learned in this area, however, are extremely valuable to all scientists, engineers, and statisticians in general.

This half-day short course will serve as an introduction to reliability data analysis, by bringing interesting cases that are important to reliability engineers, and maybe interesting to the general audience including all scientists, engineers, and statisticians. Specific subjects include censored data, analysis of censored data, analysis of non-normal data, challenges of very few data, challenges to extrapolation in space and time, analysis of recurrent events, and system reliability engineering if time allows.

The instructor is a statistician and a software engineer by training. He does not analyze reliability data daily as a profession. He, however, has dedicated 17 years to developing software for analyzing reliability data. Some prominent statisticians have recognized that it takes one statistician to fully implement the software to understand the corresponding theory and method. The instructor has gained firsthand experience with analyzing reliability data through the software implementation.

This short course also wants to spend some extra time as a reintroduction to the maximum likelihood method. Based on the instructor's own background and experience, the method of maximum likelihood is under-appreciated and less understood than it needs to be. That is an unfortunate reality among scientists, engineers, and even general statisticians. The subject will be inter-weaved into the reliability data analysis part of the short course, whenever it is relevant and appropriate. The subject is so important, the instructor believes that with it one can almost treat it as a first principle for statistical analysis, while without it one might be easily overwhelmed by statistical analysis as a huge bag of tricks.

Ethics in Statistics and Data Science

Mario Davidson, Vanderbilt University

Jennifer H. Van Mullekom, Virginia Tech

Business District

Abstract: The recent revision of the American Statistical Association (ASA) Ethical Guidelines for Statistical Practice combined with increased media attention on ethical data science algorithms has prompted our profession to renew its commitment to ethics education. We have developed DEPICT, a six-phase ethical reasoning process tailored to statistics and data science. We will present overviews of ethics paradigms and the ASA Ethical Guidelines followed by a deep dive into the DEPICT process. Participants will Define ethical dilemmas; Explore possible resolutions; Plan resolutions; anticipate issues associated with Implementation; Contemplate their actions; and Transcend to incorporate key learnings to avoid future dilemmas. The course consists of interactive exercises to learn ethical frameworks, guidance, and reasoning followed by applications in complex, nuanced case studies with applications in the physical & engineering sciences and quality & productivity. Attendees will participate in facilitated small group discussions as they apply the framework, reporting key elements of their small group discussion to the larger group. Participants will develop multi-perspective views and debate the pros and cons of various resolutions based on professional guidance. This course is appropriate for students, faculty, early career professionals, and managers in any application area. Participants will learn how to apply DEPICT as well as teach or mentor ethical reasoning in statistics and data science. Pre-reading the American Statistical Association Ethical Guidelines for Statistical Practice and a sample case study will allow participants to get the most out of the course.

PM Short Courses: Tuesday, October 8

Half day short courses will be held on Tuesday, October 1:30PM to 5:30PM

Introduction to Statistical Analysis Software Development

Peng Liu, JMP Statistical Discovery LLC

Theater District

Abstract: The short course will emphasize the instructor's main experience on the development of desktop software driven by a graphical user interface (GUI). Different from software development of arbitrary desktop GUI applications, the short course will bring attention to statistical analysis software development, which has its unique challenges and opportunities. Although the course is centered around GUI driven statistical analysis software development, the instructor will cover three other important areas of software development that are important to data scientists and machine learning engineers. Two are prominent, which are software development in the R ecosystem, and that in Python ecosystem, which are related to the GUI driven statistical analysis software development, but substantially different. They have their advantages, disadvantages, challenges, and opportunities. There is a new area of statistical analysis software: web-based

software, both single user and collaborative software. The web-based software is very likely to rely on existing R and Python eco-systems, but a purely JavaScript-based approach is also feasible. The new area has its own challenges and opportunities.
The content of the short course includes:

- basic knowledge and skill set of developing a GUI driven desktop software,
- basic knowledge and skill set of developing a statistical analysis software,
- an example of developing a maximum likelihood estimation procedure,
- basic knowledge and skill set of testing a statistical analysis software,
- basic knowledge and skill set of software engineering, and
- special knowledge and skill set for developing in R, Python, and Web ecosystems.

Technical Program At-A-Glance Schedule: Wednesday, Oct 9

8:00am – 9:00am Entertainment District	**Welcome & Plenary Session** **Gerald J. Hahn Achievement Award** *Experimental Mathematics - Friend or Foe?* Fred Faltin, Virginia Tech		
	Arts District	**Music District**	**Theater District**
9:15am – 10:00am	**1A:** *AI, BI & SI: Artificial Intelligence, Biological Intelligence and Statistical Intelligence* Dennis Lin Purdue University **Moderator:** Caleb King	**1B:** *Quality by Design for the Era of Precision Medicine* Julia O'Neill Direxa Consulting LLC **Moderator:** Amanda Yoder	**1C:** *Calibration and Uncertainty Quantification for Estimating Topographic Speedup Factors with CFD Models* Adam Pintar NIST **Moderator:** Di Michelson
10:30am – 12:00pm	**2A: Robust Experimental Design** *Minimally aliased D- and A-optimal Main-effects Designs* Mohammed Saif Ismail Hameed KU Leuven *Optimal Design Under Model Uncertainty* Xletao Zhou King's College London **Moderator:** Fred Faltin	**2B: Model Estimation** *Peelle's Pertinent Puzzle and D'Agostini Bia – Estimating the Mean with Relative Systematic Uncertainty* Scott Vander Wiel LANL *Estimation and Variable Selection of Conditional Main Effects for Generalized Linear Models* Kexin Xie Virginia Tech **Moderator:** Yeng Saanchi	**2C: Technometrics Invited** *Drift vs. Shift: Decoupling Trends and Changepoint Analysis* Toryn Schafer Texas A&M *Building Trees for Probabilistic Prediction via Scoring Rules* Sara Shashaani NC State University **Moderator:** Bobby Gramacy
12:15pm – 1:45pm Entertainment District	**Luncheon** *Reflections on a Career at Eastman in Statistics* Kevin White		
2:00pm – 3:30pm	**3A: STAT Invited Session** *XGBoost Modeling for Live Use in Manufacturing* Amanda Yoder Corning *Can You Dig It? Using Machine Learning to Efficiently Audit Utility Locator Tickets Prior to Excavation to Protect Underground Utilities* Jennifer H. Van Mullekom Virginia Tech **Moderator:** Karen Hulting	**3B: Deep GPs** *Deep Gaussian Processes for Surrogate Modeling with Categorical Data* Andrew Cooper Virginia Tech *Generating Higher Resolution Sky Maps Using a Deep Gaussian Process Poisson Model* Steven D. Barnett Virginia Tech **Moderator:** Ayumi Mutoh	**3C: DOE** *Selection of Initial Points Using Latin Hypercube Sampling for Active Learning* Roelof Coetzer North-west University *Optimal Experimental Designs for Process Robustness Studies* Peter Goos KU Leuven **Moderator:** Steven Gilmour
4:00pm – 5:00pm Entertainment District	**W.J. Youden Address** *Youden's Enduring Legacy at NIST* Adam Pintar, NIST		

Welcome & Plenary Session, Gerald J. Hahn Achievement Award

Wednesday, October 9, 8:00 – 9:00am ASA Q&P Presiding

Entertainment District

Experimental Mathematics - Friend or Foe?

Fred Faltin, Virginia Tech

Abstract: In the years since World War II, the practice of statistics has been transformed in many ways. Arguably the most substantive of these has been the field's often rocky interplay with "experimental mathematics", most of whose practitioners consider themselves computer scientists, not statisticians. The rise of experimental math has undeniably contributed a great deal to current modeling methodology (consider, for example, many of the tools of Machine Learning). But it has also led to some conspicuous failures in applications. With decisions of increasingly greater importance being based on these methods, the stakes for getting it right are becoming ever higher.

From my perspectives as a student, consultant, businessperson, and academic, I'd like to share some observations and experiences regarding the evolution of the relationship between statistics and experimental mathematics. The root causes of modeling errors are often not difficult to identify, and suggest that in some quarters a basic understanding of statistical fundamentals has been lost. Identifying where naïve modelers can go wrong provides us a better understanding of what we need to address if we are to be successful in exerting a positive influence on the practice of data science, and the societally important outcomes that will ensue.

Notes

Session 1A:

Wednesday, October 9, 9:15 – 10:00am Moderator: Caleb King

Arts District

AI, BI & SI - Artificial Intelligence, Biological Intelligence, and Statistical Intelligence

Dennis K. J. Lin, Purdue University

Abstract: Artificial Intelligence (AI) is clearly one of the hottest subjects these days. Basically, AI employs a huge number of inputs(training data), super-efficient computer power/memory, and smart algorithms to perform its intelligence. In contrast, Biological Intelligence (BI) is a natural intelligence that requires very little or even no input. This talk will first discuss the fundamental issue of input (training data) for AI. After all, not-so-informative inputs (even if they are huge) will result in not-so-intelligent AI. Specifically, three issues will be discussed: (1)input bias, (2) data right vs. right data, and (3) sample vs. population. Finally, the importance of Statistical Intelligence (SI)will be introduced. SI is somehow in between AI and BI. It employs important sample data, solid theoretically proven statistical inference/models, and natural intelligence. In my view, AI will become more and more powerful in many senses, but it will never replace BI. After all, it is said that "The truth is stranger than fiction, because fiction must make sense."

Session 1B:

Wednesday, October 9, 9:15 – 10:00am Moderator: Amanda Yoder

Music District

Quality by Design for the Era of Precision Medicine

Julia O'Neill, Direxa Consulting LLC

Abstract: The pharmaceutical industry is adopting the Quality by Design paradigm after a long history of relying on Quality by Inspection. Recent accelerated approval pathways for vaccines, gene therapies, and other pioneering biotherapies have been supported by the deep connection between knowledge and statistical methods that is the foundation of QbD. The transition from conventional to modern development requires fundamental change. Many established practices in pharmaceutical statistics represent doing the "wrong thing right". Statistical strategies must be modernized to do the "right thing right" instead.

Statistical thinking provides insights to establish effective quality control strategies. Advances in analytical testing have delivered a suite of methods better able to characterize products. However the traditional Quality by Inspection focus has mandated functional and clinical testing. Functional assays may provide an intuitively direct link to safety or efficacy, but often older methods such as cell-based potency assays are handicapped by inherently lower precision and accuracy relative to newer methods. The transition to modern methods for analytical characterization can be unfamiliar and complex. Examples from recent

successful submissions will be introduced to illustrate possibilities in this new era, and to stimulate discussion about Quality by Design thinking to bridge the gap between familiar and novel control strategies.

Session 1C: Case Studies

Wednesday, October 9, 9:15 – 10:00am Moderator: Di Michelson

Theater District

Calibration and Uncertainty Quantification for Estimating Topographic Speedup Factors with CFD Models

Adam Pintar, NIST

Abstract: During extreme wind events, such as hurricanes, the wind can be significantly accelerated by hilly or mountainous terrain. This topographic influence on wind fields has been identified as a major contributor to catastrophic damage and loss of life, as evidenced by past hurricanes, such as Iniki and Maria, which severely impacted Hawaii and Puerto Rico, respectively. Buildings located in mountainous regions could experience wind speeds that are nearly double those in flat terrain. Thus, estimating this speedup is essential to determining wind loads for designing buildings to resist such extreme wind events. Wind tunnel experiments are often used to investigate topographic effects on wind speeds; however, they face challenges in accurately generating approach flows, representing complex interactions between the natural landscape and constructions at small scales, and the spatial resolution of measurements. Computational fluid dynamics (CFD) simulations provide an alternative without the scale, flow field, or resolution limitations. However, the use of CFD simulations for this purpose requires the calibration of the simulation input parameters as well as statements about uncertainty.

This work combines CFD simulations and data from wind tunnel experiments to estimate topographic speed-up factors (TSFs) for generic topographic features . The TSF is the ratio of wind speed over topography to wind speed over flat land. The typical workflow for the calibration of computer experiments is followed. Gaussian process (GP)surrogates are created for the CFD simulations that mimic the wind tunnel experiments, one with a bare floor and one with topographic features. Maximum projection and maximin designs are utilized to develop the surrogates. The simulation output is two dimensional (wind speed and turbulence) and functional (over 2D space). For the bare floor surrogate, five tuning parameters are considered, and for the topographic surrogate, a sixth is added. The tuning parameters of each surrogate are optimized to wind tunnel measurements, and additional GPs are used to assess the discrepancy between the surrogate predictions and the wind tunnel measurements. The result is a calibrated estimate of the TSF for a full 2D cross section of the wind tunnel. The uncertainty statements account for the finite sample used to construct the surrogates, the limited number of wind tunnel measurements, and noise in the wind tunnel measurements.

Notes

Session 2A: Robust Experimental Designs

Wednesday, October 9, 10:30am – 12:00pm Moderator: Fred Faltin

Arts District

Minimally aliased D- and A-optimal Main-effects Designs

Mohammed Saif Ismail Hameed, KU Leuven

Abstract: The literature on two-level D- and A-optimal designs for the main-effects model is very exhaustive for run sizes that are multiples of four. This is due to the fact that complete catalogs of D- and A-optimal designs exist for run sizes that are multiples of four. However, for run sizes that are not multiples of four, there are no such catalogs, and experimenters resort to heuristic optimization algorithms to create designs (such as coordinate-and point-exchange algorithms). This approach has multiple weaknesses. First, it requires computing time. Second, heuristic optimization algorithms often fail to return a truly optimal design. Third, even in the event the design produced is truly optimal for the main-effects model, it often exhibits substantial aliasing between the main effects and the two-factor interactions as well as among the two-factor interactions. In this presentation, we explain how to enumerate complete catalogs of D- and A-optimal main-effects designs for run sizes that are not multiples of four, and how to select the best of these designs in terms of aliasing between the main effects and the two-factor interactions and among the two-factor interactions. As a result of our work, the use of heuristic optimization can be avoided for most optimal design problems where the run size is at most 20 in the event the primary interest of the experimenter is in a main-effects model and statistical software can provide a minimally aliased D-and A-optimal design instantaneously.

Optimal Design Under Model Uncertainty

Xietao Zhou, King's College London

Abstract: Design of experiment is a useful approach when we are studying the effects of several factors on one or more responses, and it has seen wide application in industrial research and many other areas. Traditional approaches assume that the model best fitted will be fixed and some classic optimality criteria have been applied to evaluate the designs under this model. Early extensions have been made so that a few alternative models could be considered. The QB criterion has then been proposed to offer the capacity of considering hundreds of alternative models that could appear for multifactor designs. It could avoid the risk on the estimation of the factors by going to extreme belief of the model best fitted and allow experimenter's beliefs about the model reflected in the design selection process by assigning different prior probabilities to each possible model.

Recently an alternative parameterization of factorial models called Baseline parameterization has been considered in the literature. It has been argued that such parameterization arises naturally if there is a null state of each factor, and the corresponding optimal design has been raised.

In this talk I will introduce the basic framework of QB criterion and how it could be extended to the Baseline parameterization. I will then present some QB optimal designs we have found and show that they have achieved advantages in terms of traditional A optimality versus the optimal design in previous literature for various specified prior probabilities of main effects and two-factor interactions included in the best fitted model.

Session 2B: Model Estimation

Wednesday, October 9, 10:30am – 12:00pm Moderator: Yeng Saanchi

Music District

Peelle's Pertinent Puzzle and D'Agostini Bias – Estimating the Mean with Relative Systematic Uncertainty

Scott Vander Wiel, Los Alamos National Laboratory

Abstract: Scientists often represent relative, systematic uncertainty by a covariance term proportional to the outer product of measurements (yy^T). This form of error covariance produces generalized least squares estimates that are biased toward zero, an effect known in the nuclear data community as Peelle's Pertinent Puzzle and in high energy physics as D'Agostini Bias. Decades of explanations and proposed fixes have not been connected to well-established statistical methods and theory. This work fills the gap, establishing why this covariance has unacceptable properties and providing a clear explanation of the root cause–namely, overfitting. We compare properties of such estimators to the Gaussian maximum likelihood estimator and iteratively re-weighted least squares, showing that IRLS provides stable estimates with uncertainties that correctly reflect the impact of systematic errors. Peelle's Puzzle is solved!

Estimation and Variable Selection of Conditional Main Effects for Generalized Linear Models

Kexin Xie, Virginia Tech

Abstract: In numerous engineering and healthcare applications, comprehending the interaction effects among system factors is crucial. Traditional interaction effects, while powerful, often present challenges in interpretation and clarity, particularly within the context of complex systems. The concept of Conditional Main Effects (CMEs) marks a significant development aimed at addressing this challenge. In this work, we focus on modeling the data with non-continuous responses using CME within the generalized linear model framework. The proposed method considers an appropriate penalized likelihood function for model estimation, integrating CME coupling and reduction alongside an overlapping group structure to refine bi-level variable selection. An iteratively reweighted least squares procedure is used to enhance the computational efficiency. Simulation studies are conducted to examine the proposed methods' advantage on selection and estimation accuracy. A case study in public health is used to demonstrate the merits of the proposed method.

Session 2C: Technometrics Invited Session

Wednesday, October 9, 10:30am – 12:00pm Moderator: Bobby Gramacy

Theater District

Drive vs Shift: Decoupling Trends and Changepoint Analysis

Toryn Schafer, Texas A&M

Abstract: We introduce a new approach for decoupling trends (drift) and changepoints (shifts) in time series. Our locally adaptive model-based approach for robustly decoupling combines Bayesian trend filtering and machine learning based regularization. An over-parameterized Bayesian dynamic linear model (DLM) is first applied to characterize drift. Then a weighted penalized likelihood estimator is paired with the estimated DLM posterior distribution to identify shifts. We show how Bayesian DLMs specified with so-called shrinkage priors can provide smooth estimates of underlying trends in the presence of complex noise components. However, their inability to shrink exactly to zero inhibits direct changepoint detection. In contrast, penalized likelihood methods are highly effective in locating changepoints. However, they require data with simple patterns in both signal and noise. The proposed decoupling approach combines the strengths of both, i.e. the flexibility of Bayesian DLMs with the hard thresholding property of penalized likelihood estimators, to provide changepoint analysis in complex, modern settings. The proposed framework is outlier robust and can identify a variety of changes, including in mean and slope. It is also easily extended for analysis of parameter shifts in time-varying parameter models like dynamic regressions. We illustrate the flexibility and contrast the performance and robustness of our approach with several alternative methods across a wide range of simulations and application examples.

Building Trees for Probabilistic Prediction via Scoring Rules

Sara Shashaani, NC State University

Abstract: Decision trees built with data remain in widespread use for nonparametric prediction. Predicting probability distributions is preferred over point predictions when uncertainty plays a prominent role in analysis and decision-making. We study modifying a tree to produce nonparametric predictive distributions. We find the standard method for building trees may not result in good predictive distributions and propose changing the splitting criteria for trees to one based on proper scoring rules. Analysis of both simulated data and several real datasets demonstrates that using these new splitting criteria results in trees with improved predictive properties considering the entire predictive distribution.

Notes

Luncheon

Wednesday, October 9, 12:15 – 1:45pm ASQ CPID Presiding

Entertainment District

Reflection on a Career at Eastman in Statistics

Kevin White

Abstract: Historically, the FTC has featured a local speaker who provides an engaging non-technical presentation. In this spirit and being somewhat local (we'll see about the engaging part), I will begin by offering insights into Eastman, one of the state's largest employers. Eastman boasts a long-standing history of quality and statistics, and I will discuss how the applied statistics group has operated to contribute to the company's success while addressing some of the current challenges they face. Additionally, I will explore Eastman's newer AI and Machine Learning team, sharing my experiences and observations from leading both the Applied Statistics and AI/ML groups during my career. I will conclude with some experiences and advice that was particularly helpful during my career involving professional societies and the Fall Technical Conference.

Notes

Session 3A: STAT Invited Session

Wednesday, October 9, 2:00 – 3:30pm Moderator: Karen Hulting

Arts District

XGBoost Modeling for Live Use in Manufacturing

Amanda Yoder, Corning

Abstract: Detecting defects early in the process leads to cost reduction and product capability. One key metric that is used is strength testing of filters. Many techniques have been used in the past but have required much time and many resources to complete. In addition, there was an insufficient feedback loop to manufacturing where we run continuously for them to make quick decisions to fix the process. With the combination of tools including JMP, XGBoost modeling, SQL, and many others, we have built a seamless process that provides this feedback to manufacturing. We used tabular data collected from high resolution images to build the model. Using JMP we were able to easily export the model in SQL, which links us to the web application that manufacturing sees today as a bar graph which updates automatically. Shapley values are used in conjunction with the model to help provide manufacturing reasons behind the predicted result of the black-box model, which leads to process changes and prevents leakage of defects.

Can you Dig It? Using Machine Learning to Efficiently Audit Utility Locator Tickets Prior to Excavation to Protect Underground Utilities

Jennifer H. Van Mullekom, Virginia Tech

Abstract: Virginia 811 (VA811) is a not for profit company in Virginia, USA that administers the transactional aspects of utility location prior to commencing an excavation project as mandated by Virginia law. In recent years, an increasing number of excavation tickets have been entered via web users versus through the call center. These web entry tickets have a higher number of errors, as opposed to those entered by call agents. Prior to working with the Virginia Tech Statistical Applications and Innovations Group (VTSAIG), VA811 performed random audits of their tickets to ensure quality. Beginning in 2020, the VT SAIG developed two machine learning models to predict ticket quality. The most recent of these models has been integrated into VA811's quality assurance program and detects nearly twice the poor-quality tickets as compared to the random audit process used prior to late 2022. This talk details the case study in the context of the phases of Cross Industry Standard Data Mining Practice (CRISP-DM). Statistical methods include measurement systems analysis and gradient boosted machines. Features were engineered using text mining and geographical information systems data. Practical aspects of project implementation will also be discussed including data cleaning, model implementation, and model monitoring. This case study truly harmonizes quality, statistics, and data science by employing statistical thinking, feature engineering, machine learning models, and statistical programming to more accurately and efficiently audit the quality of a transactional process.

Session 3B: Deep GPs

Wednesday, October 9, 2:00 – 3:30pm Moderator: Ayumi Mutoh

Music District

Deep Gaussian Processes for Surrogate Modeling with Categorical Data

Andrew Cooper, Virginia Tech

Abstract: Many applications of experimental design produce categorical response data. Gaussian Processes (GPs) are stochastic models that provide flexible fitting of response surfaces, but must be modified to handle non-Gaussian likelihoods. Performing fully Bayesian estimation of a GP classifier requires directly sampling from a latent GP layer, and computational bottlenecks with inverting covariance matrices make posterior estimation computationally infeasible in large-data regimes. The Vecchia approximation can reduce the cost of inverting covariance matrices by inducing a sparse Cholesky decomposition. By combining this with the Elliptical Slice Sampling (ESS) algorithm for generating valid posterior samples from a latent layer, we obtain a tractable, fully Bayesian approach to fitting a global GP classification model that can handle large training sizes. We apply our methods to a Binary Black Hole (BBH) simulator example, which contains both binary and real-valued components in its response. Our method of combining fully Bayesian classification and regression models provides us full Uncertainty Quantification (UQ) estimation of BBH formation and chirp mass. Finally, we introduce an additional latent GP layer to add "deepness" to our model, which can capture non-stationary behavior in BBH formation and improve UQ estimation.

Generating Higher Resolution Sky Maps Using a Deep Gaussian Process Poisson Model

Steven D. Barnett, Virginia Tech

Abstract: The Interstellar Boundary Explorer (IBEX) satellite was launched in 2008 in an effort to learn more about the heliosphere, which sits at the boundary between our solar system and interstellar space. IBEX detects energetic neutral atoms (ENAs) emanating from the heliosphere to create sky maps describing their rate of emission. These maps are used by physicists to inform their theoretical models about the heliosphere. However, the data collected by IBEX are both noisy and irregular. Multiple tools have been developed to smooth out this data to produce higher resolution sky maps. We propose a deep Gaussian process Poisson model for the rate of energetic neutral atoms (ENAs) emanating from the heliosphere. We believe our deep Gaussian process model constitutes a more cohesive model than those developed previously. Additionally, deep Gaussian processes have shown a greater ability to learn complex surfaces, while maintaining a simpler covariance function. We hope to develop a Markov chain Monte Carlo algorithm utilizing elliptical slice sampling and the Vecchia approximation to learn the underlying latent deep Gaussian process.

Session 3C: DOE

Wednesday, October 9, 2:00 – 3:30pm Moderator: Steven Gilmour

Theater District

Selection of Initial Points Using Latin Hypercube Sampling for Active Learning

Roelof Coetzer, North-west University

Abstract: Binary classification is a common task in machine learning where the goal is to assign binary labels to observed and new observations. However, labeling large sets of data is often a time-consuming and expensive process. Active learning learns from a few data points, while selecting the most informative unlabelled samples for labeling to improve model performance. The success of active learning is dependent on the selection of the initial points to initialize the active learning process, and the selection criteria used to identify informative samples. In this paper we illustrate the use of Latin Hypercube sampling, conditioned Latin Hypercube sampling, and a modified Latin Hypercube sampling procedure for initializing active learning for the estimation of the logistic regression classifier. In addition to the usual performance measures for classification, we consider the mean squared error of the predicted posterior probability and the classifier as performance measures. The results are demonstrated using simulated data sets and some actual case studies. We show that Latin Hypercube sampling outperforms the traditional random sampling approach for various performance measures.

Optimal Experimental Designs for Process Robustness Studies

Peter Goos, KU Leuven & University of Antwerp

Abstract: In process robustness studies, experimenters are interested in comparing the responses at different locations within the normal operating ranges of the process parameters to the response at the target operating condition. Small differences in the responses imply that the manufacturing process is not affected by the expected fluctuations in the process parameters, indicating its robustness. In this presentation, we propose an optimal design criterion, named the generalized integrated variance for differences (GID) criterion, to set up experiments for robustness studies. GID-optimal designs have broad applications, particularly in pharmaceutical product development and manufacturing. We show that GID-optimal designs have better predictive performances than other commonly used designs for robustness studies, especially when the target operating condition is not located at the center of the experimental region. In some situations that we encountered, the alternative designs typically used are roughly only 50% as efficient as GID-optimal designs. We will demonstrate the advantages of tailor-made GID-optimal designs through an application to a manufacturing process robustness study of the Rotarix liquid vaccine.

Notes

W.J. Youden Address

Wednesday, October 9, 4:00 – 5:00pm — ASQ STAT Presiding

Entertainment District

Youden's Enduring Legacy at NISH

Adam Pintar, NIST

Abstract: In this address, I highlight some of Dr. Youden's vast body of work at the National Bureau of Standards (NBS) while discussing how these works continue to impact the culture, policy, and statisticians at today's National Institute of Standards and Technology (NIST). As an example, I describe Dr. Youden's contributions to the resolution of the 1953 AD-X2 controversy at NBS that severely tested NIST's core value of integrity. The resolution of this controversy is still held high as a shining illustration of NIST's integrity.

The title of my address is derived from a later article of Dr. Youden's called "Enduring Values" where he argues that systematic errors in the measurements of physical constants are best explored and quantified through designed experiments. While this ideal may not always be fully embraced, there is recognition that multiple "independent" measurement methods are crucial for rigorous uncertainty quantification, and these ideas are found in documents governing the creation of NIST standard reference materials (SRMs).

Beyond the influence that Dr. Youden had on the policy and culture of NIST, the tools that he added to the applied statistician's toolbox are timeless. Techniques for exploratory data analysis like the Youden plot and for experiment design like the Youden square remain relevant. I highlight the power and continued relevance of these tools through their use in several recent NIST projects.

Notes

Technical Program At-A-Glance Schedule: Thursday, October 10

	Arts District	Music District	Theater District
8:00am – 9:30am	**4A: CPID Invited Session** *On the Testing of Statistical Software* Ryan Lekivtez JMP *MaLT: Machine-Learning-Guided Test Case Design and Fault Localization of Complex Software Systems* Irene Ji Duke University **Moderator:** Jennifer Kensler	**4B: New Perspectives** *On Forming Control LImits for Short Run Standardized Xbar Control Charts with Varying Subgroup Sizes* Annie Dudley & Di Michelson, JMP *Collaborative Design of Controlled Experiments in the Presence of Subject Covariates* William Fisher Clemson University **Moderator:** Shane Bookholtz	**4C: JQT Invited Session** *Nonparametric Online Monitoring of Dynamic Networks* Peihua Qiu University of Florida *A Graphical Comparison of Screening Designs using Support Recovery Probabilities* Kade Young Eli Lilly & Co. **Moderator:** Fadel Megahed
10:00am – 11:30am	**5A: SPES Invited Session** *Machine Learning, Cross Validation, and DOE* Maria Weese Miami University *Autonomy vs Safety: Joint Modeling of Disengagement and Collision Events in Autonomous Vehicle Driving Study* Simin Zheng Virginia Tech **Moderator:** Michael Crotty	**5B: Screening Designs** *A Replacement for Lenth's Method for Nonorthogonal Designs* Caleb King JMP *Optimal Two-level Designs Under Model Uncertainty* Steven Gilmour King's College London **Moderator:** Xietao Zhou	**5C: QE Invited Session** *Monitoring Univariate Processes Using Control Charts: Some Practical Issues and Advice* Bill Woodall Virginia Tech *How Generative AI Models Such as ChatGPT Can Be (Mis)Used in SPC Practice, Education and Research: An Exploratory Study* Fadel Megahed Miami University **Moderator:** Peter Parker
11:45am – 1:15pm Entertainment District	**Luncheon** *Statistics is a Core Competency for Effective Collaboration and Sound Science* Madhumita (Bonnie) Ghosh-Dastidar, RAND		
1:30pm – 3:00pm	**6A: Q&P Invited Session** *Active Learning for a Recursive Non-Additive Emulator for Multi-Fidelity Computer Experiments* Junoh Heo Michigan State University *Quantitative Assessment of Machine Learning Reliability and Resilience* Lance Fiondella Darmouth **Moderator:** Ryan Lekivetz	**6B: Computer Experiments** *Quick Input-Response Space-Filling (QIRSF) Designs* Xiankui Yang University of South Florida *A Kernel-Based Approach for Modeling Gaussian Processes with Functional Information* Andrew Brown Clemson University **Moderator:** Jennifer H. Van Mullekom	**6C: DOE II** *Optimal Experimental Designs for Precision Medicine with Multi-component Treatments* Yeng Saanchi JMP *Simulation Experiment Design for Calibration via Active Learning* Özge Sürer Miami University **Moderator:** Stephanie DeHart
3:15pm – 5:15pm Entertainment District	**Reception** followed by **SPES Special Panel Session** *How to Attract and Prepare Students for Careers in Industrial Statistics* Maria Weese, Miami University; Yeng Saanchi, JMP; Peter Parker, NASA; and Kade Young, Eli Lilly & Co **Moderator:** Michael Crotty		

Session 4A: CPID Invited Session

Thursday, October 10, 8:00 – 9:30am Moderator: Jennifer Kensler

Arts District

On the Testing of Statistical Software

Ryan Lekivetz, JMP Statistical Discovery LLC

Abstract: Testing statistical software is an extremely difficult task. What is more, for many statistical packages, the developer and test engineer are one and the same, may not have formal training in software testing techniques, and may have limited time for testing. This makes it imperative that the adopted testing approach is both efficient and effective and, at the same time, it should be based on principles that are readily understood by the developer. As it turns out, the construction of test cases can be thought of as a designed experiment (DOE). This article provides a treatment of DOE principles applied to testing statistical software and includes other considerations that may be less familiar to those developing and testing statistical packages.

MaLT: Machine-Learning-Guided Test Case Design and Fault Localization of Complex Software Systems

Irene Ji, JMP Statistical Discovery LLC

Abstract: Software testing is essential for the reliable and robust development of complex software systems. Due to the complexity of these systems, testing and fault localization can be very costly. To mitigate this cost, we outline in this work a holistic machine-learning-guided test case design and fault localization (MaLT) framework, which leverages recent probabilistic machine learning methods to accelerate the testing of complex software systems. MaLT consists of three steps: (i) the construction of a suite of test cases using a covering array for initial testing, (ii) the investigation of posterior root cause probabilities via a Bayesian fault localization procedure, then (iii) the use of such Bayesian analysis to guide selection of subsequent test cases via active learning. The proposed MaLT framework can thus facilitate efficient identification and subsequent diagnosis of software faults with limited test runs. We demonstrate the effectiveness of MaLT via a numerical experiment and an application on the Traffic Alert and Collision Avoidance System (TCAS).

Session 4B: New Perspectives

Thursday, October 10, 8:00 – 9:30am Moderator: Shane Bookholtz

Music District

On Forming Control Limits for Short Run Standardized Xbar Control Charts with Varying Subgroup Sizes

Annie Dudley and Di Michelson, JMP Statistical Discovery LLC

Abstract: Short Run control charts are commonly used when manufacturing more than one product on the same production line. Developments with High Throughput and Just In Time Manufacturing methods lead to multiple products made on the same line, together with shorter intervals between products. Short Run control charts enable the practitioner to view all products in the order they were made all on the same chart, by either centering all the data by a corresponding product mean or target, or by standardizing all the data by corresponding product target and estimate of sigma.

In the absence of Short Run control chart methods, a practitioner cannot either view all the data on a meaningful control chart or apply any of the Nelson Runs Tests, as the runs are not in order.
Current methods used for computing control limits for Short Run, Standardized and Summarized (XBar/R) Control Charts have deficiencies when the subgroup size is not constant. In practice, XBar/R charts might regularly have unequal subgroup sizes. In this paper, we explore methods to resolve the constant subgroup size constraint.

We consider different published sets of recommendations for control limit calculation of Short Run standardized XBar/R control charts. For each, we study the distributional foundations of the estimators and run a simulation study to verify chart performance and present our findings for both the estimate of sigma and control limits.

Collaborative Design of Controlled Experiments in the Presence of Subject Covariates

William Fisher, Clemson University

Abstract: In some cases, researchers may run multiple, separate controlled experiments where subjects participate in more than one experiment. Due to subjects participating in multiple experiments, there is correlation among the responses across experiments. Taking account of the correlation across experiments, Zhang et al. (2024) proposed the collaborative analysis framework and demonstrated that their framework can provide more precise estimates of treatment effects than if one were to analyze the experiments separately. In this work, we consider the experimental design problem of allocating subjects to treatment or control within each of the multiple experiments when subject covariate information is available. The goal of the allocation is to provide precise estimates of treatment effects for each experiment to further improve

precision gained through collaborative analysis. Using D-optimality as our allocation criterion, we propose semi-definite programming based randomized algorithms which provide solutions to the D-optimality problem. We showcase the performance of our algorithms in a simulation study, demonstrating their effectiveness over pure randomization methods when the number of subject covariates is large.

Session 4C: Journal of Quality Technology Invited Session

Thursday, October 10, 8:00 – 9:30am Moderator: Fadel Megahed

Theater District

Nonparametric Online Monitoring of Dynamic Networks

Peihua Qiu, University of Florida

Abstract: Network sequence has been commonly used for describing the longitudinal pattern of a dynamic system. Proper online monitoring of a network sequence is thus important for detecting temporal structural changes of the system. To this end, there have been some discussions in the statistical process control (SPC) literature to first extract some features from the observed networks and then apply an SPC chart to monitor the extracted features sequentially over time. However, the features used in many existing methods are insensitive to some important network structural changes, and the control charts used cannot accommodate the complex structure of the extracted features properly. In this paper, we suggest using four specific features to describe the structure of an observed network, and their combination can reflect most network structural changes that we are interested in detecting in various applications. After the four features are extracted from the observed networks, we suggest using a multivariate nonparametric control chart to monitor the extracted features online. Numerical studies show that our proposed network monitoring method is more reliable and effective than some representative existing methods in various cases considered.

A Graphical Comparison of Screening Designs Using Support Recovery Probabilities

Kade Young, Eli Lilly & Co

Abstract: A screening experiment attempts to identify a subset of important effects using a relatively small number of experimental runs. Given the limited run size and a large number of possible effects, penalized regression is a popular tool used to analyze screening designs. In particular, an automated implementation of the Gauss-Dantzig selector has been widely recommended to compare screening design construction methods. Here, we illustrate potential reproducibility issues that arise when comparing screening designs via simulation, and recommend a graphical method, based on screening probabilities, which compares designs by evaluating them along the penalized regression solution path. This method can be implemented using simulation, or, in the case of lasso, by using exact local lasso sign recovery probabilities. Our approach circumvents the need to specify tuning parameters associated with regularization methods, leading to more reliable design comparisons.

Notes

Session 5A: SPES Invited Session

Thursday, October 10, 10:00 – 11:30am Moderator: Michael Crotty

Arts District

Machine Learning, Cross Validation, and DOE

Maria Weese, Miami University

Abstract: Recently there is interest in fitting complex models to traditional experimental designs (i.e. machine learning models and central composite designs). Machine learning models, which require out of sample data for hyperparameter tuning, are often optimized with cross validation which involves selecting subsets of training and test data. When larger, less structured, data sets are used to train machine learning models there is little concern with the structure of the subsets since they are not likely to be different from the full data. However, when the training data is collected using a structured experimental design, creating subsets using cross validation might produce training samples that do not preserve the structure of the original design. In this work we investigate the consequences of using cross validation on data collected using various types of experimental designs. We provide a literature review that illustrates the types of models fit to designed experiments and how those models are tuned. Finally, we present designs constructed using an optimality criteria to mitigate the effects of sampling in leave-one-out cross validation and k-fold cross validation procedures.

Autonomy vs Safety: Joint Modeling of Disengagement and Collision Events in Autonomous Vehicle Driving Study

Simin Zheng, Virginia Tech

Abstract: As the popularity of artificial intelligence (AI) continues to grow, AI systems have become increasingly embedded into various aspects of daily life, leading to significant transformations in industries and changing how people live. One of the typical applications of AI systems is autonomous vehicles (AVs). In AVs, the relationship between the level of autonomy and safety is an important research question to answer, which can lead to two types of recurrent events data being recorded: disengagement and collision events. This paper proposes a joint modeling approach with multivariate random effects to analyze these two types of recurrent events data. The proposed model captures the inter-correlation between the levels of autonomy and safety in AVs. We apply an expectation-maximization (EM) algorithm to obtain maximum likelihood estimates for the functional form of fixed effects, variance-covariance components, and baseline intensity functions. This proposed joint modeling approach can be useful for modeling recurrent events data with multiple event types from various applications. We analyze disengagement and collision events data from the California Department of Motor Vehicles AV testing program to demonstrate its application.

Session 5B: Screening Designs

Thursday, October 10, 10:00 – 11:30am Moderator: Xietao Zhou

Music District

A Replacement for Lenth's Method for Nonorthogonal Designs

Caleb King, JMP Statistical Discovery LLC

Abstract: A key part of implementing a screening experiment is the correct identification of the active factors. To do this requires separating the signal from the noise. Therefore, a precise estimate of σ is desirable. Lenth (1989) provided a method for estimating σ in the context of unreplicated factorial designs that has become the standard for factor screening. Currently, optimal designs are in common use as screening designs as they exist for any desired number of runs. The price for using these designs is often the loss of orthogonality of the effects. For nonorthogonal designs then, a replacement for Lenth's method is necessary. This work provides a method for analyzing saturated nonorthogonal screening designs. We show that our method performs well even when the number of active effects is up to half the number of runs.

Optimal Two-level Designs Under Model Uncertainty

Steven Gilmour, King's College London

Abstract: Two-level designs are widely used for screening experiments where the goal is to identify a few active factors which have major effects. We apply the model-robust Q_B criterion for the selection of optimal two-level designs without the usual requirements of level balance and pairwise orthogonality. We provide a coordinate exchange algorithm for the construction of Q_B-optimal designs for the first-order maximal model and second-order maximal model and demonstrate that different designs will be recommended under different prior beliefs. Additionally, we study the relationship between this new criterion and the aberration-type criteria. Some new classes of model-robust designs which respect experimenters' prior beliefs are found.

Session 5C: Quality Engineering Invited Session

Thursday, October 10, 10:00 – 11:30am Moderator: Peter Parker

Theater District

Monitoring Univariate Processes Using Control Charts: Some Practical Issues and Advice

Bill Woodall, Virginia Tech

Abstract: We provide an overview and discussion of some issues and guidelines related to monitoring univariate processes with control charts. We offer some advice to practitioners to help them set up control charts appropriately and use them most effectively. We propose a four-phase framework for control chart set-up, implementation, use, and maintenance. In addition, our recommendations may be useful for researchers in the field of statistical process monitoring. We identify some current best practices, some misconceptions, and some practical issues that rely on practitioner judgment.

How Generative AI Models such as ChatGPT can be (Mis)Used in SPC Practice, Education, and Research: An Explorative Study

Fadel Megahed, Miami University

Abstract: Generative Artificial Intelligence (AI) models such as OpenAI's ChatGPT have the potential to revolutionize Statistical Process Control (SPC) practice, learning, and research. However, these tools are in the early stages of development and can be easily misused or misunderstood. In this paper, we give an overview of the development of Generative AI. Specifically, we explore ChatGPT's ability to provide code, explain basic concepts, and create knowledge related to SPC practice, learning, and research. By investigating responses to structured prompts, we highlight the benefits and limitations of the results. Our study indicates that the current version of ChatGPT performs well for structured tasks, such as translating code from one language to another and explaining well-known concepts but struggles with more nuanced tasks, such as explaining less widely known terms and creating code from scratch. We find that using new AI tools may help practitioners, educators, and researchers to be more efficient and productive. However, in their current stages of development, some results are misleading and wrong. Overall, the use of generative AI models in SPC must be properly validated and used in conjunction with other methods to ensure accurate results.

Notes

Luncheon

Thursday, October 10, 11:45am – 1:15pm ASA SPES Presiding

Entertainment District

Statistics is a Core Competency for Effective Collaboration and Sound Science

Madhumita (Bonnie) Ghosh-Dastidar, RAND

Abstract: The American Statistical Association vision imagines a world that relies on data and statistical thinking to drive discovery and inform decisions. We know the challenges to attaining this vision are significant, so collaboration is key. As a statistician working to inform policy and decision making, I know it will take collaboration across disciplines to address society's biggest challenges—e.g., pandemic recovery, climate change, precision medicine, education reform, or criminal justice. In an era of data ubiquity and rapid analysis, statisticians and data scientists are positioned to play a central role across application areas. The gold standard for public policy is evidence-based decision making—deliberate and strategic application of real facts and research-supported principles that yields objective evidence.

Statistical science is the foundation for evidence-based decision making. As an interdisciplinary science, it has applications to every field imaginable, making statisticians uniquely qualified to lend their expertise in multiple policy domains. Effectively informing policy requires becoming involved early in the design phase; understanding the nature of the issue; and knowing how to communicate, educate, and explain. In this talk, I will provide multiple examples from health policy to highlight both valuable contributions made by statistical scientists and lessons learned – and how this model of collaboration is relevant across other fields and application areas. I will suggest areas of improvement based on lessons learned. And extrapolating from these successes, I will suggest areas for future contributions in which the stakes are very high and involving statistics will be essential.

Notes

Session 6A: Quality & Productivity Invited Session

Thursday, October 10, 1:30 – 3:00pm Moderator: Ryan Lekivetz

Arts District

Active Learning for a Recursive Non-Additive Emulator for Multi-Fidelity Computer Experiments

Junoh Heo, Michigan State University

Abstract: Computer simulations have become essential for analyzing complex systems, but high-fidelity simulations often come with significant computational costs. To tackle this challenge, multi-fidelity computer experiments have emerged as a promising approach that leverages both low-fidelity and high-fidelity simulations, enhancing both the accuracy and efficiency of the analysis. In this paper, we introduce a new and flexible statistical model, the Recursive Non-Additive (RNA) emulator, that integrates the data from multi-fidelity computer experiments. Unlike conventional multi-fidelity emulation approaches that rely on an additive auto-regressive structure, the proposed RNA emulator recursively captures the relationships between multi-fidelity data using Gaussian process priors without making the additive assumption, allowing the model to accommodate more complex data patterns. Importantly, we derive the posterior predictive mean and variance of the emulator, which can be efficiently computed in a closed-form manner, leading to significant improvements in computational efficiency. Additionally, based on this emulator, we introduce four active learning strategies that optimize the balance between accuracy and simulation costs to guide the selection of the fidelity level and input locations for the next simulation run. We demonstrate the effectiveness of the proposed approach in a suite of synthetic examples and a real-world problem. An R package RNAmf for the proposed methodology is provided on CRAN.

Quantitative Assessment of Machine Learning Reliability and Resilience

Lance Fiondella, University of Massachusetts Dartmouth

Abstract: Advances in machine learning (ML) have led to applications in safety-critical domains, including security, defense, and healthcare. These ML models are confronted with dynamically changing and actively hostile conditions characteristic of real-world applications, requiring systems incorporating ML to be reliable and resilient. Many studies propose techniques to improve the robustness of ML algorithms. However, fewer consider quantitative techniques to assess the reliability and resilience of these systems. To address this gap, this study demonstrates how to collect relevant data during the training and testing of ML suitable for the application of software reliability, with and without covariates, and resilience models and the subsequent interpretation of these analyses. The proposed approach promotes quantitative risk assessment of machine learning technologies, providing the ability to track and predict degradation and improvement in the ML model performance and assisting ML and system engineers with an objective approach to compare the relative effectiveness of alternative training and testing methods. The approach is illustrated in the context of an image recognition model, which is subjected to two generative adversarial attacks and then iteratively retrained to improve the system's performance. Our results indicate that software reliability models incorporating covariates characterized the misclassification discovery process more accurately than

models without covariates. Moreover, the resilience model based on multiple linear regression incorporating interactions between covariates tracks and predicts degradation and recovery of performance best. Thus, software reliability and resilience models offer rigorous quantitative assurance methods for ML-enabled systems and processes.

Session 6B: Computer Experiments

Thursday, October 10, 1:30 – 3:00pm Moderator: Jennifer H. Van Mullekom

Music District

Quick Input-Response Space-Filling (QIRSF) Designs

Xiankui Yang, University of South Florida

Abstract: Space-filling designs have been broadly used in computer experiments to guide efficient and informative data collection. Traditional space-filling designs primarily focus on uniformly spreading design points throughout the input space. Recent development on input-response space-filing (IRSF) designs offer additional advantages when having a nice coverage over the range of response values is also desirable for some applications. The original IRSF designs use the maximin distance criterion and a modified point exchange algorithm to balance the uniform spread of design points across the input space and response values. In this paper, we develop a new quick input-response space-filling (QIRSF) design which utilizes hierarchical clustering techniques and the minimax point to achieve desirable coverage in input and response spaces. The new method dramatically reduces the computing time by at least 20 folds while maintaining high efficiency of approximating the IRSF designs. The performance and computational efficiency of the proposed methods are demonstrated through multiple examples with different input and response dimensions and varied characteristics of response surfaces. An R Shiny App is offered to facilitate easy construction of QIRSF designs of flexible size and dimension.

A Kernel-Based Approach for Modeling Gaussian Processes with Functional Information

Andrew Brown, Clemson University

Abstract: Gaussian processes are commonly used tools for modeling continuous processes in machine learning and statistics. This is partly due to the fact that one may employ a Gaussian process as an interpolator for a finite set of known points, which can then be used for prediction and straight forward uncertainty quantification at other locations. However, it is not always the case that the available information is in the form of a finite collection of points. For example, boundary value problems contain information on the boundary of a domain, which is an uncountable collection of points that cannot be incorporated into typical Gaussian process techniques. In this paper, we propose and construct Gaussian processes that unify, via reproducing kernel Hilbert space, the typical finite case with the case of having uncountable information by exploiting the equivalence of conditional expectation and orthogonal

projections. We show the existence of the proposed Gaussian process and that it is the limit of a conventional Gaussian process conditioned on an increasing but finite number of points. We illustrate the applicability via numerical examples and proof-of-concept.

Session 6C: DOE II

Thursday, October 10, 1:30 – 3:00pm Moderator: Stephanie DeHart

Theater District

Optimal Experimental Designs for Precision Medicine with Multi-component Treatments

Yeng Saanchi, JMP Statistical Discovery LLC

Abstract: In recent years, many medical and behavioral interventions have evolved to combine multiple therapeutic options. We refer to treatments that consist of a combination of therapeutic options as multi-component treatments. Cancer treatment is one area in which the use of such treatments is increasingly becoming common. Considering several therapeutic options in a clinical trial can result in a potentially large number of unique treatment combinations at each decision point. Ignoring the overlap among components and viewing each combination as unique in evaluating the treatment combinations is inefficient. We propose a strategy that leverages classical optimal design methods to assign combination treatments to maximize power for primary and secondary analyses of interest. Our method uses the c and D optimality criteria in constructing a composite objective function to be minimized to obtain the optimal allocation of multi-component treatments to promote the patient's long-term well-being.

Simulation Experiment Design for Calibration via Active Learning

Özge Sürer, Miami University

Abstract: Simulation models often have parameters as input and return outputs to understand the behavior of complex systems. Calibration is the process of estimating the values of the parameters in a simulation model in light of observed data from the system that is being simulated. When simulation models are expensive, emulators are built with simulation data as a computationally efficient approximation of an expensive model. An emulator then can be used to predict model outputs, instead of repeatedly running an expensive simulation model during the calibration process. Sequential design with an intelligent selection criterion can guide the process of collecting simulation data to build an emulator, making the calibration process more efficient and effective. This presentation focuses on two novel criteria for sequentially acquiring new simulation data in an active learning setting by considering uncertainties on the posterior density of parameters. Analysis of several simulation experiments and real-data simulation experiments from epidemiology demonstrates that proposed approaches result in improved posterior and field predictions.

Notes

SPES Wine & Cheese Reception and Special Session

Thursday, October 10, 3:15 – 5:15pm Moderator: Michael Crotty

Entertainment District

How to Attract and Prepare Students for Careers in Industrial Statistics

Panelists: Maria Weese, Miami University
Yeng Saanchi, JMP Statistical Discovery LLC
Peter Parker, NASA
Kade Young, Eli Lilly & Co

Abstract: In a world of big data, cloud analytics, and artificial intelligence, graduate students in statistics and data science programs might not readily think of careers in industrial statistics. This panel will explore strategies for recruiting students into careers in industrial statistics and preparing them for those careers. On the recruiting front, we will discuss recruitment locations, effective advertising strategies, and the types of problems to highlight. On the preparation front, we will explore the essential technical and non-technical skills that are needed for success and whether current programs are adequately preparing new hires for industry roles. We hope to generate a lively discussion on these issues and more.

Notes

Biographies

Steven D. Barnett, Virginia Tech sdbarnett@vt.edu

Steven Barnett is a fifth-year graduate student at Virginia Tech pursuing a Ph.D. in Statistics. He received his M.S. in Statistics from Virginia Tech in 2022 and a B.S in Computer Science from Brigham Young University in 2016. His research focuses on Gaussian process regression, surrogate modeling, and computer model calibration, with applications in the physical and engineering sciences. Steven currently works as a Graduate Research Assistant at Los Alamos National Laboratory.

Andrew Brown, Clemson University ab7clemson.edu

Andrew Brown earned his MS and PhD in Statistics from the University of Georgia under the direction of Nicole Lazar and Gauri Datta. He subsequently took a faculty position in the School of Mathematical and Statistical Sciences at Clemson University where he is now an Associate Professor. He maintains research interests in functional and structural neuroimaging data analysis from his dissertation work, and has since expanded into uncertainty quantification, computer experiments, Bayesian computation, and inverse problems. This is in addition to collaborative work he has done in areas such as veterinary medicine, engineering design, and materials science. His research has been supported by the National Science Foundation and the Department of Education. He has served as an elected officer with the Industrial Statistics section of the International Society for Bayesian Analysis, the Uncertainty Quantification Interest Group of the American Statistical Association, and is past president of the South Carolina chapter of the ASA. He has had visiting positions at the Statistical and Applied Mathematical Sciences Institute, the School of Industrial and Systems Engineering at Georgia Tech, and at Los Alamos National Laboratory.

Roelof Coetzer, North-west University roelof.coetzer@nwu.ac.za

Roelof LJ Coetzer, PhD, Pr.Sci.Nat. joined the North-West University on 1 June 2021. He is currently a professor in Business Development, Faculty of Engineering, and active member of the Research Focus Area for Pure and Applied Analytics, Faculty of Natural and Agricultural Sciences. He obtained his PhD in Mathematical Statistics from the University of the Witwatersrand in 2004. His research interests include the design of experiments, statistical modeling and optimization, statistical learning and multivariate process monitoring. He is co-author of 53 peer-reviewed articles in national and international scientific journals and conference proceedings, and has one patent in coal preparation related to coal particle size distribution.

Dr. Coetzer has been successful in driving Data Science projects and Big Data Analytics solutions in Industry. He has extensive experience in data-driven process modeling and optimization, leading multidisciplinary highly technical teams and developing technology packages. Specifically, he initiated, developed, and led technical teams in implementing complete online AI, machine learning and diagnostic software platforms in Industry. The Sasol MSPEMTM Technology Package is a complete online process monitoring, modeling, and diagnostic platform, which is used daily for decision making, control and predictive analytics across the whole coal value chain. He has been at Sasol for about 23 years, and in total, has 30 years of experience in Industry. Other career achievements include: Council Member of the South African Council for Natural Scientific Professions (SACNASP) since 1 May 2015; Chair of the SACNASP Professional Accreditation Committee for Statistical Sciences; President of the South African Statistical Association (SASA) in 2013. Extraordinary Associate Professor in the School of Industrial Engineering, North-West University, since 2019; Affiliated Associate Professor in the Department of Mathematical Statistics and Actuarial Science, University of the Free State, from January 2013 – December 2021.

Andrew Cooper, Virginia Tech ahcooper@vt.edu

Andrew Cooper is a 5th-year PhD candidate in Virginia Tech's Department of Statistics. He received his bachelors and masters degrees in Statistical Science from Duke University. His research areas include computer experiments and surrogate modeling, as well as Bayesian methodology.

Mario Davidson, Vanderbilt University mario.davidson@vumc.org

Dr. Mario Davidson is an Associate professor in biostatistics at Vanderbilt University School of Medicine in Nashville, TN. Holding the position of Associate Vice-Chair of Equity, Diversity, and Inclusion within the Department of Biostatistics, he plays a pivotal role in fostering a culture of inclusivity. Dr. Davidson earned his Ph.D. in Statistics Education from The Ohio State University.

As the lead biostatistician for medical students and educators at Vanderbilt, Dr. Davidson directs the second-year research course known as PLAN. This innovative program employs weekly topics to guide students in comprehending research methodologies and crafting protocols. Additionally, he has facilitated numerous case-based learning courses for medical students and oversees the Classroom Peer Reviews, where he administers training for faculty peer reviewers.

Dr. Davidson is credited with developing the department's integral course, Statistical Collaboration in Health Sciences. This course emphasizes communication, professionalism, and ethics, showcasing his expertise in statistical collaboration, education, and ethical considerations. Actively engaged in service committees at Vanderbilt, he contributes to discussions on diversity, equity, inclusion, assessment, bias, and admissions.

A member of the Academy of Excellence in Education, Dr. Davidson has served on its board, reflecting his commitment to advancing educational excellence.

Annie Dudley, JMP Statistical Discovery LLC Annie.Dudley@jmp.com

Annie Dudley currently works as a Senior Research Statistical Developer on the JMP product at JMP Inc., and has worked in the JMP group at SAS for 30 years. Her M.S. is in statistics from Virginia Tech. She currently develops control charts in JMP, including the Control Chart Builder, EWMA and CUSUM charts.

Olga Egorova, King's College London o.egorova@kcl.ac.uk

Olga Egorova is a Research Associate at King's College London and has been working on methodological developments in optimal experimental design and providing application-tailored design and analysis support for experimenters across academia and industry. She has more than 5 years of research experience, working on multi-disciplinary projects in different collaborative settings, along with teaching various statistical methods. Olga specializes in multi-objective experimental design, and most recently in screening designs and sequential planning of experiments – building comprehensive statistical methodologies for general experimentation as well as particular applications.

Fred Faltin, Virginia Tech ffaltin@vt.edu

Over his 40+ year career, **Dr. Fred Faltin** has made significant contributions as a researcher, consultant, author, and educator. He has successfully applied statistical quality and productivity methods across many problem domains, and through his publications and presentations, has promoted their use by others. Fred is best known for his work in process control and statistical process monitoring, design of experiments, and as a corporate research manager and international consultant. He was elected a Fellow of ASA in 2002.

Fred was deeply involved in quality and productivity improvement across many of General Electric's businesses and founded the data science laboratory at GE Research. After his career at GE, Fred co-founded a consulting company which supported organizations around the world in domains ranging from manufacturing to financial services, healthcare, and corporate real estate.

Fred has successfully promoted quality and productivity techniques through the publication of technical and non-technical papers and articles in refereed journals, magazines, and book chapters. He was co-Editor-in-Chief of three books in applications of statistics, including the highly cited Encyclopedia of Statistics in Quality and Reliability. Through the many professionals he has trained and mentored in courses, tutorials and seminars, his work has delivered incalculable value to individuals' careers and to the product and process design and improvement of their employers. Fred's ability to translate sophisticated statistical concepts into real-world applications that non-statistical operators and managers could understand and

apply has been a key to his impact. Methods that Fred developed and taught are now routinely used by practitioners across the globe as part of their everyday jobs.

Fred currently serves as Professor of Practice in the Department of Statistics at Virginia Tech, where he teaches the theory and practice of Data Science in VT's Computational Modeling and Data Analytics curriculum. In this role, he is promoting awareness and use of statistics for organizational quality and productivity improvement to the next generation of professional data scientists. Fred is a natural teacher whose wealth of personal and business experience greatly benefits his students.

Fred Faltin has demonstrated outstanding and sustained achievement and leadership in developing, promoting, and successfully improving the quality and productivity of products, services, and organizational performance using statistical concepts. He is an exceptionally deserving recipient of the prestigious ASA Gerald J. Hahn Q&P Achievement Award.

Lance Fiondella, University of Massachusetts Dartmouth lfiondella@umassd.edu

Dr. Lance Fiondella is an Associate professor of Electrical and Computer Engineering and the Director of the University of Massachusetts Dartmouth Cybersecurity Center, an NSA/DHS National Center of Academic Excellence in Cyber Research (CAE-R).

William Fisher, Clemson University wsfishe@clemson.edu

William Fisher is a PhD candidate at Clemson University's School of Mathematical and Statistical Sciences. His research interests are broadly in experimental design, with a particular focus on discrete choice experiments. He is currently interning with JMP as a research statistician software developer.

Madhumita (Bonnie) Ghosh-Dastidar, RAND bonnieg@rand.org

Madhumita (Bonnie) Ghosh-Dastidar is head of the Statistics Group and a senior statistician and data scientist at RAND, a nonprofit organization that improves policy and decision making through research and analysis. Her policy applications are focused on health and social justice. Specific examples include evaluation of neighborhood-level intervention to improve diet, health; cluster randomized trials to improve outcomes among HIV positive persons; and measurement or surveillance in health disparity. Her statistical expertise includes design, sampling, measurement, missing data, evaluations, and longitudinal/multilevel modeling. Ghosh-Dastidar's major studies include assessment of sexual assault and harassment in the U.S. military and the first Singapore mental health study to assess national prevalence. She is the 2024 President of the American Statistical Association (ASA), the largest professional organization for statisticians and data scientists. She is an elected fellow of the ASA. The ASA Committee on Women in Statistics recognized her work in statistics and data science by naming her "Top 20 Woman in Statistics and Data Science" (March 2019. Ghosh-Dastidar received her Ph.D. and M.A. in statistics from Penn State, University Park, PA; she also holds a dual Bachelor's in mathematics and computer science from Albright College, Reading, PA.

Steven Gilmour, King's College London steven.gilmour@kcl.ac.uk

Steven Gilmour is Professor of Statistics at King's College London, where he also served as Chair of the Department of Mathematics from 2019-24. He has more than 35 years of research experience in the theory, methodology and applications of the design and analysis of experiments. Steve is co-author of the book Statistical Principles for the Design of Experiments, as well as more than 100 research articles in both Statistics journals and diverse application areas. He has taught design and analysis of experiments to many different audiences, as well as teaching most other areas of statistics. Steve's current research includes various aspects of multi-objective optimal design, model-robust design, and design in the context of big data. He is known in both his teaching and research for always trying to build new ideas on the strong classical foundations of the design of experiments.

Peter Goos, KU Leuven & University of Antwerp peter.goos@kuleuven.be

Peter Goos is a full professor at the Faculty of Bio-Science Engineering of KU Leuven, and at the Faculty of Business and Economics of the University of Antwerp, where he teaches various introductory and advanced courses on statistics and probability. His main research area is the statistical design and analysis of experiments. Besides numerous influential articles in various kinds of scientific journals, he published the books The Optimal Design of Blocked and Split-Plot Experiments, Optimal Experimental Design: A Case-Study Approach, Statistics with JMP: Graphs, Descriptive Statistics and Probability and Statistics with JMP: Hypothesis Tests, ANOVA and Regression. For his work, Peter Goos has received four Shewell Awards, two Lloyd S. Nelson Awards, a Brumbaugh Award and the Youden Award of the American Society for Quality, the Ziegel Award and the Statistics in Chemistry Award from the American Statistical Association, and the Young Statistician Award of the European Network for Business and Industrial Statistics (ENBIS). Peter Goos is known for this ability to introduce new design of experiments concepts in an accessible fashion to non-academics, and he is a co-founder of the statistical software company Effex®.

Junoh Heo, Michigan State University heojunoh@msu.edu

Junoh Heo is a Ph.D. student in the Department of Statistics and Probability at Michigan State University, and currently advised by Dr. Chih-Li Sung. His research interests include computer experiment, uncertainty quantification, multi-fidelity simulation in the field of industrial/engineering statistics.

Mohammed Saif Ismail Hameed, KU Leuven mohammedsaif.ismailhameed@kuleuven.be

Mohammed Saif Ismail Hameed is a postdoctoral researcher at KU Leuven's MeBioS research group (Mechatronics, Biostatistics and Sensors). He obtained a PhD for his work on design and analysis of experiments under the supervision of Prof. Peter Goos at KU Leuven, Belgium. He holds a masters degree in Manufacturing Engineering from the University of Michigan, USA, and a bachelors degree in Mechanical Engineering from Ramaiah Institute of Technology, India. He is a certified Black Belt from the University of Michigan in both Lean Six Sigma and Design for Six Sigma.

Alex Jaimes-Sandoval, Virginia Tech alexjaimes9@vt.edu

Alex Jaimes-Sandoval received his undergraduate degree from Virginia Tech in Computational Modeling and Data Analytics in 2024 with minors in Mathematics, Statistics, and Computer Science. He served as the head of the CMDA Computing Consultants, an in-major tutoring group that provides help with Python, R, Java, and other languages that are regularly used by students in CMDA courses. He has received multiple awards including the Alice-Luther Hamlett Scholarship and the CMDA Outstanding Senior Award. Alex has also won the grand prize in multiple American Statistical Association (ASA) DataFest competitions and has interned at JP Morgan Chase as an AI Analyst. He is currently a graduate student in Computer Science and will be focusing on Machine Learning and Numerical Analysis.

Irene Ji, JMP Statistical Discovery LLC Irene.Ji@jmp.com

Yi (Irene) Ji obtained her Ph.D. degree in statistics from Duke University in May 2024. She is now working as a Research Statistician Developer at JMP. Her research interests include design of experiments, uncertainty quantification, and machine learning methods in physics and engineering.

Caleb King, JMP Statistical Discovery LLC Caleb.King@jmp.com

Dr. Caleb King is a Senior Research Statistician Developer in the DOE& Reliability Group at JMP Statistical Discovery, LLC. His areas of research include design of experiments, accelerated testing, and reliability analysis among others. He also consults on a variety of projects, including those with WildTrack, a non-profit organization focused on animal tracking and conservation through footprint imaging. Dr. King received his PhD in statistics from Virginia Tech. Prior to joining JMP, he worked for 3 years as a senior statistician at Sandia National Laboratories.

Ryan Lekivetz, JMP Statistical Discovery LLC Ryan.Lekivetz@jmp.com

Ryan Lekivetz is a Senior Manager of Advanced Analytics R&D at JMP, heading the Design of Experiments (DOE) and Reliability Development team. Ryan earned his doctorate in statistics from Simon Fraser University in Burnaby, British Columbia. He has published papers on DOE topics in peer-reviewed journals and holds many patents that he shares with his team members.

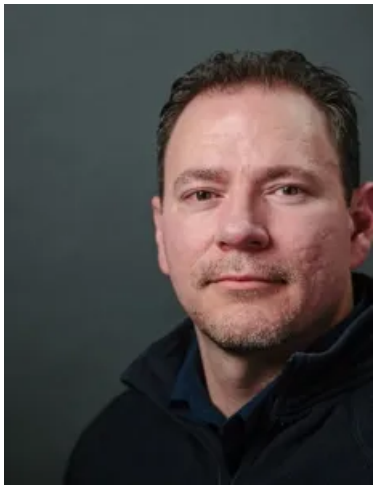
Dennis K. J. Lin, Purdue University djklin@purdue.edu

Dr. Dennis K. J. Lin is a Distinguished Professor of Statistics at Purdue University. He served as the Department Head during 2020-2022. Prior to this current job, he was a University Distinguished Professor of Supply Chain Management and Statistics at Penn State, where he worked for 25 years. His research interests are data quality, industrial statistics, statistical inference, and data science. He has published nearly 300 SCI/SSCI papers in a wide variety of journals. He currently serves or has served as an associate editor for more than 10 professional journals and was a co-editor for Applied Stochastic Models for Business and Industry. Dr. Lin is an elected fellow of ASA, IMS, ASQ, & RSS, an elected member of ISI, and a lifetime member of ICSA. He is an honorary chair professor for various universities, including a Chang-Jiang Scholar at Renmin University of China, Fudan University, and National Taiwan Normal University. His recent awards include, the Youden Address (ASQ, 2010), the Shewell Award (ASQ, 2010), the Don Owen Award (ASA, 2011), the Loutit Address (SSC, 2011), the Hunter Award (ASQ, 2014), the Shewhart Medal (ASQ, 2015), and the SPES Award (ASA, 2016). He was the Deming Lecturer Award at 2020 JSM. His most recent award is "The 2022 Distinguished Alumni Award" (National Tsing Hua University, Taiwan).

Peng Liu, JMP Statistical Discovery LLC peng.liu@jmp.com

Peng Liu is a Principal Research Statistician Developer at JMP Statistical Discovery LLC. He holds a Ph.D. in statistics from NCSU. He has been working at JMP since 2007. He specializes in computational statistics, software engineering, reliability data analysis, reliability engineering, time series analysis, and time series forecasting. He is responsible for developing and maintaining all JMP platforms in the above areas. He has a broad interest in statistical analysis research and software product development.

Christian Lucero, Virginia Tech chlucero@vt.edu



Christian Lucero is a Collegiate Assistant Professor of Statistics in the College of Science at Virginia Polytechnic and State University (Virginia Tech). He earned his PhD in Mathematical and Computer Sciences from the Colorado School of Mines in 2013 where his research focused upon optimal experimental designs for ill-posed inverse problems. His general interests include statistical computing, statistical machine learning, inverse problems, uncertainty quantification, and education in statistics and data science. While his home department is Statistics, he is a core faculty member teaching in the Computational Modeling and Data Analytics (CMDA) program. He regularly teaches courses on statistical machine learning and statistical computing at both the undergraduate and graduate levels. He advocates the use of experiential learning and prefers to teach using real-world datasets to answer questions that stakeholders care about as much as possible. One of his goals is to give students many opportunities to work with real data as much as possible, and as early as possible, and he regularly achieves this through hosting data competitions each semester at Virginia Tech. The short course he will be presenting is based upon the most popular CMDA course that is regularly taken by hundreds of Virginia Tech students each year.

Fadel Megahed, Miami University fmegahed@miamioh.edu

Dr. Fadel M. Megahed is a Miami University Faculty Scholar and an Information Systems & Analytics Professor. He received his Ph.D. and M.S. in Industrial and Systems Engineering from Virginia Tech and a B.S. in Mechanical Engineering from the American University in Cairo. His main research streams include applied artificial intelligence (AI) and statistical surveillance, with applications in manufacturing, public health, and occupational health & safety. His work in these areas has been funded by Aflac, The American Society for Safety Engineers (ASSE) Foundation, GE Research, Gore, The National Institute for Occupational Safety and Health (NIOSH), and The National Science Foundation (NSF). Dr. Megahed has 56 peer-reviewed journal papers, four invited editorials/discussions, and 12 conference proceedings. His research findings and views have been covered in over 50 media articles.

Dr. Megahed is the Editor of the Case Study Section for the Journal of Quality Technology. He has received numerous research and teaching awards. He was named a University Faculty Scholar in 2023. He was the Endres Associate Professor (2022-2024) and an FSB Research Fellow (2023-2025). He also received the Student Recognition of Teaching Excellence Award (Fall 2020) and 27 faculty commendations for impacting students' learning and development (2018-2023). Before joining Miami University, he received the Career Development Award from the NIOSH Deep South Center for Occupational Health and Safety (2013) and the Mary G. and Joseph Natrella Scholarship from the Quality and Productivity Section of the American Statistical Association (2012).

Di Michelson, JMP Statistical Discovery LLC di.michelson@jmp.com

Di Michelson received a Ph.D. in Statistics in 1994 from Texas A&M University. She started her career as an industrial statistician in the semiconductor industry and now has the job of her dreams, teaching engineers and scientists how to get information out of their data using her favorite software. Her research interests include statistical process control and design of experiments, especially when the data are auto-correlated and factors are random.

Julia O'Neill, Direxa Consulting LLC Julia@DirexaConsulting.com

Since 2005, **Julia O'Neill** has supported approval of multiple accelerated products including vaccine, gene therapy, microbiome, and regenerative medicines. From 2020 to 2023 she was a Distinguished Fellow on the Moderna Technical Development Leadership Team, where she built and led a new department, CMC Modeling and Statistics, while a member of the Spikevax vaccine technical development team. She was named a Fellow of the American Society for Quality in 2020 in recognition of her passion for connecting people and data across disciplines to accelerate delivery of life-changing medicines to patients at commercial quality scale. She is the 2023 recipient of the ASA Gerald J. Hahn Quality & Productivity Achievement Award, and an Associate Director of CASSS.

From 2005 to 2015 O'Neill worked at Merck (MSD) as Senior Scientific Fellow – Statistics in Regulatory & Analytical Sciences; and Director in Global Technical Operations, with a primary focus on vaccines and biologics. Her experience includes development of specifications; stability analysis; development, qualification and validation of analytical methods; process development and qualification; control strategy authoring; comparability assessment; design and implementation of Continued Process Verification programs; and expert Design of Experiments support for a wide range of development programs in vaccines, biologics, gene therapies, small molecules, and other products derived from biological materials. She developed and delivered an invited full-day course on statistical applications in development and manufacturing for the FDA Staff College in 2016. She is a Six Sigma Master Black Belt who built and directed multiple teams which successfully resolved complex investigations, drove sustained improvements, and represented solutions to regulators. Her mentoring has provided encouragement for five colleagues to complete graduate degrees in statistics.

She has over 30 years of experience bridging statistics and chemical engineering in the pharmaceutical and chemical industries. Her education synthesizes statistics and engineering, with an MS in Statistics from the University of Wisconsin, and a BS in Chemical Engineering from the University of Maine.

Peter Parker, NASA peter.a.parker@nasa.gov

Dr. Peter Parker is Team Lead for Advanced Measurement Systems at the National Aeronautics and Space Administration's Langley Research Center in Hampton, Virginia. He serves as an Agency-wide statistical expert across NASA's Aeronautics, Exploration, and Science mission directorates to infuse statistical thinking and methods. His expertise is in collaboratively integrating engineering objectives with statistical and measurement sciences to enable rigorous decision-making in aerospace research and development. After eight years in private industry, Dr. Parker joined NASA Langley Research Center in 1997.

He holds a B.S. in Engineering, a M.S. in Applied Physics and Computer Science, and a M.S. and Ph.D. in Statistics from Virginia Tech. He is a licensed Professional Engineer in the Commonwealth of Virginia. Dr. Parker is an Associate Fellow of the American Institute of Aeronautics and Astronautics, and Senior Member of the American Society for Quality (ASQ) and American Statistical Association. Dr. Parker is Editor Emeritus of Quality Engineering and past-Chair of ASQ's Publication Management Board. He currently serves as past-Chair of the International Statistical Engineering Association.

Adam Pintar, NIST adam.pintar@nist.gov

Dr. Adam Pintar earned a Ph.D. in Statistics from Iowa State University in 2010, and has been a Mathematical Statistician with NIST's Statistical Engineering Division since October of the same year. His work is primarily collaborative research in many different areas, e.g., Engineering, Chemistry, and more recently Nanoscience. He also teaches in the Georgetown environmental metrology and policy program. Adam is a Past Chair of the Statistics Division of the American Society for Quality (ASQ), and served as the General Conference Chair of the 2019 FTC held in Gaithersburg, MD. He currently serves on the editorial board for the journals Transactions on Mathematical Software and Statistical Analysis and Data Mining. and he is a member of the American Statistical Association and a senior member of the ASQ.

Peihua Qiu, University of Florida pqui@ufl.edu

Peihua Qiu is Dean's Professor and Founding Chair of the Department of Biostatistics at the University of Florida. He received his PhD in statistics from the Department of Statistics at the University of Wisconsin at Madison in 1996. He has made substantial contributions in the research areas of jump regression analysis, image processing, statistical process control, survival analysis, dynamic disease screening, and spatio-temporal disease surveillance. So far, he has published three books and over 170 research papers in refereed journals in these areas. He is an elected fellow of the American Association for the Advancement of Science (AAAS), an elected fellow of the American Statistical Association (ASA), an elected fellow of the American Society for Quality (ASQ), an elected fellow of the Institute of Mathematical Statistics (IMS), and an elected member of the International Statistical Institute (ISI). He served as associate editor for a number of top statistical journals, including Journal of the American Statistical Association, Biometrics, and Technometrics. He was the Editor-in-Chief of the flagship statistical journal Technometrics during 2014-2016, and the recipient of the Shewhart Medal from ASQ in 2024.

Yeng Saanchi, JMP Statistical Discovery LLC Yeng.Saanchi@jmp.com

Yeng Saanchi is an Analytics Software Tester at JMP Statistical Discovery, LLC. She received her M.S. and Ph.D. in Statistics from the University of Michigan, Ann Arbor, and North Carolina State University, respectively. Her research is primarily in the areas of stochastic optimization and applications of optimal experimental designs to precision medicine, but she's always open to exploring new areas whenever the opportunity presents itself.

Sara Shashaani, NC State University sshasha2@ncsu.edu

Sara Shashaani is an Assistant Professor in Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. Her research interests lie in the intersection of simulation, analytics, and stochastic optimization with a focus on continuous zeroth-order problems as well as high-dimensional binary decision spaces. Quantifying bias in the model outputs with nonparametric approaches is another direction of her research, seeking to broaden the advantage of Monte Carlo methodology for machine learning. She is an active member and elected treasurer of the INFORMS Simulation Society and the co-creator of SimOpt: the open-source library and benchmarking platform for simulation optimization. Her research has contributed to renewable energy, climate adaptation, advanced manufacturing, and public health application areas.

Toryn Schafer, Texas A&M tschafer@tamu.edu

Toryn Schafer is an Assistant Professor and 2024 ConocoPhillips Data Science Faculty Fellow in the Department of Statistics at Texas A&M University. Dr. Schafer's research interests span many topics, but are primarily related to spatio-temporal modeling, Bayesian statistics, and alternative learning frameworks such as machine learning, deep learning, and reinforcement learning. She is highly motivated by ecological and environmental applications.

Özge Sürer, Miami University surero@miamioh.edu

Özge Sürer is an Assistant Professor of Business Analytics at Miami University. Before joining Miami University, she was a postdoctoral research fellow at the Northwestern Argonne Institute of Science and Engineering, where she specialized in developing novel techniques utilizing Bayesian uncertainty quantification and computational statistics. Dr. Sürer earned her Ph.D. in Industrial Engineering and Management Sciences from Northwestern University in December 2020, with her dissertation centered on creating interpretable predictive models for large data sets.

Jennifer H. Van Mullekom, Virginia Tech vanmuljh@vt.edu

After a 20-year career in industry, **Dr. Jennifer Van Mullekom** joined Virginia Tech in Fall 2016 as the Director of the Statistical Applications and Innovations Group (SAIG) where she is a Professor of Practice in the Department of Statistics. In addition to directing SAIG, she teaches collaboration skills and design of experiments to graduate students while serving as an active member of the global statistical practice community.

Formerly, she was a Senior Consulting Statistician and Certified Six Sigma Master Black Belt in DuPont's Applied Statistics Group, supporting the DuPont Protection Technologies business. At DuPont, she provided statistical leadership to the Tyvek® Medical Packaging Transition Project in the areas of product development, quality, commercialization, and regulatory. Her contributions to this project earned her a DuPont Engineering Excellence Award, one of the company's highest honors. She continues to collaborate with DuPont on various material science projects as permitted by Virginia Tech's external consulting policy.

Dr. Van Mullekom is active in professional societies, holding leadership roles in the American Statistical Association (ASA) and the American Society for Quality (ASQ). She is an inventor on two US Patents and has also worked at Lubrizol and Capital One. Dr. Van Mullekom is a regular participant at the Conference on Statistical Practice on topics such as communication, collaboration, leadership, and ethics. In 2024, she was honored with the American Statistical Association's Section on Statistical Consulting Mentoring Award for her role mentoring junior employees, colleagues, and students. She holds an MS and PhD in Statistics from Virginia Tech and a BS in Mathematics and a BS ED in Mathematics Education from Concord University.

Maria Weese, Miami University weeseml@miamioh.edu

Maria Weese is an Associate Professor of Business Analytics at Miami University. She holds a bachelor's degree in Chemical Engineering from Virginia Tech and worked as a Process Improvement Engineer at Celanese Acetate before returning to graduate school to pursue a degree in Statistics at the University of Tennessee. Her research interests focus on industrial analytics, particularly in the areas of screening experiments, experimental analysis, and process monitoring. She is an Associate Editor for *Technometrics* and serves on the editorial review boards of *Quality Engineering* and *Quality and Reliability Engineering International*.

Scott Vander Wiel, Los Alamos National Laboratory scottv@lanl.gov

Scott Vander Wiel is a fellow of the American Statistical Association, conducting statistics research at Los Alamos National Laboratory since 2005. He collaborates with engineers and scientists to analyze data and develop statistical methods for a rich set of problem domains such as nuclear forensics, meso-scale behavior of materials, radio astronomy, malware detection, power grid operation, rare event estimation, anomaly detection, and numerical optimization. Scott holds patents on methods for network traffic modeling and incremental quantile estimation. At LANL he is the Uncertainty Quantification project lead for the Advanced Certification Campaign. Scott earned a Ph.D. in Statistics from Iowa State University in 1991 and was a researcher at Bell Laboratories until 2005.

Kevin White timothy.kevin.white@gmail.com

Kevin White is a recent retiree from Eastman where he worked as a statistician including leading a team of statisticians who provided experimental design and other statistical modeling solutions in support of the top company growth programs. He also led the Eastman AI and machine learning technology team that delivered advanced analytics solutions across a range of business problems. He received his bachelor's and master's degrees in statistics from the University of Tennessee. He is an ASQ Fellow, ASQ Certified Quality Engineer, a past chair of the ASQ Chemical and Process Industries Division, a past chair of the Fall Technical Conference, and a past steering committee member of the Fall Technical Conference.

Bill Woodall, Virginia Tech bwoodall@vt.edu

William H. Woodall is Professor Emeritus of Statistics at Virginia Tech. He is a former Editor of the *Journal of Quality Technology* (2001–2003) and Associate Editor of *Technometrics* (1987–1995). He will be the next editor of *Quality Engineering*. He has published over 180 refereed journal papers, most on various aspects of process monitoring. He is the recipient of the ASQ Shewhart Medal (2002), ENBIS Box Medal (2012), William G. Hunter Award (2019), Jack Youden Prize (1995, 2003), ASQ Brumbaugh Award (2000, 2006), Ellis Ott Foundation Award (1987), Soren Bisgaard Award (2012), Lloyd S. Nelson Award (2014), and an IIE Transactions on Quality and Reliability Engineering best paper award (1997). He is a Fellow of the American Statistical Association, a Fellow of the American Society for Quality, and an elected member of the International Statistical Institute. Current research interests include sampling issues related to process monitoring, risk-adjusted outcome monitoring in healthcare applications, and social network monitoring.

Kexin Xie, Virginia Tech kexin@vt.edu

Kexin Xie is a fourth-year PhD student in statistics at Virginia Tech under the supervision of Xinwei Deng. Her research interests include variable selection, modeling and analysis of high-dimensional data, and uncertainty quantification.

Xiankui Yang, University of South Florida xiankuiyang@usf.edu

Xiankui Yang is a Ph.D. candidate in the Department of Mathematics and Statistics at the University of South Florida, where he is pursuing advanced research in statistics. He holds a Master of Arts in Statistics from the University of South Florida and a Bachelor of Science in Mathematics and Applied Mathematics from Central China Normal University in China. His primary research interests lie in the design of experiments, response surface methodology, optimization, statistical modeling, statistical learning, and travel behavior analysis. His proficiency in R allows him to efficiently manage large datasets, implement advanced statistical techniques, and visualize results effectively, making significant contributions to his research in experimental design and statistical learning. His research focuses on experimental design, statistical modeling, and travel behavior analysis, particularly in transportation and public health contexts.

Amanda Yoder, Corning yodera27@gmail.com

Amanda Yoder is a true data enthusiast with two Master's degrees: one in Mathematics from Appalachian State and one in Statistics from Virginia Tech, which is where she discovered her love for all things data. Over the past seven years, she's been a Statistical Engineer at Corning Incorporated, spearheading cool, cutting-edge automation and machine learning projects—one of which snagged a Manufacturing Leadership Award in Machine Learning & Artificial Intelligence in 2022. On top of her work at Corning, she's the lead facilitator at eCornell for AI and Data Science courses, where she gets to geek out about AI with other professionals. Not only is Amanda a tech whiz, but she's also been an active volunteer with the ASQ CPID (American Society for Quality - Chemical and Process Industries Division) for around seven years, serving as Chair in 2023. When she's not immersed in data, you'll find her in Blacksburg, Virginia, playing strategy board games, exploring the outdoors, traveling, or getting lost in a good book.

Kade Young, Eli Lilly & Co kyoungstats@gmail.com

Kade Young is a statistician at Eli Lilly & Company specializing in drug manufacturing. He earned his PhD in statistics from North Carolina State University in 2023. His research focuses on optimal design of experiments for penalized estimation and generalized linear models.

Simon Zheng, Virginia Tech siminzheng@vt.edu

Simin Zheng, a fourth-year PhD student in the Department of Statistics at Virginia Tech, is interested in the area of reliability analysis of Artificial Intelligence systems in her research. Prior to her academic pursuits at Virginia Tech, Simin Zheng held positions at Fannie Mae as a Statistical Data Modeler and at Deloitte as a Senior Data Scientist.

Xietao Zhou, King's College London xietao.zhou@kcl.ac.uk

Xietao Zhou is currently a postgraduate researcher at King's College London and his current research interest lies in obtaining optimal designs under model uncertainty. Prior to this, his has obtained a Bachelor's degree in Mathematics, management and finance (2020) and a Master's degree in Complex systems modeling (2021) from the same university.

Call for Papers

67th Annual Fall Technical Conference

Big Data, Big Energy, Innovations in Quality, Statistics, and Data Science

October 7-9, 2025

Houston, Texas

Photo courtesy of Trace Hudson at pexels.com.

Co-Sponsored By:

American Society for Quality
Chemical and Process Industries Division
Statistics Division

American Statistical Association
Section on Physical and Engineering Sciences
Quality & Productivity Section

We invite you to submit abstracts for presentation at the 67th Annual Fall Technical Conference to be held on October 7-9, 2025 (short courses to be held October 7) in Houston, TX. Houston is known for space exploration, energy, and its culinary scene. The Fall Technical Conference has long been a forum for both statistics and quality, and is co-sponsored by the American Society for Quality (Chemical and Process Industries Division and Statistics Division) and the American Statistical Association (Section on Physical and Engineering Sciences and Section on Quality and Productivity). The goal of this conference is to engage researchers and practitioners in a dialogue that will lead to the more effective use of statistics to improve quality and foster innovation.

Please submit an abstract online if you are interested in presenting an applied or expository paper in any of the categories of: Statistics, Quality, Experiment Design, Machine Learning, or Tutorial/Case Studies. The work in a presentation should be strongly justified by an application to a problem in engineering, manufacturing, the process/chemical industry, the physical sciences, or a service industry. The mathematical level of the papers may range from introductory to advanced (e.g., that of the *Journal of Quality Technology* or *Technometrics*). Please note which level of audience is targeted (Introductory, Intermediate, or Advanced) so that the committee can assign papers appropriately and plan a balanced program. The program committee welcomes any suggestions for special

session topics or speakers. If you have any suggestions or proposals, please contact one of the program committee members listed below. For more information please visit http://www.falltechnicalconference.org

Program Committee

SPES
Oksana Chkrebtii
The Ohio State University

Q&P
Lauren Wilson
Sandia National Lab

STAT
Karen Hulting
Hulting Consulting LLC

CPID
Annie Booth
NC State University

General Conference Chair
Jennifer Kensler
Shell
jennifer.kensler@shell.com

Abstract Submission Information

The abstract submission deadline is February 28, 2025.

Online Abstract Submission
www.falltechnicalconference.org/submit-an-abstract

Abstract Content

The abstract should include a session preference, target audience, and the following three components:
1. Motivation or background
2. Description of work done
3. Significance. Are there improvements, applications, new abilities, new points of view, etc.? How will the status quo be changed?

Session Preference (choose one)
___ Statistics
___ Quality
___ Experimental Design
___ Tutorial/Case Study
___ Machine Learning
___ Data Science

Target Audience (choose one)
___ Introductory/Practitioner
___ Intermediate
___ Advanced/Theoretical

2025 ASA/IMS Spring Research Conference on Statistics in Industry and Technology

New York, New York

Annual ASA/IMS Spring Research Conference (SRC) on Statistics in Industry and Technology will be hosted by Paul H. Chook Department of Information Systems and Statistics, Baruch College, City University of New York. Plans are underway.

Join us at **SRC 2025**!

Location: New York, NY
Dates: June 3-5, 2025