

The Fourth International Conference on the Interface between Statistics and Engineering

20 - 22 June 2016

University of Palermo, Italy

Organizer

Department of Chemical and Managerial Engineering,
Manufacturing and Information Sciences, University of Palermo

Co-organizers

Department of Systems Engineering and Engineering Management,
City University of Hong Kong

School of Industrial and Systems Engineering,
Georgia Institute of Technology



TABLE OF CONTENTS

INTRODUCTION	1
WELCOME ADDRESS FROM THE HOST ORGANIZER	2
GREETINGS FROM CONFERENCE CHAIR	3
CONFERENCE COMMITTEES	4
CONFERENCE VENUE AND TRANSPORTATION	6
CONFERENCE PROGRAM	8
KEYNOTE SPEECH I	10
KEYNOTE SPEECH II	11
BANQUET SPEECH	12
PRESENTATION SCHEDULE	13
ABSTRACTS	17

University of Palermo

Viale delle Scienze



INTRODUCTION

Organizer

- Department of Chemical and Managerial Engineering,
Manufacturing and Information Sciences,
University of Palermo

Co-organizer

- Department of Systems Engineering and Engineering Management,
City University of Hong Kong
- School of Industrial and Systems Engineering,
Georgia Institute of Technology

Aim and Scope

The ICISE Conference aims to provide a platform for innovative creation, development, and dissemination of research ideas and results on the interface between statistics and engineering for the support of complex system design and operation, quality and reliability improvement, and optimal proactive decision-making.

The topics of ICISE include, but not limited to, the followings:

- Applied statistics
- Data mining and machining learning
- Statistical data analysis integrated with engineering models
- Quality and reliability engineering
- Optimal decision making
- Optimal manufacturing and service enterprise systems
- Statistical controls in semiconductor industry
- System informatics and control

WELCOME ADDRESS FROM THE HOST ORGANIZER



It is a great pleasure and great honor for us to host here, at the University of Palermo, the fourth edition of the International Conference on the Interface between Statistics and Engineering.

As researchers and teachers, we have devoted our career to explore and to improve such interface, looking for innovative solutions to new engineering problems and - at the same time - finding new statistical methods with more general validity; therefore, we are convinced that being on such interface, although a bit uncomfortable, is an increasingly important challenge and a strong need for the society.

Reducing the waste from experimentations and minimizing their cost is one of the top priorities in our age of sustainable growth.

Finding solutions to promptly analyzing the massive data sets generated by more and more intricate networks and sensors is another great priority in the age of information and communication technologies.

Interpreting and anticipating the explicit and implicit needs and desires of people is another top priority in the present age characterized by excellent services.

These are only three examples of research areas lying in the interface statistics-engineering.

Engineers equipped with statistical tools and educated with a statistical mindset are ever more needed in the labor market and can realistically aim to qualified jobs.

Our University of Palermo has a quite long tradition on the statistical education of engineers, since the founding of the “managerial engineering” school, here started at the beginning of the 80’s.

It has been a long journey, not always easy, but still – we can admit – very rewarding.

The presence here today of such honorable participation to this conference is a precious reward of such history and repayment for our efforts on this field.

We want to thank the Conference Chair and the Advisory Committee, and all participants, authoritative researchers and scientists.

We hope you will enjoy the Conference with so wide range of presentations and will appreciate the organization of all related events.

Moreover, we hope you will enjoy the beautiful town of Palermo, its signs of an ancient history of mixed cultures and vivid civilization.

Stefano BARONE & Alberto LOMBARDO
University of Palermo

GREETINGS FROM CONFERENCE CHAIR



On behalf of the Department of Systems Engineering and Engineering Management (SEEM) and the ICISE Program Committee, I am pleased to extend my warmest greetings to everyone attending the fourth International Conference of Interface between Statistics and Engineering 2016 (ICISE2016) at University of Palermo.

The aim of the ICISE Conference is to provide a platform for innovative creation, development, and dissemination of research ideas and results on the interface between statistics and engineering for the support of complex system design and operation, quality and reliability improvement, and optimal proactive decision-making. The previous ICISE conferences in Beijing, China (2009), Tainan, Taiwan (2012) and Hong Kong (2014) had been extremely successful in providing a platform for applied/theoretical statisticians and science/engineering researchers on interesting discussions and exchanges about various important research topics.

This year we are particularly fortunate to have two well famous researchers as our keynote speakers, Prof. Henry Wynn (London School of Economics and Political Science) and Prof. Jianjun Shi (Georgia Institute of Technology). I am sure the conference participants will greatly benefit from their insightful speech and discussions, as well as many reputable researchers and scholars from North America and the rest of the world.

Finally, I would like to offer you my best wishes for a most enjoyable and productive meeting.

Kwok L. TSUI

Head and Chair Professor of Industrial Engineering
Department of Systems Engineering and Engineering Management
City University of Hong Kong

CONFERENCE COMMITTEES

Advisory Committee

Chair

Prof. Jeff Wu

Georgia Institute of Technology

Members

Prof. Way Kuo

City University of Hong Kong

Prof. Alberto Lombardo

University of Palermo

Prof. Vijay Nair

University of Michigan

Program Committee

Co-Chairs

Prof. Stefano Barone

University of Palermo

Prof. William Li

University of Minnesota

Prof. Kwok L. Tsui

City University of Hong Kong

Members

Prof. Ching-Shui Cheng

Academia Sinica

Prof. MK Jeong

Rutgers University

Prof. Wei Jiang

Shanghai Jiaotong University

Prof. Judy Jin

University of Michigan

Prof. Changsoon Park

Chung Ang University

Prof. Roshan Joseph

Georgia Institute of Technology

Prof. Xavier Tort-Martorell

Univeritat Politècnica de

Catalunya

Prof. Fugee Tsung

The Hong Kong University of

Science and Technology

Prof. Min Xie

City University of Hong Kong

Prof. Dan Yu

China Academy of Science

Prof. Ji Zhu

University of Michigan

Prof. Shiyu Zhou

University of Wisconsin-Madison

Conference Organizing Committee

Co-Chairs

Dr. K. S. Chin

Prof. Stefano Barone

City University of Hong Kong

University of Palermo

Members

Dr. Louis Liu

Ms. Joana Li

Ms. Lolli Lee

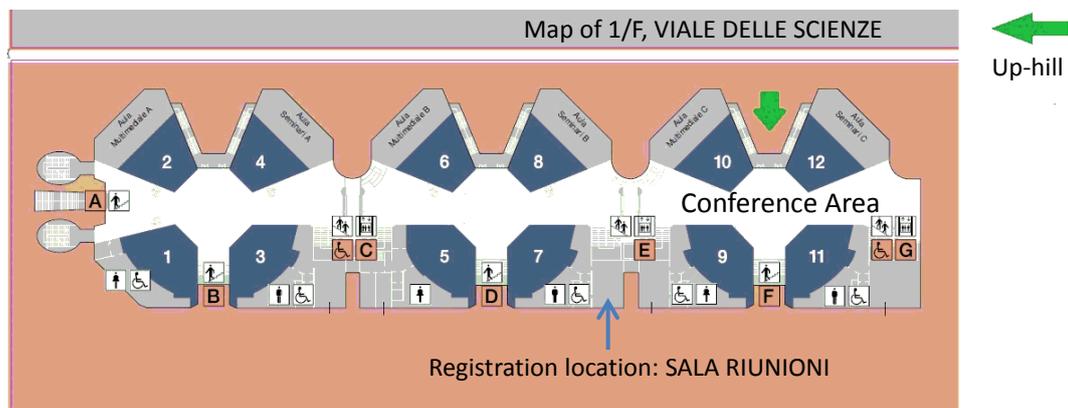
Ms. Yuki Lui

City University of Hong Kong

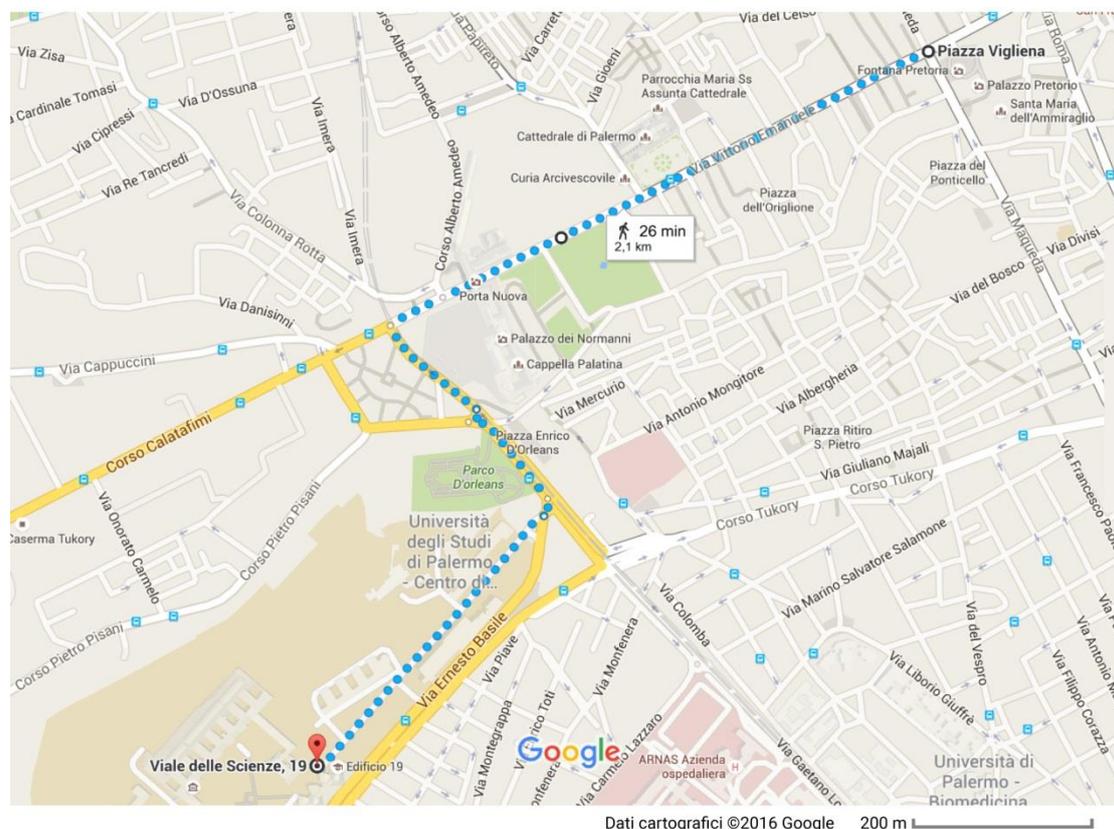
CONFERENCE VENUE AND TRANSPORTATION

Conference main venue: Room 9, 1/F, Viale delle Scienze,
University of Palermo

Parallel session rooms: Room 9 to 12



The walking route from Piazza Vigliena (i.e. Quattro Canti) to
Conference site (2.1 Km and 26 minutes walk):



Transportation information:

From Palermo airport to the city center it is possible to take a shuttle bus or taxi. The shuttle bus leaves every 30 minutes from outside the arrivals building (follow the signs airport bus-"Prestia e Comandé"). There are several bus-stops in town. The last one is at central station. Taxi from Palermo airport to city center hotel cost 35-40 € depending on the location of the hotel. Better to check with the driver before.

It is also possible to fly to/from Trapani airport. Trapani airport is a Ryanair hub and it is not too distant from Palermo city center (about 1 hour 15 minutes drive). There is a bus directly connecting Trapani airport to Palermo city center, see Terravision website, http://www.terravision.eu/airport_transfer/bus-trapani-airport-palermo/

Car rent:

Do you wish to rent a car to explore Palermo with its surroundings and the whole Sicilian Island? If yes, please look at the following website. <https://www.autoeuropa.it/rent/index.aspx>

Motorbike rent:

Wish to rent a motorbike to explore Palermo with its surroundings and the whole Sicilian island? Look at this website: <http://www.sicilymotorent.it>

Parking:

For those who have rented a car/motorbike and stay in a hotel without parking, a recommended garage with video surveillance is underground "Piazza Vittorio Emanuele Orlando" which is the square in front of the Court of Palermo so it is a very safe area.

With Google maps you can check the path to follow and the distance, by car or on foot from your hotel. You may drop off your luggage at the hotel and drive your car to the garage. Rates are per hour but with a ceiling of € 8 for a maximum duration of 24 hour stopover.

CONFERENCE PROGRAM

Day 1 of Conference – 20 Jun 2016, Monday	
08:30-09:00	Registration (Location: SALA RIUNIONI)
09:00-09:20	Conference Opening by Maurizio CARTA, Kwok L. TSUI, Stefano BARONE and William LI (Room 9)
09:20-10:30	Keynote Speech I by Henry WYNN (Room 9)
10:30-11:00	Refreshment Break
11:00-12:30	Parallel Session I (Room 9 to 12)
12:30-13:30	Lunch (near conference venue)
13:30-15:00	Parallel Session II (Room 9 to 12)*
15:00-15:30	Refreshment Break
15:30-17:30	Parallel Session III (Room 9 to 12)*
17:30-18:00	Break
18:00-20:30	Banquet ¹ and Banquet Speech by Xiao-Li MENG

Day 2 of Conference – 21 Jun 2016, Tuesday	
09:00-09:20	Registration (Location: SALA RIUNIONI)
09:20-10:30	Keynote Speech II by Jianjun SHI (Room 9)
10:30-11:00	Refreshment Break
11:00-12:30	Parallel Session IV (Room 9 to 12)
12:30-13:30	Lunch (near conference venue)
13:30-15:00	Parallel Session V (Room 9 to 12)
15:00-15:30	Refreshment Break
15:30-17:30	Parallel Session VI (Room 9 to 12)
17:30-18:00	Break
18:00-20:30	Seafood Dinner ² (Registered participants only)

Day 3 of Conference – 22 Jun 2016, Wednesday	
09:00-10:30	Parallel Session VII (Room 9 to 10)
10:30-11:00	Refreshment Break
11:00-12:30	Parallel Session VIII (Room 9 to 11)
12:30-13:30	Lunch (near conference venue)
13:40-17:30	Industrial Visits (Registered participants only)

¹ Caffè del Teatro Massimo, Piazza Giuseppe Verdi, Palermo

² Grand Hotel Villa Igiea, Salita Belmonte, 43, Palermo (GPS: 38.144653, 13.369724)

*City walking tour is arranged in parallel to Parallel Session II and III.

KEYNOTE SPEECH I

UNCERTAINTY AND ROBUSTNESS IN ENGINEERING DESIGN

Prof. Henry WYNN

London School of Economics and Political Science



Abstract

Robust Engineering Design (RED) has drawn in many areas of statistics, but maybe especially experimental design and stochastic processes. The work has been highly successful, particularly when fused with computer experiments (surrogate or meta-modelling). But there remains some separation: statisticians are perhaps too content to treat engineering models as black boxes and engineers resist making use of the more advanced statistical methods. But under the umbrella of Uncertainty Quantification (UQ), these barriers are finally breaking down. The area is sketched briefly and examples given of new interface areas taken from the speaker's own collaborations.

About the Speaker

Henry P. Wynn (BA University of Oxford, PhD Imperial College, London 1970) is Emeritus Professor of Statistics at the London School of Economics (LSE). He was head of the LSE Department of Statistics in 2003-2006, and part-time Scientific co-Director of EURANDOM (Netherlands) in 2000-2005. Before that he was Professor of Mathematical Statistics, Dean of Mathematics and co-founder of the Engineering Design Centre at City University, London, and then Professor of Industrial Statistics and founding Director of the Risk Initiative and Statistical Consultancy Unit at the University of Warwick. He has published widely in theoretical and applied statistics, with an emphasis on engineering applications. He holds the Guy Medal in Silver from the Royal Statistical Society, the Box Medal from the European Network for Business and Industrial Statistics, is an Honorary Fellow of the UK Institute of Actuaries and a Fellow of the Institute of Mathematical Statistics.

KEYNOTE SPEECH II

MANUFACTURING ANALYTICS: SYNERGIES BETWEEN ENGINEERING AND STATISTICS

Prof. Jianjun SHI

Georgia Institute of Technology



Abstract

In advanced manufacturing systems, the rapid advances in cyber-infrastructure ranging from sensor technology and communication networks to high-powered computing have resulted in temporally and spatially dense data-rich environments. With massive data readily available, there is a pressing need to develop advanced methodologies and associated tools that will enable and assist (i) the handling of the rich data streams communicated by the contemporary complex engineering systems, (ii) the extraction of pertinent knowledge about the environmental and operational dynamics driving these systems, and (iii) the exploitation of the acquired knowledge for more enhanced design, analysis, diagnosis, and control of them.

Addressing this need is considered very challenging because of a collection of factors, which include the inherent complexity of the physical system itself and its associated hardware, the uncertainty associated with the system's operation and its environment, the heterogeneity and the high dimensionality of the data communicated by the system, and the increasing expectations and requirements posed by real-time decision-making. It is also recognized that these significant research challenges, combined with the extensive breadth of the target application domains, will require multidisciplinary research and educational efforts.

This presentation will discuss some research challenges, advancements, and opportunities in synergies of engineering and statistics for system performance improvement. Specific examples will be provided on research activities related to the integration of statistics and engineering knowledge in various applications.

Real case studies will be provided to illustrate the key steps of system research and problem solving, including (1) the identification of the real need and potential in problem formulation; (2) acquisition of a system perspective of the research; (3) development of new methodologies through interdisciplinary methods; and (4) implementation in practice for significant economical and social impacts. The presentation will emphasize the examples of research achievements, as well as how the achievements were achieved.

About the Speaker

Dr. Jianjun Shi is the Carolyn J. Stewart Chair and Professor at H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. Prior to joining Georgia Tech in 2008, he was the Lawton and Johnson Chair Professor of Engineering, Professor of the Department of Industrial and Operations Engineering, and Professor of the Department of Mechanical Engineering at the University of Michigan. He received his B.S. and M.S. in Electrical Engineering at the Beijing Institute of Technology in 1984 and 1987, and his Ph.D. in Mechanical Engineering at the University of Michigan in 1992.

Professor Shi's research interests focus on system informatics and control for the design and operational improvements of manufacturing and service systems. He has published one book and more than 160 papers (110+ Journal papers, and collectively received about 5500+ paper citations). Professor Shi is the founding chairperson of the Quality, Statistics and Reliability (QSR) Subdivision at INFORMS. He is currently serving as the Focus Issue Editor of IIE Transactions on Quality and Reliability Engineering; Associate Editor of ASME Transactions, Journal of Manufacturing Science and Engineering; Editor of Journal of Systems Science and Complexity; and Senior Editor of Chinese Journal of Institute of Industrial Engineering. He is a Fellow of IIE, a Fellow of ASME, a Fellow of INFORMS, an Elected Member of the International Statistical Institute (ISI), an Academician of the International Academy for Quality, and a life member of the American Statistics Association (ASA).

BANQUET SPEECH

IT IS THE TIME TO CONSIDER MARRIAGES . . .

Prof. Xiao-Li MENG

Harvard University



About the Speaker

Xiao-Li Meng, Dean of the Harvard University Graduate School of Arts and Sciences (GSAS), Whipple V. N. Jones Professor and former chair of Statistics at Harvard, is well known for his depth and breadth in research, his innovation and passion in pedagogy, and his vision and effectiveness in administration, as well as for his engaging and entertaining style as a speaker and writer. Meng has received numerous awards and honors for the more than 150 publications he has authored in at least a dozen theoretical and methodological areas, as well as in areas of pedagogy and professional development; he has delivered more than 400 research presentations and public speeches on these topics, and he is the author of "The XL-Files," a regularly appearing column in the IMS (Institute of Mathematical Statistics) Bulletin. His interests range from the theoretical foundations of statistical inferences (e.g., the interplay among Bayesian, frequentist, and fiducial perspectives; quantify ignorance via invariance principles; multi-phase and multi-resolution inferences) to statistical methods and computation (e.g., posterior predictive p-value; EM algorithm; Markov chain Monte Carlo; bridge and path sampling) to applications in natural, social, and medical sciences and engineering (e.g., complex statistical modeling in astronomy and astrophysics, assessing disparity in mental health services, and quantifying statistical information in genetic studies). Meng received his BS in mathematics from Fudan University in 1982 and his PhD in statistics from Harvard in 1990. He was on the faculty of the University of Chicago from 1991 to 2001 before returning to Harvard as Professor of Statistics, where he was appointed department chair in 2004 and the Whipple V. N. Jones Professor in 2007. He was appointed GSAS Dean on August 15, 2012.

PRESENTATION SCHEDULE

Day 1 of Conference – 20 Jun 2016, Monday

Room No.:	Room 9	Room 10	Room 11	Room 12
	Parallel Session I			
	S1 Peihua QIU	S17 Henry WYNN/ Ron BATES	S5 K. S. CHIN	S15 Dave WOODS/ Alberto LOMBARDO
11:00 – 11:30	Xiaoming HUO ICISE-017	Ron BATES ICISE-095	William MYERS ICISE-086	Maria ADAMOU ICISE-070
11:30 – 12:00	Kamrna PAYNABAR ICISE-014	Francisco Alejandro DIAZ DE LA O ICISE-036	Haiying WANG ICISE-029	Kalliopi MYLONA ICISE-016
12:00 – 12:30	Zhisheng YE ICISE-024	Mark ATHERTON ICISE-089	Frederick PHOA ICISE-021	Jong-Seok LEE ICISE-059
12:30 – 13:30	Lunch			
	Parallel Session II			
	S6 C.S. CHENG/ Qiang ZHOU	S4 Samuel KOU	S21 Ji ZHU/ Min XIE	S7 Fugee TSUNG
13:30 – 14:00	Boxin TANG ICISE-031	Samuel KOU ICISE-072	Jian KANG ICISE-065	Fugee TSUNG ICISE-107
14:00 – 14:30	Hung YING ICISE-085	Fan LI ICISE-073	Sijian WANG ICISE-066	Bart De KETELAERE ICISE-064
14:30 – 15:00	Steven GILMOUR ICISE-061	Bin NAN ICISE-074	Anthony Joseph YEZZI ICISE-093	Chang-Yun LIN ICISE-033
15:00 – 15:30	Refreshment Break			
	Parallel Session III			
	S25 Qingpei HU	S3 Sheng-Tsaing TSENG	S14 Wei JIANG	S18 Guido MASAROTTO
15:30 – 16:00	Alain BENSOUSSAN ICISE-084	Sheng-Tsaing TSENG ICISE-004	Wei JIANG ICISE-102	Peihua QIU ICISE-011
16:00 – 16:30	Joshua LANDON ICISE-105	Tsai-Hung FAN ICISE-026	Lianjie SHU ICISE-092	Bianca Maria COLOSIMO ICISE-027
16:30 – 17:00	Sajid ALI ICISE-044	Shuen-Lin JENG ICISE-075	Dong HAN ICISE-068	Panagiotis TSIAMYRTZIS ICISE-023
17:00 – 17:30	Chi-Hyuck JUN ICISE-032	LIM Yong Bin ICISE-088		
17:30 – 18:00	Break			
18:00 – 20:30	Banquet			

Day 2 of Conference – 21 Jun 2016, Tuesday

Room No.:	Room 9	Room 10	Room 11	Room 12
	Parallel Session IV			
	S20 Grazia VICARIO	S11 Changsoon PARK	S8 Judy JIN	S24 Dan YU
11:00 – 11:30	Grazia VICARIO ICISE-041	Guido MASAROTTO ICISE-012	Matthew PLUMLEE ICISE-080	Hao PENG ICISE-001
11:30 – 12:00	Pietro TARANTINO ICISE-022	Nan CHEN ICISE-013	Qingpei HU ICISE-077	Dan YU ICISE-081
12:00 – 12:30	Bart De KETELAERE ICISE-063	I-Chen LEE ICISE-103	Judy JIN ICISE-094	
12:30 – 13:30	Lunch			
	Parallel Session V			
	S19 Grazia VICARIO	S9 Roshan VENGAZHAYIL/ William LI	S16 Geoffrey VINING	S12 Jianjun SHI
13:30 – 14:00	Oliver ROUSTANT ICISE-060	Bradley JONES ICISE-091	Geoffrey VINING ICISE-097	Nan CHEN ICISE-007
14:00 – 14:30	Rossella BERNI ICISE-054	David STEINBERG ICISE-100	Anne DRISCOLL ICISE-098	Qiang ZHOU ICISE-006
14:30 – 15:00	John Solve TYSSDAL ICISE-037	Matthew PRATOLA ICISE-076	Rob GOEDHART ICISE-028	Jianguo WU ICISE-002
15:00 – 15:30	Refreshment Break			
	Parallel Session VI			
	S23 Bjarne BERGQUIST	S2 Nozer SINGPURWALLA/ Edward CRIPPS	S13 Regina LIU/ Guido MASAROTTO	S22 Sijian WANG
15:30 – 16:00	Bjarne BERGQUIST ICISE-069	Edward CRIPPS ICISE-042	Regina LIU ICISE-010	Annie QU ICISE-079
16:00 – 16:30	Erik VANHATALO ICISE-045	Frank SAMANIEGO ICISE-096	Aurore DELAIGLE ICISE-101	Kedong CHEN ICISE-078
16:30 – 17:00	Francesca CAPACI ICISE-051	Joshua LANDON ICISE-099	Han XIAO ICISE-048	Alberto LOMBARDO ICISE-055
17:00 – 17:30	Stefan ENGLUND ICISE-052	Subrata KUNDU ICISE-043	Xiao-Li MENG ICISE-106	
17:30 – 18:00	Break			
18:00 – 20:30	Seafood Dinner (registered participants only)			

Day 3 of Conference – 22 Jun 2016, Wednesday

Room No.:	Room 9	Room 10	Room 11
	Parallel Session VII		
Time	S10 Shiyu ZHOU	S26 Anthony Joseph YEZZI	No presentation
09:00 – 09:30	Qiang HUANG ICISE-009	Wen-Liang CHANG ICISE-018	No presentation
09:30 – 10:00	Zhisheng YE ICISE-005	Xianjia WANG ICISE-035	No presentation
10:00 – 10:30	Nan CHEN ICISE-008	Chenglong LI ICISE-050	No presentation
10:30 – 11:00	Refreshment Break		
	Parallel Session VIII		
	S27 Zhisheng YE	S28 Alain BENSOUSSAN	S29 Alberto LOMBARDO
11:00 – 11:30	Huu Du NGUYEN ICISE-030	Diego ZAPPA ICISE-040	Natalie VOLLERT ICISE-053
11:30 – 12:00	Ernest FOKOUE ICISE-062	Konrad WAELDER ICISE-046	Xiaohu HUANG ICISE-087
12:00– 12:30	Riccardo BORGONI ICISE-038	Max SPOONER ICISE-090	
12:30 – 13:30	Lunch		
13:40 – 17:30	Industrial Visits (registered participants only)		
	End of Conference		

Session No.	Session Title
S1	Technometrics invited session: Novel Methods for Analyzing Image and Lifetime Data and for efficient Computation
S2	Aspects of Reliability and Survival
S3	Quality and Reliability Modeling and Inference
S4	Statistical challenges in neuroimaging and protein folding
S5	Design and Analysis of Experiments
S6	Design of Experiments
S7	JQT invited session
S8	Bayesian methods for model estimation and calibration
S9	Design of Experiment and Applications
S10	Topics in Industrial Data Analytics
S11	Multivariate monitoring and sequential testing plan
S12	Multivariate Profile Data Monitoring: Methodology and Applications
S13	Statistical Challenges in Complex Data Settings and Foundation
S14	SPC and Applications
S15	Design and Analysis of Industrial and Engineering Experiments
S16	Perspectives on Statistical Process Monitoring
S17	Uncertainty in engineering design
S18	Recent Developments in Statistical Process Monitoring
S19	Design and Analysis of physical and computer experiments
S20	ENBIS invited session
S21	Computer aided decision making
S22	Recent advance in statistical learning methods in data science
S23	Recent developments of data analysis in Swedish industry
S24	New Advancements of System Reliability Research in Chinese Academia Sinica
S25	Forecasting and Decision Making
S26	Statistical Policy Making
S27	Failure and Modeling
S28	Application Case Studies
S29	Statistical Technology

ICISE-001

**RELIABILITY OPTIMIZATION FOR SERIES SYSTEMS UNDER UNCERTAIN
COMPONENT RELIABILITIES IN THE DESIGN PHASE**

Qianru Ge¹, Hao Peng², Geert-Jan Van Houtum³, Ivo Adan⁴

¹ Eindhoven University of Technology
q.ge@tue.nl

² Academy of Mathematics and Systems Science
penghao@amss.ac.cn

³ Eindhoven University of Technology
g.j.v.houtum@tue.nl

⁴ Eindhoven University of Technology
i.adan@tue.nl

Abstract

We develop an optimization model to select a design from all possible alternatives with different reliability parameters for each critical component in a system during the design phase. Since nowadays many engineering systems are under service contracts, penalty costs should be paid by the OEMs when the total system downtimes exceed the predetermined levels. In this case, the evaluation of the life cycle costs should consider these penalty costs, which complicates the analysis. Furthermore, in the design phase for each critical component, the outcome of a development process for a certain design is in most cases uncertain with respect to the reliability requirement. Hence, all the possible designs are subject to uncertain component reliabilities. An efficient approximation method is proposed to evaluate the life cycle costs considering the penalty cost of service contracts and the uncertain component reliabilities of the design phase. Numerical experiments are conducted to compare the optimization results based on this approximation method with the results generated from other evaluation methods.

Keywords: Reliability optimization

ICISE-002

**BAYESIAN HIERARCHICAL LINEAR MODELING OF PROFILE DATA WITH
APPLICATIONS TO QUALITY CONTROL OF NANOMANUFACTURING**

Jianguo Wu¹, Yuhang Liu², Shiyu Zhou³

¹ University of Texas at El Paso
jwu2@utep.edu

² University of Wisconsin-Madison
liu427@wisc.edu

³ University of Wisconsin-Madison
shiyuzhou@wisc.edu

Abstract

Profile monitoring has been an important area in quality control where the quality of a process or product is characterized by profiles. This paper presents a general framework to connect the profile data with both explanatory variables and intrinsic processing/product parameters for simultaneous profile monitoring and diagnosis. Specifically, a hierarchical linear model with level-2 variance heterogeneity is proposed to model the relationship between profiles, explanatory variables, and intrinsic processing/product parameters. An integrated Bayesian framework for model estimation, model selection, and inference of the intrinsic parameters is proposed through blocked Gibbs sampling, intrinsic Bayes factor, and importance sampling. The effectiveness of the proposed approach is illustrated through intensive numerical studies and application to attenuation profiles for quality control of nanocomposites manufacturing.

Keywords: *Profile monitoring, Hierarchical linear model, Variance heterogeneity, Monte Carlo Markov chain, Metal-matrix nanocomposites, quality control*

ICISE-004

**NANO-SOLS SHELF-LIFE PREDICTION VIA ACCELERATED PROFILE-
DEGRADATION MODEL**

Sheng-Tsaing Tseng¹

¹Institute of Statistics, National Tsing-Hua university
sttseng@stat.nthu.edu.tw

Abstract

Nano-particles tend to aggregate and become large particles when nano-sol is in-use. This paper presents a real case study that the shelf-life prediction of nano sol can be successfully obtained by adopting pH value as an accelerating variable. First, an accelerated profile-degradation model is proposed to describe the time-evolution of the particle size distributions under three different pH values. Then, we can analytically construct a 95% confidence interval (CI) for the shelf-life of nano-sol products under a normal use condition. Note that the conventional accelerated life tests (ALTs) usually adopted temperature (or voltage) as the accelerating variable for shortening the product's life-testing time. In the counterpart, this research demonstrates an interesting study that ALT of nano-sol products can be achieved by using the pH value as accelerating variable.

Keywords: *nano-sol products, shelf-life prediction, accelerated profile-degradation model*

ICISE-005

**STRATEGIC ALLOCATION OF TEST UNITS IN AN ACCELERATED
DEGRADATION TEST PLAN**

Ye Zhisheng¹

¹ Department of Industrial & Systems Engineering, National University of Singapore
yez@nus.edu.sg

Abstract

Degradation is often defined in terms of the change of a key performance characteristic over time. It is common to see that the initial performance of the test units varies and it is strongly correlated with the degradation rate. Motivated by a real application in the semiconductor sensor industry, this study advocates an allocation strategy in accelerated degradation test (ADT) planning by capitalizing on the correlation information. In the proposed strategy, the initial degradation levels of the test units are measured and the measurements are ranked. The ranking information is used to allocate the test units to different factor levels of the accelerating variable. More specifically, we may prefer to allocate units with lower degradation rates to a higher factor level in order to hasten the degradation process. The allocation strategy is first demonstrated using a cumulative-exposure degradation model. Likelihood inference for the model is developed. The optimum test plan is obtained by minimizing the large sample variance of a lifetime quantile at nominal use conditions. Various compromise plans are discussed. A comparison of the results with those from traditional ADTs with random allocation reveals the value of the proposed allocation rule. To demonstrate the broad applicability, we further apply the allocation strategy to two more degradation models which are variants of the cumulative-exposure model.

Keywords: ADT, Sequential analysis

ICISE-006

**ESTIMATION AND APPLICATIONS OF AN EFFICIENT MULTIVARIATE
GAUSSIAN PROCESS MODEL WITH NONSEPARABLE COVARIANCE
FUNCTION**

Yongxiang Li¹, Qiang Zhou²

¹ City University of Hong Kong
yongxili-c@my.cityu.edu.hk

² City University of Hong Kong
q.zhou@cityu.edu.hk

Abstract

Multivariate GP (MGP) models are very useful in meta-modeling for complex computer simulations and characterizing multivariate profile data in statistical process control. They naturally incorporate both within-level spatial autocorrelation and cross-level correlations. For greater flexibility, MGP models can be equipped with nonseparable covariance functions. However, such models are notoriously difficult to build due to significantly increased model parameters and high computation. This work proposes a pairwise modeling technique for such models. Our method is very computationally efficient and highly scalable to larger scale problems. We will investigate its properties in the applications of both computer experiments and statistical profile monitoring.

Keywords: *Gaussian process model, computer experiment, multivariate profile monitoring*

ICISE-007

**MONITORING WAFERS' GEOMETRIC QUALITY USING AN ADDITIVE
GAUSSIAN PROCESS MODEL**

Linmiao Zhang¹, Kaibo Wang², Nan Chen³

¹ Micron

zhanglinmiao@u.nus.edu

² Tsinghua University

kwang@tsinghua.edu.cn

³ National University of Singapore

isecn@nus.edu.sg

Abstract

The geometric quality of a wafer is an important quality characteristic in the semiconductor industry. However, it is difficult to monitor this characteristic during the manufacturing process due to the challenges created by the complexity of the data structure. In this article, we propose an Additive Gaussian Process (AGP) model to approximate a standard geometric profile of a wafer while quantifying the deviations from the standard when a manufacturing process is in an in-control state. Based on the AGP model, two statistical tests are developed to determine whether or not a newly produced wafer is conforming. We have conducted extensive numerical simulations and real case studies, the results of which indicate that our proposed method is effective and has potentially wide application.

Keywords: *Geometric quality, Gaussian process, GLR*

ICISE-008

**A GENERAL WIENER PROCESS MODEL FOR HETEROGENEOUS
DEGRADATIONS BASED ON KRIGING**

Nan Chen¹, Eunshin Byon²

¹National University of Singapore
isecn@nus.edu.sg

²University of Michigan
ebyon@umich.edu

Abstract

Wiener process has been widely used as a convenient mathematical model for degradation processes. It has a variety of variants to handle heterogeneity in degradation processes of multiple units, which can incorporate random effects to account for hidden factors, or incorporate covariates to account for observable factors. In this paper, we propose a novel Wiener process model based on Kriging to characterize heterogeneous degradation processes of multiple units under different conditions. Compared with models in the literature, the proposed model can effectively account for hidden factors and observable covariates simultaneously. In addition, it automatically allows for flexible covariate link functions to describe how the covariates can influence the degradation rates nonparametrically. As a result, the proposed model includes many well-known models as special cases, and extends beyond their application scopes. To demonstrate its usefulness in practice, extensive simulation studies as well as real case examples are conducted, which clearly illustrate the promising features of the proposed model.

Keywords: Wiener Process, degradation, Kriging model, heterogeneity

ICISE-009

AN ANALYTICAL FOUNDATION FOR OPTIMAL COMPENSATION OF THREE-DIMENSIONAL SHAPE DEFORMATION IN ADDITIVE MANUFACTURING

Qiang Huang¹

¹Epstein Department of Industrial and Systems Engineering
University of Southern California
qiang.huang@usc.edu

Abstract

Additive Manufacturing (AM) or three-dimensional (3D) printing is a promising technology that enables the direct fabrication of products with complex shapes without extra tooling and fixturing. However, control of 3D shape deformation in AM built products has been a challenging issue due to geometric complexity, product varieties, material phase-changing and shrinkage, and interlayer bonding. One viable approach for accuracy control is through compensation of the product design to offset the geometric shape deformation. This work provides an analytical foundation to achieve optimal compensation for high-precision AM. We first present the optimal compensation policy or the optimal amount of compensation for 2D shape deformation. By analyzing its optimality property, we propose the minimum area deviation (MAD) criterion to offset 2D shape deformation. This result is then generalized by establishing the minimum volume deviation (MVD) criterion and by deriving the optimal amount of compensation for 3D shape deformation. Furthermore, MAD and MVD criteria provide convenient quality measure or quality index for AM built products that facilitate online monitoring and feedback control of shape geometric accuracy.

Keywords: 3D Printing, Accuracy Control, Optimal compensation policy, Accuracy measure

ICISE-010

NONPARAMETRIC TOLERANCE TUBES FOR FUNCTIONAL DATA

Regina Liu¹

¹ Rutgers University
rliu@stat.rutgers.edu

Abstract

Tolerance intervals and tolerance regions are important tools for statistical quality control and process monitoring of univariate and multivariate data, respectively. This paper discusses the generalization of tolerance intervals/regions to tolerance tubes in the infinite dimensional setting for functional data. In addition to the generalizations of the commonly accepted definitions of the tolerance level of beta-content or beta-expectation, we introduce the new definition of alpha-exempt beta-expectation tolerance tube. The latter loosens the definition of beta-expectation tolerance tube by allowing alpha (usually pre-set by domain experts) portion of each functional be exempt from the requirement. More specifically, an alpha exempt beta-expectation tolerance tube of a sample of n functional data is expected to contain $[n \times \beta]$ functionals in such a way that at least $(1-\alpha) \times 100\%$ portion of each functional is contained within the boundary of the tube.

Those proposed tolerance tubes are completely nonparametric and thus broadly applicable. We investigate their theoretical justification for and properties. We also show that the alpha exempt beta-expectation tolerance tube is particularly useful in the setting where occasional short term aberrations of the functional data are deemed acceptable if those aberrations do not cause substantive deviation of the norm. This desirable property is elaborated and illustrated further with both simulations and real applications in continuous monitoring of blood glucose level in diabetes patients as well as of aviation risk pattern during aircraft landing operations.

This joint work with Yi Fan, Rutgers University.

Keywords: *tolerance tube, functional data, data depth*

ICISE-011

**DESIGN SPC CHARTS USING P-VALUES AND PROCESS MONITORING USING
A DYNAMIC SAMPLING SCHEME**

Peihua Qiu¹

¹Department of Biostatistics, University of Florida
pqi@ufl.edu

Abstract

Conventional statistical process control (SPC) charts are designed using control limits; a chart gives a signal of process distributional shift when its charting statistic exceeds a properly chosen control limit. To do so, we only know whether a chart is out-of-control (OC) at a given time. It is therefore not informative enough about the likelihood of a potential distributional shift. In a recent paper, we suggested designing the SPC charts using p-values. By this approach, at each time point of process monitoring, the p-value of the observed charting statistic is computed, under the assumption that the process is in-control (IC). If the p-value is less than a pre-specified significance level, then a signal of distributional shift is delivered. This p-value approach has several benefits, compared to the conventional design using control limits. First, after a signal of distributional shift is delivered, we could know how strong the signal is. Second, even when the p-value at a given time point is larger than the significance level, it still provides us useful information about how stable the process performs at that time point. The second benefit is especially useful when we adopt a variable sampling scheme, by which the sampling time can be longer when we have more evidence that the process runs stably, supported by a larger p-value. A resulting dynamic sampling scheme will also be introduced in this talk. This is a joint research with Drs. Ansu Chatterjee, Zhonghua Li, and Zhaojun Wang.

Keywords: *statistics process control, control limit, p-value*

ICISE-012

PHASE I DISTRIBUTION-FREE ANALYSIS OF MULTIVARIATE DATA

Guido Masarotto¹, Giovanna Capizzi²

¹Department of Statistical Sciences - University of Padua - Italy
guido.masarotto@unipd.it

²Department of Statistical Sciences - University of Padua - Italy
giovanna.capizzi@unipd.it

Abstract

In the talk, a new distribution-free Phase I control chart for retrospectively monitoring multivariate data is presented. The suggested approach, based on the multivariate signed ranks, can be applied to individual or subgrouped data for detection of location shifts with an arbitrary pattern (e.g. isolated, transitory, sustained, progressive, etc.). Furthermore, since in many practical situations shifts involve only a small number of variables, we complement the procedure with a LASSO-based method for identifying the variables that are likely to be responsible for an out-of-control condition. A simulation study shows that the method compares favorably with parametric control charts when the process is normally distributed, and largely outperforms other, recently proposed, multivariate nonparametric control charts when the process distribution is skewed or heavy-tailed. An easy-to-use R package allows practitioners to perform the proposed Phase I analysis.

Keywords: *Change-point detection, Control charts, LASSO, Multivariate signed ranks, Nonparametric methods, Statistical process monitoring*

ICISE-013

**ROBUST MULTIVARIATE CONTROL CHART BASED ON GOODNESS-OF-FIT
TEST**

Chen Zhang¹, Nan Chen², Changliang Zou³

¹National University of Singapore
zhangchen@u.nus.edu

²National University of Singapore
isecn@nus.edu.sg

³Nankai University
nk.chlzou@gmail.com

Abstract

This paper proposes a distribution-free multivariate statistical process control (MSPC) chart to detect general distributional changes in multivariate process variables. The chart is deployed based on a multivariate goodness-of-fit test, which is extensible to high dimensional observations. The chart also employs data-dependent control limits, which are computed on line along with the charting statistics, to ensure satisfactory and robust charting performance of the proposed method. Through theoretical and numerical analyses, we have shown that the proposed chart is exactly distribution-free, and able to operate with unknown in-control distribution or limited reference samples. The chart also has robust IC performance as well as satisfactory OC detection power for general process changes without any assumption of the process distribution. A real-data example in semiconductor production process is presented to demonstrate the application and effectiveness of our method.

Keywords: MSPC, goodness-of-fit, nonparametric, data-dependent limits

ICISE-014

**IMAGE DENOISING AND ANOMALY DETECTION VIA SMOOTH-SPARSE
DECOMPOSITION**

Hao Yan¹, Kamran Paynabar², Jianjun Shi³

¹ ISyE at Georgia Tech
yanhao@gatech.edu

² ISyE at Georgia Tech
kpaynabr3@gatech.edu

³ ISyE at Georgia Tech
Jianjun.shi@isye.gatech.edu

Abstract

In various manufacturing applications such as steel, composites, and textile production processes, anomaly detection in noisy images is of special importance. Although there are several methods for image denoising and anomaly detection, most of these methods perform denoising and detection separately, which affects detection accuracy. Additionally, the low computational speed of some of these methods is an issue in realtime inspection. In this paper, we develop a novel methodology for anomaly detection in noisy images with smooth background. The proposed method, named as smooth-sparse decomposition, exploits regularized high-dimensional regression to decompose an image and separate anomalous regions by solving a large-scale optimization problem. To enable the proposed method for real-time implementation, a fast algorithm for solving the optimization model is proposed. Using simulations and a case study, we evaluate the performance of the proposed method and compare it with exiting methods. Numerical results show the superiority of the proposed method in terms of the detection accuracy as well as computation time.

Keywords: *Smooth-Sparse Decomposition*

ICISE-016

**APPLYING SUPERSATURATED DESIGNS UNDER RESTRICTED
RANDOMIZATION TO A TRIBOCORROSION EXPERIMENT**

Kalliopi Mylona¹, Emily Matthews², David C. Woods³

¹ Department of Statistics, University Carlos III, Getafe (Madrid) 28903, Spain, and
Southampton Statistical Sciences Research Institute, University of Southampton, UK
mkalliop@est-econ.uc3m.es

² Southampton Statistical Sciences Research Institute, University of Southampton, UK
E.S.Matthews@soton.ac.uk

³ Southampton Statistical Sciences Research Institute, University of Southampton, UK
D.Woods@soton.ac.uk

Abstract

Designing and analysing small factorial experiments with restricted randomisation is complicated by the need to (i) minimise the degree of complex aliasing between factorial effects and (ii) estimate variance components with few degrees of freedom. We explore such situations through an experiment to optimise diamond-like carbon coatings for use in orthopaedic implants. The tests explore the effects of seven factors; six factors with hard-to-change levels and one easy-to-change factor. Interactions are also of interest. Three replicates of each treatment are needed to perform all the tests and the total number of available samples is 30. In this situation, not only is there no efficient design available but the notion of an efficient design in this context has not even been defined in a meaningful manner. In addition, novel methodology will be needed to analyse the data in order to draw the correct conclusions. We propose a Bayesian optimality criterion for the design construction. For model selection, we combine ideas from Bayesian inference, REML and shrinkage regression to identify important factorial effects.

Keywords: *design of experiment, supersaturated split-plot designs, Bayesian optimality, shrinkage regression*

ICISE-017

FAST COMPUTING FOR DISTANCE COVARIANCE

Xiaoming HUO¹, Gabor J. SZEKELY²

¹School of Industrial and Systems Engineering, Georgia Institute of Technology
huo@gatech.edu

²National Science Foundation
gszekely@nsf.gov

Abstract

Distance covariance and distance correlation have been widely adopted in measuring dependence of a pair of random variables or random vectors. If the computation of distance covariance and distance correlation is implemented directly accordingly to its definition then its computational complexity is $O(n^2)$, which is a disadvantage compared to other faster methods. In this article we show that the computation of distance covariance and distance correlation of real-valued random variables can be implemented by an $O(n \log n)$ algorithm and this is comparable to other computationally efficient algorithms. The new formula we derive for an unbiased estimator for squared distance covariance turns out to be a U-statistic. This fact implies some nice asymptotic properties that were derived before via more complex methods. We apply the fast computing algorithm to some synthetic data. Our work will make distance correlation applicable to a much wider class of problems. A supplementary file to this article, available online, includes a Matlab and C-based software that realizes the proposed algorithm.

Keywords: *Distance correlation, Fast algorithm, Statistical dependence*

ICISE-018

**OPTIMAL PREVENTIVE REPLACEMENT POLICIES FOR TWO-COMPONENT
SYSTEMS WITH WEIBULL LIFETIME DISTRIBUTIONS**

Wen-Liang Chang¹, Ruey-Huei Yeh², Ernestine Liu³

¹ General Education of Holistic Education Center, Cardinal Junior Tien College of Healthcare and Management
chang@ctcn.edu.tw

² Department of Industrial Management, National Taiwan University of Science and Technology
rhyeh@mail.ntust.edu.tw

³ Marketing Manager, Pressure Sensitive Adhesive, 3M Taiwan Ltd.
eliu3@mmm.com

Abstract

Considering a simple two-component system, the purpose of this paper is to investigate whether the components should be replaced at the same time (called group replacement) or separately (called individual replacement) under two different structures: series and parallel. For both structures, when any component fails, a minimal repair is performed to rectify the failed component. To reduce the number of failures, a preventive replacement action is carried out. Suppose that a replacement requires a team of professional technicians to perform the action. Then, the resulting cost (called setup cost) might be relatively high compared to the minimal repair cost. In this case, it might be worthwhile to replace both components at the same time instead of replacing them separately. In this paper, the comparisons of individual and group replacement policies for series and parallel systems are conducted. Furthermore, the impacts of the downtime cost and setup cost on the optimal replacement policy are analysed, and the criterion for choosing the group replacement policy instead of the individual replacement policy is demonstrated through numerical examples. The results obtained in this paper can provide some insights of deriving the optimal replacement policy for a multi-component system.

Keywords: two-component system, minimal repair, group replacement, individual replacement

ICISE-021

A GENERALIZED CLASS OF QUATERNARY CODE DESIGNS WITH GOOD PROPERTIES

Frederick Kin Hing Phoa¹

¹Institute of Statistical Science, Academia Sinica
fredphoa@stat.sinica.edu.tw

Abstract

The study of good nonregular fractional factorial designs has received significant attention over the last two decades. Recent research indicates that designs constructed from quaternary codes (QC) are very promising in this regard due to its good design properties. This talk introduces a generalized class of QC designs by a sequential column deletion process. The resulting QC designs possess at least equivalent, if not better, resolutions than the traditional class of QC designs, and their wordlength patterns are mostly improved. To the extreme, it reveals a connection to regular designs when half number of columns are sequentially deleted. In addition, three new criteria are introduced to classify a general fractional factorial designs when resolution and wordlength pattern are insufficient to differentiate between the goodness of designs. These criteria clearly identify the generalized class of QC designs are better than the traditional class of QC design and the regular designs of the same sizes.

Keywords: Factorial Designs, Quaternary Codes, Aliased Structure

ICISE-022

**ONE METHODOLOGY, SEVERAL APPLICATIONS: WHERE AND HOW DESIGN
OF EXPERIMENTS SUPPORTS PRODUCT DEVELOPMENT**

Pietro Tarantino¹, Carlo Leardi²

¹Tetra Pak Packaging Solutions
pietro.tarantino@tetrapak.com

²Tetra Pak Packaging Solutions
carlo.leardi@tetrapak.com

Abstract

Purpose of this Presentation

Experimentation is a natural approach for human beings. Analytically speaking a turning point has been the introduction of statistical design of experiments (DOE) by Fisher back in 1926. Since then, DOE has been extensively used for screening purpose, optimization and robustness testing. The new trend is, instead, to run the so called computer experiments where statistical methods are used to build a simplified mathematical model to approximate the “true” behaviour of the functional model response-design factors. This work aims at showing the versatility of DOE by placing different possible applications of it along an industrial product development process like a VEE-model.

Approach

The differences in scope, execution and analysis of DOE in the different phases of VEE-model will be highlighted and the most critical aspect and opportunities of running a physical versus a computer experiments discussed.

Findings

Product development is often a complex process, starting from customer’s requirements understanding till system design, verification and validation. Therefore, the same methodology has to be declined in different ways. Time, cost and test feasibility elements play a critical role is the type of DOE that best fit the specific stage of the process.

Value of Presentation

Real applications of DOE in the food packaging industry will be critically presented with the aim of stimulating prospective discussions on the role of this methodology for industry advancement.

Keywords: Design of Experiments (DOE), Computer Experiment, Product Development, VEE-model

ICISE-023

A BAYESIAN APPROACH FOR ONLINE MONITORING OF PHASE I DATA

Konstantinos Bourazas¹, Dimitrios Kiagias², Panagiotis Tsiamyrtzis³

¹ Dept. of Statistics, Athens University of Economics and Business
kbourazas@aueb.gr

² School of Mathematics and Statistics, University of Sheffield
kiagias.dim@gmail.com

³ Dept. of Statistics, Athens University of Economics and Business
pt@aueb.gr

Abstract

In Statistical Process Control/Monitoring (SPC/M) of either discrete or continuous data, various frequentist based techniques have been developed, like Shewhart type charts, CUSUM and EWMA. All these methods require the knowledge of the in control process parameter(s), something that in practice is handled with the employment of an offline calibration (phase I) period prior to the online control/monitoring of the process (phase II). Typically, phase I estimation demands a relatively long sequence of independent and identically distributed (iid) data from the in-control distribution. Undetected phase I issues (like the presence of masked outlying observations) will contaminate the parameter estimates, seriously affecting the phase II performance. Online monitoring (free of phase I) has been proposed by frequentist self-starting methods.

In this work we propose a Bayesian alternative that intends to tackle all the aforementioned problems utilizing the (usually) available prior information. The predictive distribution will introduce the Predictive Control Chart (PCC), which will be able to perform online phase I monitoring, right after the first observable becomes available. PCC will be presented in its most general form, allowing data of any (discrete or continuous) distribution as long as it is a member of the regular exponential family. We will establish that PCC generalize self-starting methods. Simulations will test its performance against the frequentist based phase I analysis and two data sets (one continuous and one discrete) will illustrate its use.

Keywords: *Statistical Process Control and Monitoring, Predictive distribution, Exponential family*

ICISE-024

**ESTIMATION OF FIELD RELIABILITY BASED ON AGGREGATE LIFETIME
DATA**

CHEN Piao¹, YE Zhisheng²

¹ Industrial & Systems Engineering, National University of Singapore
cp@u.nus.edu

² Industrial & Systems Engineering, National University of Singapore
yez@nus.edu.sg

Abstract

Because of the exponential distribution assumption, many reliability databases recorded data in an aggregate way. Instead of individual failure times, each aggregate data point is a summation of a series of collective failures representing the cumulative operating time of one component position from system commencement to the last component replacement. The data format is different from traditional lifetime data and the statistical inference is challenging. We first model the individual component lifetime by a gamma distribution. Confidence intervals for the gamma shape parameter can be constructed using a scaled X^2 approximation to a modified ratio of the geometric mean to the arithmetic mean, while confidence intervals for the gamma rate and mean parameters, as well as quantiles, are obtained using the generalized pivotal quantity method. We then fit the data using the inverse Gaussian (IG) distribution, a useful lifetime model for failures caused by degradation. Procedures for point estimation and interval estimation of parameters are developed. We also propose an interval estimation method for the quantiles of an IG distribution based on the generalized pivotal quantity method. An illustrative example demonstrates the proposed inference methods.

Keywords: Gamma distribution, IG distribution

ICISE-026

**BAYESIAN RELIABILITY ANALYSIS OF ACCELERATED GAMMA
DEGRADATION PROCESSES UNDER TIME-SCALE TRANSFORMATION WITH
RANDOM EFFECTS**

Tsai-Hung Fan¹, Ya-Ling Huang²

¹ Graduate Institute of Statistics, National Central University
thfanncu@gmail.com

² Graduate Institute of Statistics, National Central University
erin98102@gmail.com

Abstract

Accelerated degradation tests have been widely used to assess the lifetime information of highly reliable products. In this work, we apply Bayesian approach to the accelerated degradation test based on gamma processes with random effects under power transformation of the time-scale. A mixture prior is considered to identify the parameter of time-scale transformation. Reliability inference of the failure time distribution under normal use condition will be described through the posterior sample of the underlying parameters obtained by the Markov chain Monte Carlo procedure. Simulation study is presented to evaluate the performance of the proposed method and the model fitting issue. The proposed method is applied to an LED light intensity data set as well.

Keywords: *Constant-stress accelerated degradation test, gamma process, random effects, mixture prior, MCMC, DIC*

ICISE-027

**STATISTICAL QUALITY MONITORING IN ADDITIVE MANUFACTURING:
OPPORTUNITIES AND CHALLENGES**

Bianca Maria COLOSIMO¹, Marco Grasso²

¹ Department of Mechanical Engineering, Politecnico di Milano
biancamaria.colosimo@polimi.it

² Department of Mechanical Engineering, Politecnico di Milano
marcoluigi.grasso@polimi.it

Abstract

The market of metal additive manufacturing (AM) is growing at an impressively fast rate. EU and US visions agree to consider additive manufacturing a key enabling technology to foster competitiveness in many industrial sectors (e.g., aerospace, dental and medical implants, tooling and molds – Fig. 1).



Fig. 1 – Examples of metal AM products in different industrial sectors¹

Despite of this potential, AM technology is still far from being a reliable and repeatable technology and this can represent a barrier to the further spread of the technology, especially in SMEs.

At this point in time, data gathering during metal AM processes is possible [1-5], but no assessed solutions exist to provide AM systems with actual intelligent capabilities for integrated monitoring and control, neither in the literature nor in the industrial practice. However, most of the producers of metal AM systems recognize that this is the direction and are equipping AM systems with sensors to enhance the possibility of developing new solutions to keep the AM process under control.

This contribution discusses existing solutions for data gathering, possible problems and viable solutions to implement statistical quality monitoring in real industrial practice on metal additive manufacturing.

An approach combining image data analysis, principal component analysis and clustering is presented and applied to a real case study of selective laser melting in order to show how image data and statistical modelling can be effectively combined to detect defects onset in metal additive manufacturing.

Keywords: Statistical quality monitoring, In-situ monitoring, additive manufacturing

¹ Courtesy EOS, Renishaw Concept, shapeways, n-e-r-v-o-u-s.com

ICISE-028

GUARANTEED IN-CONTROL PERFORMANCE FOR THE SHEWHART X AND X-BAR CONTROL CHARTS

Rob Goedhart¹, Marit Schoonhoven², Ronald J.M.M. Does³

¹ IBIS UvA, University of Amsterdam
R.Goedhart2@uva.nl

² IBIS UvA, University of Amsterdam
M.Schoonhoven@uva.nl

³ IBIS UvA, University of Amsterdam
R.J.M.M.Does@uva.nl

Abstract

When in-control parameters are unknown, they have to be estimated using a reference sample. The control chart performance in Phase II, which is generally measured in terms of the Average Run Length (ARL) or False Alarm Rate (FAR), will vary across practitioners due to the use of different reference samples in Phase I. This variation is especially large for small sample sizes. Although increasing the amount of Phase I data improves the control chart performance, others have shown that the amount required to achieve a desired in-control performance is infeasibly high. Thus, in order to deal with this variation, new corrections for Shewhart control charts are proposed that guarantee a minimum in-control performance with a specified probability. However, a minimum in-control performance guarantee generally lowers the out-of-control performance. To balance the tradeoff between in-control and out-of-control performance, the minimum performance threshold and specified probability can be adjusted as desired. The corrections are given in a closed form so that the bootstrap method, which has recently been suggested, is no longer required.

Keywords: *ARL, Conditional Distribution, Parameter Estimation, Statistical Process Control*

ICISE-029

**INFORMATION-BASED OPTIMAL SUBDATA SELECTION FOR BIG DATA
REGRESSION**

HaiYing Wang¹, Min Yang, John Stufken

¹ University of New Hampshire
haiying.wang@unh.edu

Abstract

For big data, a critical step to draw useful information is data reduction. Existing investigations have been focusing subsampling-based method and normalized statistical leverage scores are often used as the subsampling distribution. This approach brings in sampling errors and the information is typically obtained at the scale of the subdata size and not the full data size. In this paper, we proposed information-based optimal subdata selection (IBOSS) method from big data under the context of linear regression. This establish a framework to deterministically select informative subdata from big data. We consider two optimality criteria: the D-optimality and T-optimality. We first theoretically characterized the IBOSS subdata under these optimality criteria and then use these characterizations to develop efficient algorithms for parameter estimation. Asymptotic properties are derived for the IBOSS method motivated from the D-optimality. Practical performance of the IBOSS methods from both optimality criteria are evaluated using various simulated and real data sets. We also derive a lower bound of covariance matrices for existing subsampling-based methods using the proposed IBOSS framework.

Keywords: *Big data, Information matrix, Linear regression, Optimality criterion, Subdata*

ICISE-030

**INFERENCE FOR COMMON-CAUSE FAILURE MODELS WITH INCOMPLETE
DATA**

Huu Du NGUYEN¹, Evans GOUNO²

¹ LMBA - University de South Brittany, France
huu-du.nguyen@univ.ubs.fr

² LMBA - University de South Brittany, France
evans.gouno@univ-ubs.fr

Abstract

We consider a system with m -identical components statistically independent. The system is exposed to external shocks that may produce simultaneous failures called common-cause failures (CCFs). When k components are implied in the failure of the system, the CCF is said to be of order k and is denoted $CCF[k]$. The failure of one component of the system may not be due to an external shock. In this case, the failure is called an intrinsic failure. In the same way, we assume that the whole components of the system can fail because of a lethal shock (which is not an external shock). In practice, it is not possible to distinguish $CCF[1]$ and intrinsic failure. Likewise, it is impossible to distinguish lethal shock and $CCF[m]$. We are dealing with incomplete data. The different failure processes are described with Poisson processes along with the Binomial failure rate (BFR). We propose to estimate the parameters of the involved models. Because of the situation of incomplete data, we develop an EM (Expectation-Maximisation) algorithm to obtain the maximum likelihood estimates of the parameters. We demonstrate the efficiency of the method with simulated data. Additionally, we investigate a Bayesian approach of the problem. Choices of prior distributions and elicitation of their parameters are presented. A comparison with the previous maximum likelihood approach is conducted. The set of methods proposed can be useful for practitioners involved in risk assessment. To end with we give some recommendations in order to effectively apply the suggested approaches.

Keywords: *Common-cause failure, Binomial failure rate model, Missing data, Maximum likelihood inference, EM algorithm, Bayesian inference*

ICISE-031

A METHOD OF CONSTRUCTING SPACE-FILLING ORTHOGONAL DESIGNS

Boxin Tang¹

¹ Simon Fraser University
boxint@sfu.ca

Abstract

This paper presents a method of constructing a rich class of orthogonal designs that include orthogonal Latin hypercubes as special cases. Two prominent features of the method are its simplicity and generality. In addition to orthogonality, the resulting designs enjoy some attractive space-filling properties, making them very suitable for computer experiments.

Keywords: *computer experiment*

ICISE-032

**MARKOV BLANKET FEATURE SELECTION FOR CLASSIFICATION OF MIXED
DATA**

Junghye Lee¹, Chi-Hyuck Jun²

¹ Department of Industrial and Management Engineering, Pohang University of Science and Technology
junghye_lee@postech.ac.kr

² Department of Industrial and Management Engineering, Pohang University of Science and Technology
chjun@postech.ac.kr

Abstract

Feature selection in multivariate analysis is an important step because a model with appropriately-selected features can provide better prediction performance than a model that includes all features. A Markov blanket (MB) is the minimal set of features to explain the target variable on the basis of conditional independence (CI). Although a MB has an advantage to be applied in classification and regression problems by using different CI tests, a MB can only be used when features and the target variable are either all continuous or all categorical. In this paper, we developed a new MB discovery algorithm called StepMB, and a generalized CI test based on a likelihood-ratio that enables to discover a MB that consists of mixed continuous and categorical features. Then, we embedded a generalized CI test into StepMB to develop a new Markov blanket feature selection method called StepMB-LR. Experimental results show that StepMB-LR is effective in improving the performance of classification models in mixed data, and it is more accurate overall as compared to other methods in classification problems.

Keywords: *Conditional independence, Likelihood-ratio test, Markov blanket discovery, Filter*

ICISE-033

**RESPONSE SURFACE METHODOLOGY USING SPLIT-PLOT DEFINITIVE
SCREENING DESIGNS**

Chang-Yun Lin¹, Po Yang²

¹National Chung Hsing University, Taiwan
chlin6@nchu.edu.tw

²University of Manitoba, Canada
Po.Yang@umanitoba.ca

Abstract

Definitive screening designs are a new class of three-level designs. We investigate the performance of definitive screening designs in split-plot structures for one-step response surface methodology. The result of the projection eligibility and the study of D-efficiency and I-efficiency show that split-plot definitive screening designs perform well when the number of important factors is small. To reduce the risk of being unable to fit second-order models for response surfaces, we provide the column indexes of projections. Experimenters can assign potentially important factors to those columns to avoid ineligible projections. An example is presented to demonstrate how to analyze data for response surface methodology using the split-plot definitive screening design.

Keywords: *Central Composite Design, D-Efficiency, Eligible Projection, Equivalent-Estimation Design, Generalized Least Squares, I-Efficiency*

ICISE-035

**OPTIMAL MANUFACTURING STRATEGIES FOR STAKEHOLDERS IN
REMANUFACTURING SYSTEM**

WANG Xianjia¹, lu XIAO², Kwai-Sang CHIN³

¹Economics and Management School, Wuhan University
wangxj@whu.edu.cn

²Economics and Management School, Wuhan University, and
Department of Systems Engineering and Engineering Management, City University of Hong Kong
xiaolu2008happy@163.com

³Department of Systems Engineering and Engineering Management, City University of Hong Kong, and
School of Management, Wuhan University of Technology
mekschin@cityu.edu.hk

Abstract

Nowadays, customers have various perceptions about remanufactured products and therefore show distinct preferences. This will influence the stakeholders in remanufacturing system to develop different manufacturing strategies, such as simple or delicate remanufacturing process, in which different levels of costs will be incurred.

The purpose of this paper is to present the preliminary findings of the authors' study on the optimal operations strategies related to remanufacturing, including production, recycling and remanufacturing aspects, to be adopted by various stakeholders in a remanufacturing system. Our research emphasizes to answer the following questions: (1) What is the optimal purchase strategy for customers among remanufactured products of different quality levels? (2) In case the original manufacturer is also a remanufacturer, what is the optimal manufacturing policy for new products and remanufactured products in different quality levels? (3) If the original manufacturer and remanufacturer are separate players and the recycling is accomplished by third-party player, what are the respective optimal strategies for new product production, remanufacturing and recycling? (4) How to promote more high-quality remanufactured products? (5) What is the influence of customer's preference on manufacturing strategy for remanufactured products of different quality levels?

Mathematical analysis and models will be employed for addressing the above problems and developing analytic conclusions or simulation results by optimization and game-theory approaches. The strategic behaviors and operations strategies of various stakeholders in the remanufacturing system will be formulated and interpreted.

Keywords: Remanufacture, Manufacturing Strategies, Game theory, Optimization

ICISE-036

A SAMPLING SCHEME FOR ROBUST DESIGN AND MODEL CALIBRATION

F.A. DiazDelaO¹, A. Garbuno-Inigo¹, K.M. Zuev²

¹Institute for Risk and Uncertainty, University of Liverpool,
{fado,agarbuno}@liv.ac.uk

²Department of Computing and Mathematical Sciences, California Institute of Technology,
kostia@caltech.edu

Abstract

The goal of engineering design is to create systems that satisfy specific performance objectives subject to constraints over a period of time. Many feasible designs may satisfy the required objectives, for which it is desirable to choose an optimal design according to some set of criteria. Since modern engineering systems are inherently complex, endogenous (geometry, material properties) and exogenous (loads) information is never complete. This lack of information can be captured by modelling uncertainties probabilistically. Therefore, the objective of performance-based design is to minimise an expected loss function which depends on both the characteristics of the design space and the model parameters of the system under study.

In this work, we present an efficient stochastic optimisation method that samples from the set of optimal design configurations and associated parameter space for performance-based design. The method combines ideas such as Simulated Annealing, Importance Sampling and Markov Chain Monte Carlo. Apart from applications related to engineering design, it will be demonstrated how this sampling scheme can be used to produce statistical approximations to the output of expensive computer simulators and for the robust calibration of complex models.

Keywords: *Engineering design, stochastic optimisation, simulated annealing, sampling, MCMC, calibration*

ICISE-037

POSSIBILITIES WITH THE NONREGULAR 16-RUN TWO LEVEL DESIGNS

John Tyssedal¹

¹Department of Mathematical Sciences, NTNU
tyssedal@stat.ntnu.no

Abstract

Among two-level nonregular designs the 12 run and the 20 run Plackett-Burman designs seem to be the ones that are used the most. The two-level nonregular 16 run designs seem to be forgotten. Though several attempts have been made to promote desirable properties of these designs, it is very hard to find any application. One reason may be that practitioners are uncertain about their analysis. In this presentation we will focus on methods to analyse some of these designs, both graphical and quantitative methods, as well as pointing out some (more) desirable properties in general and with respect to blocking and restrictions on randomization in particular.

Keywords: 16 run nonregular designs

UPDATING MONITORING NETWORKS IN OVERLAY DATA MODELLING

Riccardo Borgoni¹, Diego Zappa²

¹ Dipartimento di Economia, Metodi Quantitativi e Strategie d'Impresa, Piazza dell'Ateneo Nuovo 1, Milan
riccardo.borgoni@unimib.it

² Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore, Largo Gemelli 1, Milan, Italy
diego.zappa@unicatt.it

Abstract

Purpose *Integrated circuits are built by a sequence of patterning steps that form subsequent layers onto a semiconductor wafer. The pattern created at each step should be aligned to pre-existing patterns. The difference between the current layer position and the substrate geometry is called overlay. The wafer area is divided in portions called fields that are scanned sequentially to measure overlay displacements. In order to estimate models for adjusting overlay in subsequent steps, data are collected using a network of monitoring points (see fig. 1), called target points, located at the border of the fields. However, measuring procedures are time consuming and expensive, hence, it is worth trying to reduce the number of points that are necessary to evaluate accurately the displacement in order to speed up the fabrication process.*

Methodology *In this paper, we propose a maxent and a maxmin strategy, based on the tree spanning the sampling points, to select an optimal subsample of a given size of the network. The objective function measures how much the network is spread to cover the wafer area. The computational issue in subgrid selection is that inspecting all the possible configurations of n points out of the N points originally present in the network is practically unfeasible when the number of measurement locations is even moderately high. Metaheuristic optimization methods are employed to tackle this problem. Fig 2 shows an example of subnetwork obtained by using the simulated annealing and the maxent criterion.*

Findings & Practical Implications *We compared the results obtained using the reduced network to those obtained using the full sample in terms of the precision of both the predicted overlay values and the estimates of the regression coefficients of the calibrating model. It has been found, that even halving the sample size, the performance of the reduced network remains substantially unchanged.*

Value of Presentation *The paper shows a procedure to reduce a monitoring network of a multivariate process according to an optimal criterion facing the computational burden through metaheuristic optimization. This allows fab practitioners to implement smaller sample designs to reduce time and costs of data acquisition. The procedure can be easily modified to include different objective/cost functions or to include constraints in the sampling strategy.*

Research Limitations/perspectives *Although in this paper we refer to a geometrical criterion to select the subnetwork, a model-based approach can be sensibly considered instead. Since the final objective of the analysis is to calibrate the overlay model to correct the fabrication process, the objective function can be defined in terms of the overall precision in estimating the coefficients of the model. This has some difficulties related to the multivariate nature of the regression (one needs to model the overlay in the x and y direction simultaneously) and to adjust for the heteroscedasticity typically shown by overlay data. This matter has been left for future research.*

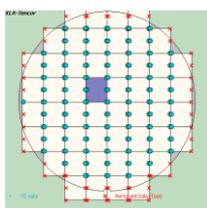


Fig. 1: Fields and targets of the original network

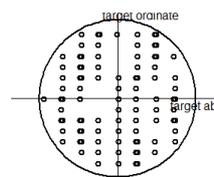


Fig. 2: Maxent subnetwork

Keywords: *Microelectronics, Overlay, Lithography, Optimal design, Spatial network reduction, Metaheuristic optimization*

ICISE-040

PROCESS CAPABILITY INDICES FOR BOUNDED DISTRIBUTIONS: THE DRY-ETCHING SEMICONDUCTOR CASE-STUDY

Riccardo Borgoni¹, Laura Deldossi², Diego Zappa³

¹ Dipartimento di Economia, Metodi Quantitativi e Strategie d'Impresa, Piazza dell'Ateneo Nuovo 1, Milan
riccardo.borgoni@unimi.it

² Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore, Largo Gemelli 1, Milan, Italy
laura.deldossi@unicatt.it

³ Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore, Largo Gemelli 1, Milan, Italy
diego.zappa@unicatt.it

Abstract

Purpose Capability indices are well known tools to estimate the mean-deviance performance of a process with respect to both a target and specification limits. A flow of contributions is available in the literature both for univariate and multivariate cases. Most of the references refer to processes with two sided specification limits but in many applications only a one sided specification limit is necessary. In this paper we consider how to compute appropriately the capability in the so called dry-etching phase in semiconductor manufacturing processes where the distribution of the outcome – the residual oxide materials on wafer surface after the dry etching phase - not only is not gaussian but it is also left censored.

Methodology One of the most controversial fact in the estimation of semiconductor process capability indices is represented by the use of rational subgroups. Since chipsets are built by splitting wafers into smaller pieces, process capability indices should reflect the precision of the production at the wafer stage. The uniqueness of this case study is that measurements are filtered (truncated) above a pre-fixed commercially driven threshold and all the wafers with at least one measurement below it, should be excluded in the computation of the capability indices. As a consequence the choice of the threshold value becomes crucial since it might determine the computation of capability indices on a limited number of wafers. To avoid this, given a sample grid of size k , we have proposed to compute capability indices for each subset of size i , $i=2, \dots, k$, excluding all the estimates that involve not feasible measurements and averaging results. Computational effort is not negligible. For example, for $i=4$ and $k=9$ we have 126 subsets and then 126 different estimates of the capability index.

Findings & Practical Implications We have compared : a) the estimate of C_{pk} by using rational subgroups and the hypothesis of censored lognormal distribution; b) the standard estimate of C_{pk} under normality assumptions without rational subgroups; c) the standard estimate of C_{pk} under normality assumptions with rational subgroups. Results with rational subgroups are significantly better than the ones obtained without it.

Value of Presentation The case study includes complexity common to many industrial process. A not marginal computation effort is required to evaluate all the subgroups combinations. A secondary outcome of the procedure is that it allows to select the subgrid that on average corresponds to the worst/best capability index and that may be helpful to identify critical/optimal production region.

Research Limitations/perspectives Priors on the weight to be assigned to each point of the grid may be applied. That depends on how much relevant is the engineering expertise of the process. Exploratory study of the dataset shows that a mixture of distributions is present. The best mixing is a methodological issue that may be used to improve the quality of the fitting. To avoid unnecessary complexities this point will not be fully treated.

Keywords: Microelectronics, Process capability indices, Dry-etching, Rational subgroups, Map reduction, Censored distributions

ICISE-041

**INFERENCE ON ERRORS IN INDUSTRIAL PARTS VIA KRIGING AND
VARIOGRAMS: A SIMULATION STUDY**

Grazia Vicario¹, Giovanni Pistone²

¹ Department of Mathematical Sciences, Politecnico di Torino, Torino
grazia.vicario@polito.it

² Collegio Carlo Alberto, Moncalieri, Torino
giovanni.pistone@carloalberto.org

Abstract

Industrial parts are routinely affected by dimensional and geometric errors due to the manufacturing processes used for their production. These errors, that usually have a typical pattern related to the employed manufacturing process, are partially controlled by means of dimensional and geometrical tolerances (such as straightness, roundness, flatness, profile) that have to be verified on the manufactured parts.

In the present paper we focus on the inference on the error of different planar surfaces whose tolerances are verified using a Coordinate Measuring Machines (CMM), the most common equipment for 3D measurement because of their accuracy and flexibility.

For this purpose we suggest the prediction of the surface model using a Kriging model on a set of measured points. Kriging is a stochastic linear interpolation technique that predicts the response values at untried locations with weights assigned to the tried locations. The weights are selected so that the estimates are unbiased and they have minimum variance. The fundamentals is the rate at which the variance between points changes over space. This is expressed as a variogram which shows how the average difference between values at points changes; it is a function of the distance and of the corresponding direction of any pair of points depicting their correlation extent. Theoretically, it is defined as the variance of the difference between the response values at two locations and it is equivalent to the correlation function for intrinsically stationary processes. The use of the variogram instead of the correlation function is recommended by the geostatisticians even if the process is not stationary.

In this paper we resort to variograms to detect possible manufacturing signatures, i.e. systematic pattern that characterizes all the features manufactured with a particular production process, and systematic errors of the CMM measurement process. We simulate different and typical manufacturing signatures of a planar surface and possible errors of a measurement process with CMM, adding a white noise. The variograms are estimated using the most robust empirical estimator in the case at hand and the likelihood (or restricted likelihood) estimator. The behavior of the omnidirectional variogram suggests the spatial correlations, giving evidence of possible non isotropy.

Keywords: *Kriging Model, Spatial Correlation, Variogram, Anisotropy, Geometric Errors*

References

- Cressie, N. A., (1997): Spatial Prediction and ordinary kriging, Mathematical Geology, 20(4), 407-421.*
Haslett, J., (1997): On sample variogram and the sample autocovariance for non-stationary time series. The Statistician, 46, 475-485.
Jin, N., Zhou, S., (2006): Signature construction and matching for fault diagnosis in manufacturing processes through fault space analysis, IIE Transactions, 38:4, 341-354

ICISE-042

**CHANGE POINT DETECTION USING BAYESIAN MIXTURE MODELS FOR
INDUSTRIAL ASSETS**

Edward Cripps¹, John Lau², Sally Wood³

¹ University of Western Australia
edward.cripps@uwa.edu.au

² University of Western Australia
john.au@uwa.edu.au

³ University of Sydney
sally.wood@sydney.edu.au

Abstract

Periodic inspection data are often used to assess the performance of industrial assets and it may be of interest to assess when/if the observations of an asset, or groups of assets, undergo changes that may imply degradation or improved performance. We present some Bayesian mixture methods designed to flexibly estimate latent partitions of asset trajectories, where each partition represents a shift in the data generating mechanism for that asset. An asset's trajectory is then estimated by averaging over all possible models, weighted by their posterior probabilities. Trans-dimensional Markov chain Monte Carlo techniques are used to estimate the mixture models. Challenges and advantages of our models are described in the context of several real time series and longitudinal data examples.

Keywords: *Bayesian mixture models, Model averaging, Change point detection, Time series, longitudinal data*

ICISE-043

**A STATISTICAL DESCRIPTION OF THE SPATIAL EXTENT OF A SPELL OF
RAINFALL**

Subrata Kundu¹

¹George Washington University
kundu98@gmail.com

Abstract

The spatial extent of a spell of rainfall is a connected region with positive rainfall at a particular time. The probabilistic behavior of the spatial extent of a spell of rainfall and various attributes of it, are issues of interest in meteorological studies. While the spatial extent can be viewed as a shape object, scale and rotational invariance of the shape are not necessarily desirable attributes from meteorological considerations. For modeling objects of the above type, we propose a computationally efficient multivalued functional representation of the shape of the rainfall region and an appropriate linear space, with an associated distance measure. While a probability density function does not exist in this situation, it is possible to develop a meaningful surrogate for a density when functional data are considered in the space determined by eigen-functions in a principal component analysis. We develop a method for deriving the probability distribution of a general functional of the shape from the surrogate probability density function for the shape and propose a nonparametric method to estimate this probability distribution. Strong consistency of the proposed estimator is established. This method is used to analyze an open access satellite data set over the West Bengal, India.

Keywords: *Shape Analysis, Functional Data Analysis*

ICISE-044

**A PREDICTIVE BAYESIAN APPROACH TO TIME-BETWEEN-EVENTS
MONITORING**

Sajid Ali¹, Antonio Pievatolo², Sonia Petrone³

¹Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milan, Italy
sajid.ali@phd.unibocconi.it

²CNR-IMATI, via Bassini 15, 20133 Milan, Italy
antonio.pievatolo@mi.imati.cnr.it

³Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milan, Italy
sonia.petrone@unibocconi.it

Abstract

In the traditional process monitoring setup, a large Phase-I data set is required to establish control limits and overcome estimation error. However, a large Phase-I data set is often problematic, especially when the sampling is expensive or not available. Moreover, with the advancement in technology, quality practitioners are more and more interested in online process monitoring. For sequential and adaptive learning, the Bayesian methodology provides a natural solution that overcomes the restrictive assumption about the Phase-I data set needed to set up a monitoring structure. In this study, we propose Bayesian control charts for time-between-events (TBE) of homogenous Poisson process. In particular, a predictive approach is adopted to introduce predictive limit control charts. Moreover, we refine the procedure by introducing a 'double-check' in case of in-control decisions, taking into account that an observation inside but very close to the control limits may provide an early signal of deterioration or out-of-control status. We show in simulated and real data studies that the suggested procedure, combining predictive control limits and 'false in-control' test, greatly improve the performance.

Keywords: Bayesian process monitoring, Deterioration check, Phase-I data, Poisson Process, Predictive control limits, Time-between-events control charts

ICISE-045

LAG STRUCTURE IN DYNAMIC PRINCIPAL COMPONENT ANALYSIS

Erik Vanhatalo¹, Murat Kulahci^{1,2}, Bjarne Bergquist³, Francesca Capaci⁴

^{1,2,3,4}Luleå University of Technology, Luleå, Sweden
erik.vanhatalo@ltu.se
bjarne.bergquist@ltu.se
francesca.capaci@ltu.se

²Technical University of Denmark, Kongens Lyngby, Denmark
muku@dtu.dk

Abstract

Purpose of Presentation

Automatic data collection schemes and abundant availability of multivariate data increase the need for latent variable methods in statistical process control (SPC) such as SPC based on principal component analysis (PCA). However, process dynamics combined with high-frequency sampling will often cause successive observations to be autocorrelated which can have a negative impact on PCA-based SPC, see Vanhatalo and Kulahci (2015).

Dynamic PCA (DPCA) proposed by Ku et al. (1995) has been suggested as the remedy ‘converting’ dynamic correlation into static correlation by adding the time-lagged variables into the original data before performing PCA. Hence an important issue in DPCA is deciding on the number of time-lagged variables to add in augmenting the data matrix; addressed by Ku et al. (1995) and Rato and Reis (2013). However, we argue that the available methods are rather complicated and lack intuitive appeal.

The purpose of this presentation is to illustrate a new and simple method to determine the maximum number of lags to add in DPCA based on the structure in the original data.

Findings

We illustrate how the maximum number of lags can be determined from time-trends in the eigenvalues of the estimated lagged autocorrelation matrices of the original data. We also show the impact of the system dynamics on the number of lags to be considered through vector autoregressive (VAR) and vector moving average (VMA) processes. The proposed method is compared with currently available methods using simulated data.

Research Limitations / Implications

The method assumes that the same numbers of lags are added for all variables. Future research will focus on adapting our proposed method to accommodate the identification of individual time-lags for each variable.

Practical Implications

The visualization possibility of the proposed method will be useful for DPCA practitioners.

Originality/Value of Presentation

The proposed method provides a tool to determine the number of lags in DPCA that works in a manner similar to the autocorrelation function (ACF) in the identification of univariate time series models and does not require several rounds of PCA.

Design/Methodology/Approach

The results are based on Monte Carlo simulations in R statistics software and in the Tennessee Eastman Process simulator (Matlab).

Keywords: Vector autoregressive process, Vector moving average process, Autocorrelation, Simulation, Visualisation

References:

- Vanhatalo, E. & Kulahci, M. (2015). *Qual. & Rel. Engn. Int.* DOI: 10.1002/qre.1858.
- Ku, W., Storer, RH. & Georgakis, C. (1995). *Chem. & Intell. Lab. Sys.* 30: 179-196.
- Rato, TJ. & Reis, MS. (2013); *Chem. & Intell. Lab. Sys.* 125: 74-86. DOI: 10.1016/j.chemolab.2013.03.009.

ICISE-046

MEASURING THE ACCORDANCE OF COLOR SHADES BASED ON THE INTERSECTION VOLUME OF CONFIDENCE ELLIPSOIDS

Konrad Waelder¹, Olga Waelder²

¹ BTU Cottbus-Senftenberg
waelder@b-tu.de

² BTU Cottbus-Senftenberg
olga.waelder@b-tu.de

Abstract

In many applications the colors of products and parts have to match with respect to defined color parameters or dimensions. Obviously, parts painted in a plant should not be distinguished from parts of the same color painted in another part. Therefore, a measure for the accordance of color shades is needed. Of course, a useful representation for color shades is necessary.

The Lab representation provides a common approach for defining colors. It is based on three parameters L, a and b. L represents lightness, a and b stand for the color-opponent dimensions, based on nonlinearly compressed coordinates.

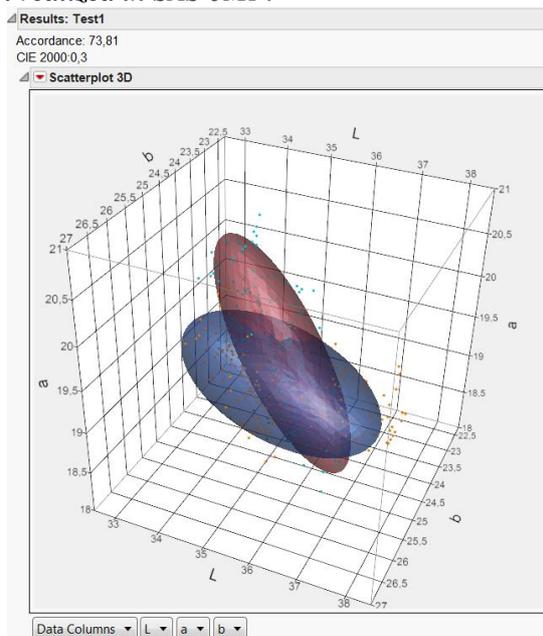
For measuring accordance we need two samples with measured color dimensions, where each sample represents a certain producing or painting process.

Now, variation within each sample can be represented by the three-dimensional confidence ellipsoid, especially by the volume of this ellipsoid. Obviously, if the ellipsoids of both samples overlap completely, no differences are recognized, at least from a statistical point of view.

For determining a measure of accordance we calculate the volume of the intersection of the two ellipsoids. Its percentage share of the sum of the volumes of the two ellipsoids defines the measure of accordance.

From a mathematical point of view it is not trivial to calculate this volume of the intersection. But Monte Carlo integration provides a quite simple method with sufficient accuracy.

In our talk we want to present a tool realized in SAS-JMP.



The attached figure shows example.

Keywords: accordance measure, confidence ellipsoid, Lab representation, SAS-JMP

ICISE-048

LEAD LAG AMONG HIGH DIMENSIONAL TIME SERIES

Han Xiao¹, Die Sun²

¹ Rutgers University
hxiao@stat.rutgers.edu

² Rutgers University
diesun@eden.rutgers.edu

Abstract

Multiple time series often exhibits lead-lag relationship among its component series. It is very challenging to identify this type of relationship when the number of series is large. We study the lead-lag relationship in the high dimensional context, using the maximum cross correlations and some other variants. Our result can also be used to test whether there is a correlation between two time series.

Keywords: *High dimensional time series, cross correlation*

ICISE-050

**A DISTRIBUTION-FREE PHASE I CONTROL SCHEME FOR SUBGROUP
LOCATION AND SCALE BASED ON A SINGLE STATISTIC**

Chenglong Li¹, Amitava Mukherjee², Min Xie³

¹Department of Systems Engineering and Engineering Management,
City University of Hong Kong, Hong Kong, China
School of Management, Xi'an Jiaotong University, Xi'an, China
lchenglon2-c@my.cityu.edu.hk

²XLRI-Xavier School of Management, Production Operations and Decision Sciences Area,
Jamshedpur, India
amitmukh2@gmail.com

³Department of Systems Engineering and Engineering Management,
City University of Hong Kong, Hong Kong, China
minxie@cityu.edu.hk

Abstract

There is a large number of literature for Phase-I or retrospective analysis based on the assumption of normality which often does not hold in practice. Over the last ten years, significant developments have taken place in area of distribution-free (nonparametric) procedure for the process location and the process scale. Most of these works are, however, intended for monitoring either the location parameter or the scale parameter. Nevertheless, practitioners recommend using a two-chart scheme, where two charts, one for monitoring the location parameter and the other monitoring the scale parameter are used in tandem. In this paper, we note that this two-chart scheme has certain practical disadvantages. There is a plethora of literature on the distribution-free (nonparametric) Phase-II control schemes for simultaneously monitoring location and scale parameters of a process distribution using a single statistic. We, however, find no research for the Phase-I or retrospective analysis in this context. We propose a distribution-free Phase-I control scheme for simultaneously monitoring location and scale parameters using a single statistic. This scheme can be used to define the in-control (IC) state of the process location as well as the scale and to facilitate identifying a proper IC reference sample. We compare the proposed scheme with nonparametric two-chart schemes based on the Kruskal-Wallis and the Fligner and Killeen statistics and the parametric two-chart schemes based on \bar{X} -bar and R as well as \bar{X} -bar and S. We see clear performance advantage of the proposed scheme in various situations. We also offer implementation guidelines and an illustrative example.

Keywords: *Distribution-free, Joint Monitoring, Lepage Statistic, Phase-I, Retrospective Analysis, Shewhart Scheme*

ICISE-051

**A TWO-STEP PROCEDURE FOR FAULT DETECTION IN THE TENNESSEE
EASTMAN PROCESS SIMULATOR**

Francesca Capaci¹, Murat Kulahci², Erik Vanhatalo³, Bjarne Bergquist⁴

¹ Luleå University of Technology, Luleå, Sweden
francesca.capaci@ltu.se

² Technical University of Denmark, Kongens Lyngby, Denmark
Luleå University of Technology, Luleå, Sweden
muku@dtu.dk

³ Luleå University of Technology, Luleå, Sweden
erik.vanhatalo@ltu.se

⁴ Luleå University of Technology, Luleå, Sweden
Bjarne.bergquist@ltu.se

Abstract

Most modern production processes involve automated data collection schemes that allow for generating large amounts of multi- or even mega-dimensional and high frequency data. Process surveillance using multivariate data has always been an academic and practical challenge. The common approach is the use of latent structures methods such as PCA and PLS (Kourti, MacGregor 1995). These methods aim to capture the static relations among the multiple variables and hence reduce the dimensionality of the problem. However the data collected in time also show serial dependence. To account for these dynamic relations, one of the proposed methods is Dynamic PCA where the original data set is expanded by adding the time-lagged variables (Ku et al. 1995). Many of these methods are successfully applied in process industries such as chemical, biological and pharmaceutical. However these processes often involve engineering control as in the case of feedback controllers. The engineering process control aims to keep the variable of interest on target by making adjustments to a manipulated variable(s). The statistical process control in the presence of engineering process control can potentially fail to identify a fault (i.e. out-of-control situation) as the latter will immediately attempt to minimize the impact of such a fault. This very crucial fact seems to have been ignored in many applications of SPC in process industry. This can lead to delay in detection and even potentially to failure to detect at all. We illustrate a two-step procedure where [1] the variables are qualitatively pre-classified prior to the analysis as manipulated and controlled variables and [2] the monitoring scheme based on latent variables is implemented for all variables as well each of the two groups separately. This allows for understanding signatures of various faults in the data and establishing more effective process surveillance methods.

A case study based on the data available from the Tennessee Eastman Process simulator under feedback control loops (Matlab) is presented. The results from the proposed method are also compared with currently available methods through simulations in R statistics software.

Keywords: *Statistical Process Control, Latent structure methods, Engineering Process Control, Close-loop systems*

References

- KOURTI, T. and MACGREGOR, J.F., 1995. Process Analysis, monitoring and diagnosis, using multivariate projection methods. Chemometrics and Intelligent Laboratory Systems, 28, pp. 3-21.*
KU, W., STORER, R.H. and GEORGAKIS, C., 1995. Disturbance Detection and Isolation by Dynamic Principal Component Analysis. Chemometrics and Intelligent Laboratory Systems, 30, pp. 179-196.

ICISE-052
**AN EXPERIMENTAL STUDY OF GRANULAR FLOW AND SEGREGATION
BEHAVIOR**

Stefan Englund¹, Bjarne Bergquist², Erik Vanhatalo³

^{1,2,3} Quality Technology, Luleå University of Technology

stefan.englund@ltu.se

Bjarne.Bergquist@ltu.se

erik.vanhatalo@ltu.se

Abstract

1. Purpose of the presentation.

Creating traceability in continuous or semi-continuous granular flow transportation chains, such as in the mining industry, faces many challenges. For example, silos or binges constitute major segregation and mixing points and warehousing in silos subject the granular media to flow or impact induced stresses. Granular materials such as grain, gravel, powder or pellets may up to some point sustain shear like a solid material but can also flow like a liquid (Behringer 1995). It is known that flow segregation of granular materials can be induced by differences in particle size, shape or density. While segregation on the surface of granular materials is well-described segregation mechanisms under the surface are not. Understanding of segregation mechanisms is important to create traceability models in granular flows.

One way to trace material in a granular flow is to add in-situ sensors, so called PATs. The traceability is reduced if the PATs segregate from the surrounding granules. The PATs may not register the 'correct' physical stresses or follow the granular flow appropriately. Available PATs today often need to be larger than the bulk particles to accommodate required electronics. The purpose of this presentation is to report on segregation behaviour of larger particles added to a flow of homogeneous sized particles and the effect of sensor casings design on segregation.

2. Results. *We show which factors that significantly affect segregation behaviour.*

3. Research Limitations/Implications. *As data collection requires manual mapping of each individual particle and surrounding bulk material the amount of data in modelling is limited. Future research will explore Particle Image Velocimetry technology (PIV) and customised software to analyse metadata from experiments more efficiently.*

4. Practical implications. *Research results will help practitioners to improve traceability in continuous and semi-continuous supply chains. Improved traceability will enable delivery of a customized quality of the granular material to different customers as well as improved root cause analyses of quality issues.*

5. Value of presentation. *Improved understanding of segregation in granular transportation chains is of value for industries such as the mining industry and especially valuable for law and regulation controlled industries where traceability is essential (e.g. food and pharmaceuticals).*

6. Method. *Experiments have been performed using granules of different shapes and densities to study flow and segregation behaviour in a transparent 2D model of a silo. The experiments are designed to mimic warehouses along an iron ore pellets distribution chain. Bulk material was discharged together with larger objects of different shapes, size and density. Responses such as estimated mixing, flow, and segregation behaviour were captured using video analysis and statistically analysed to identify significant factors affecting these behaviours.*

Keywords: *Velocity profile of granular flow inside silos, Silo discharge, Granular material flow, Size segregation*

ICISE-053

**CONSTRUCTION OF SMOOTH BORDERS FOR TREED GAUSSIAN PROCESS
MODELS**

Natalie Vollert¹, M. Ortner, J. Pilz

¹CTR Carinthian Tech Research AG / Alpen-Adria Universität Klagenfurt
natalie.vollert@ctr.at

Abstract

When usual Gaussian process models fail to adequately represent the output of computer experiments for complex physical models, one reason may be the violation of fundamental assumptions of the approach, mainly regarding stationarity. To meet this problem one can use modeling techniques based on fully nonstationary correlation functions [1], however, such an approach is often difficult to fit and computationally tractable only for a small number of data points and dimensions.

An alternative is the consideration of a partitioning of the data based on binary splits, which results in a tree structure, followed by fitting separate Gaussian process models in each of the leaves [2]. It is known that Gaussian process models scale poorly with the number of data points n , due to the inversion of $n \times n$ covariance matrices resulting in execution times of $O(n^3)$. Hence, by splitting the data also the overall computational demand can be decreased. In addition, the treed process modeling automatically leads to more locally concentrated prediction errors, which can be especially helpful when it comes to optimization.

However, whereas the non-stationary models enforce continuity, treed processes generally yield discontinuities at the leaf borders. This is quite problematic as most of the time smoothness is a key requirement when describing a physical process. Furthermore, the treed processes are mostly extrapolating the data in the vicinity of the leaf borders while information from adjacent leaf data is disregarded.

In this work, different methods for the construction of globally continuous and differentiable treed Gaussian process models are proposed. They include methods for simple spline interpolation and Gaussian process models based on combinations of neighboring processes and their derivatives. In addition, soft borders are tested by overlapping leaves to eliminate extrapolation. Finally the methods are illustrated and compared based on a scheme introduced by [3].

[1] Sampson P. D. and Guttorp P. (1992). Nonparametric Estimation of Nonstationary Spatial Covariance Structure. *Journal of the American Statistical Association*, 87, 108-119

[2] Gramacy R. B. and Lee H. K. H. (2008). Bayesian Treed Gaussian Process Models With an Application to Computer Modeling. *Journal of the American Statistical Association*, 103, 1119-1130

[3] DiMatteo I. and Genovese C. R. and Kass R. E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika*, 88, 1055-1071

Keywords: *Gaussian process, nonstationary modeling, tree partitioning, continuous borders, spline interpolation, derivative samples*

ICISE-054

**SPLIT-PLOT DESIGN AND MODELLING FOR NOVEL GAS SENSING
MATERIALS**

Rossella Berni¹, Francesco Bertocci²

¹ Department of Statistics, Computer Science, Applications - University of Florence
berni@disia.unifi.it

² Department of Information Engineering and Mathematics, University of Siena, Italy
francesco.bertocci@live.com

Abstract

This paper deals with split-plot planning and modelling for the process improvement of novel gas sensing materials. The planned design aims at optimizing the sensor response by studying the main experimental variables, such as gas sensing material, gas concentration and working temperature as well as sub-experimental variables, e.g. noise and block variables. Furthermore, mixed response surface models are applied to study all these sources of variability, by involving fixed and random effects. The theoretical issues involve a dual-response modelling approach, in which two models are evaluated and iteratively fitted through the conditional response variance modelling. It addresses three main issues: i) the planning of a split-plot design for gas sensing materials in order to optimize the response through a small number of trials, and satisfying some stringent requirements, such as low power consumption and response stability over time; ii) the introduction of a new approach for the split-plot modelling in a robust design context, where a specific model is evaluated for the response conditioning to the variance response model; iii) process optimization is directly conducted via the dual-response modelling approach.

The theoretical issues are also confirmed by the empirical results; more precisely, the dual-response modelling allows for achieving satisfactory estimates for the process variables and, simultaneously, good diagnostic evaluations. Optimal solutions are obtained for each gas sensing materials conditioning to the working temperature and by target gas, by improving the results achieved through previous studies.

Keywords: Split-plot design, MOX semiconductors, Generalized Linear Mixed Models- GLMMs

ICISE-055

A REVIEW OF THE RISK MATRIX

Stefano Barone¹, Alberto Lombardo²

¹ Università di Palermo, Department of Chemical, Managerial, Information and Mechanical Engineering
stefano.barone@unipa.it

² Università di Palermo, Department of Chemical, Managerial, Information and Mechanical Engineering
alberto.lombardo@unipa.it

Abstract

So far, Risk has been mostly defined as the expected value of a loss, in formula as P times L , being P the probability of an adverse event and L the loss incurred as a consequence of the event. The so called risk matrix widely adopted in risk management is based on such definition.

Also for favorable events one usually refers to the expected gain P times G , being G the gain incurred as a consequence of the positive event.

These “measures” are generally violated in practice. The case of insurances (on the side of losses, i.e. negative risk) and the case of lotteries (on the side of gains, positive risk) are the most obvious cases. In these cases a single person is willing to pay a higher price than the one stated by the mathematical expected value, according to measures, more or less theoretically justified. The higher the risk, the higher the unfair accepted price.

The definition of risk as expected value is justified in a long term “manager’s” perspective, in which it is conceivable to dilute the effects of an adverse event over a large number of subjects or a large number of recurrences. In other words, this definition is mostly justified on frequentist terms. Moreover, according to this definition, in two extreme situations (high-probability/low-consequence and low-probability/high-consequence), the estimated risk is low. This logic is against the principle of sustainability and the practice of continuous improvement, which should instead impose both a continuous search for lower probabilities of adverse events (higher and higher reliability) and a continuous search for lower impact of adverse events (in accordance with the fail-safe principle).

In this work a different definition of risk is proposed, which stems from the idea of Safeguard = $(1 - \text{Risk}) = (1 - P) \cdot (1 - L)$. According to this definition, the risk levels can be considered low, only when both the probability of the adverse event and the consequent loss are small.

Such perspective, in which the calculation of safeguard replaces the calculation of risk, will possibly avoid exposing the society to catastrophic consequences, sometimes due to wrong or oversimplified use of probabilistic models. Therefore, it can be seen as the citizen’s perspective to the definition of risk.

Keywords: Risk matrix, Risk & sustainability, Managerial perspective of risk, Social perspective of risk

ICISE-059

**A COST-FREE BOOSTING METHOD FOR IMBALANCED DATA
CLASSIFICATION**

Jong-Seok Lee¹, Hanhee Yun²

¹Department of Industrial Engineering, Sungkyunkwan University
jongseok@skku.edu

²Deloitte Touche Tohmatsu Limited
hanhyun@deloitte.com

Abstract

This research presents a modification of the AdaBoost (adaptive boosting) algorithm for imbalanced data classification. While the original AdaBoost was designed to minimize misclassification rate, the proposed method, which is named AdaAUC, attempts to maximize the AUC (Area Under ROC Curve) of a strong classifier. It implies that the proposed method aims at coping with class imbalance in data. This research begins by defining a loose upper bound of (1-AUC) of the strong classifier and then discovers the weights of weak learners and the sample weights in that the (1-AUC) upper bound is decreased at each iteration of the boosting framework. Unlike the cost-sensitive boosting algorithms that were previously developed for imbalanced data classification, our method is not required to determine its parameters such as misclassification costs. An experimental study was performed on eight real datasets from the machine learning repository at the University of California, Irvine. The proposed AdaAUC outperformed the existing boosting methods in terms of AUC.

Keywords: Classification, Adaptive boosting, Imbalanced data, Area under ROC curve, Cost-free approach

Acknowledgments

This research was supported by the MSIP, Korea, under the G-ITRC support program (IITP-2016-R6812-16-0001) supervised by the IITP.

ICISE-060

KRIGING AND DESIGN OF EXPERIMENTS ON CIRCULAR DOMAINS

Roustant Olivier¹, Padonou Espéran²

¹ Mines Saint-Etienne, France
roustant@emse.fr

² Mines Saint-Etienne, France
STMicroelectronics, France
padonou@emse.fr

Abstract

This research is motivated by the problem of reconstructing a spatial profile in microelectronics. More precisely, the aim is to reconstruct a variable defined on a disk (called wafer) from few measurements, typically less than 20. Furthermore, the spatial profile sometimes contains radial or angular patterns, due to the technological processes involved in their fabrication, such as rotations or diffusions.

Among spatial statistics techniques, Kriging (or Gaussian process regression) may be the preferred choice of modelizers. Firstly it provides a measure of uncertainty, due to its stochastic nature. Secondly, it is parameterized in a flexible way by a function, called kernel, that allows incorporating a priori information.

We introduce so-called polar Gaussian processes, defined as Gaussian processes on the cylinder of polar coordinates. The corresponding kernel is defined as a combination of a kernel for the radius, and a kernel on the circle for the angle. This typically allows taking into account radial and angular correlations. A construction from the ANOVA decomposition also allows a complete visualization of the different effects (radial, angular and interaction).

Of course, the problem of learning on a disk is closely linked to design of experiments. After reviewing the main designs classes, we introduce Latin cylinder designs (LCD), that generalize Latin hypercubes to polar coordinates, and propose two kinds of maximin LCDs.

The whole methodology is applied to toy functions, as well as case studies. We observe that polar Gaussian processes significantly outperform the standard Kriging technique, when the spatial profile contains radial or angular patterns.

Finally, we evoke two connected works: A relocation strategy based on the IMSE criterion, and an extension in higher dimensions. In particular, it is observed that reconstructing a radial function is done much more accurately with polar Gaussian processes.

Keywords: *Kriging, Gaussian processes, Circular domain, Polar coordinates, Design of experiments, Space-filling designs*

ICISE-061

MULTI-STRATUM DESIGNS FOR STATISTICAL INFERENCE

Steven Gilmour¹, Luzia Trinca²

¹ King's College London
steven.gilmour@kcl.ac.uk

² UNESP / Sao Paulo State University - Botucatu
ltrinca@ibb.unesp.br

Abstract

It is increasingly recognized that many industrial and engineering experiments use split-plot or other multi-stratum structures. Much recent work has concentrated on finding optimum, or near-optimum, designs for estimating the fixed effects parameters in multi-stratum designs. However, often inference, such as hypothesis tests or interval estimation, will also be required and valid inference requires pure error estimates of the variance components.

Most optimal designs provide few, if any, pure error degrees of freedom. Gilmour and Trinca (2012) introduced design optimality criteria for inference in the context of completely randomized and block designs. Here these criteria are used stratum-by-stratum in order to obtain multi-stratum designs. It is shown that these designs have better properties for performing inference than standard optimum designs.

Compound criteria, which combine the inference criteria with traditional point estimation criteria, are also used and the designs obtained are shown to compromise between point estimation and inference. Designs are obtained for two real split-plot experiments and an illustrative split-split-plot structure.

Keywords: *A-optimality, D-optimality, hard-to-change factor, mixed model, response surface, split-split-plot design*

ICISE-062

**EFFICIENT ESTIMATION OF THE FAILURE RATE USING ENSEMBLES OF
PARAMETRIC AND NONPARAMETRIC BASE LEARNERS**

Wilcox, Kenneth Tyler¹, Gouno, Evans², Fokoué, Ernest³

¹ Rochester Institut of Technology, USA
ktw5691@rit.edu

² University of South Brittany, France
evans.gouno@univ-ubs.fr

³ Rochester Institut of Technology, USA
epfeqa@rit.edu

Abstract

Failure rate estimation plays a central role in reliability studies. One of the major challenges is to describe the relationship between the propensity of failure and environmental conditions. Many existing models are limited because they impose a form on the underlying function which may not correspond to the true underlying function. The work presented here explores the promising potential of efficient estimation of the failure rate using ensemble learning estimators with base learners other than trees. The genesis of this idea comes from noticing that in the context of machine learning for lifetime data analysis, almost all existing research papers have resorted to classification and regression trees as their preferred base learners, with random forests being the method of choice and adaptive boosting being the second. In this research work, we combine both parametric and nonparametric base learners in our functional aggregate, and we intend to exploit existing computing power through parallelization for instance to offset any computation burden brought about by the use of many computationally intensive base learners like support vector machines and the relevance vector machine. We present the initial computational results of our ensemble methods on both real and simulated data.

Keywords: *Machine learning, Reliability, Failure rate, Random Forest, Random Subspace Learning, Support Vector Machine*

ICISE-063

**ENGINEERING APPLICATIONS OF THE DESIGN AND ANALYSIS OF
COMPUTER EXPERIMENTS**

B. De Ketelaere¹, K. D’huys, J. Keresztes, K. Nona, R. Watté and W. Saeys

¹Department of Biosystems – Division MeBioS, Belgium
Leuven Statistics Research Centre, KU Leuven, Belgium
bart.deketelaere@biw.kuleuven.be

Abstract

Computer simulations, such as Finite Element Modelling (FEM), non-sequential ray-tracing (NSRT), Monte Carlo light propagation modelling and Discrete Element Modelling (DEM) are often used when an analytical solution to an engineering problem is not at hand. Those methods allow for describing complex behaviour and are valid alternatives for physical experiments. Although computer power is increasing at great pace, such simulations still require a substantial time to compute. Therefore, efficient approaches to provide an accurate result in a limited time frame are highly desired. The Design and Analysis of Computer Experiments (DACE), a rapidly growing field within statistics, is a framework that can provide this efficiency. DACE combines novel approaches to the design of experiments with state-of-the-art analysis techniques that allow for modelling the often complex data that result from performing computer simulations. Those models interpolate between the simulations performed at the design points using so-called ‘meta-models’ or ‘emulators’. In this talk, we will demonstrate the advantage of DACE in four different engineering applications. The first application deals with an active thermography setting where the heat transfer and spatial temperature profiles are modelled during periodic excitation of a specimen. In a second example, the merits of DACE in optimizing the illumination of a hyperspectral set-up through NSRT are discussed, and it will be shown how DACE can handle the highly constrained input space encountered in this application. In the third application, DACE is used to develop a computationally faster alternative for a Monte Carlo light propagation model. This model links optical properties to a spatially resolved reflectance profile, by modelling the trajectories of millions of photons. The last example deals with a simulation study of the compression behaviour of fibrous biological materials in an extruder. We show how DACE is used to obtain a robust set of parameters with a minimal amount of required experiments.

Keywords: *DACE, engineering application, active thermography, ray tracing, illumination design*

ICISE-064

PROPERTIES OF PCA-BASED METHODS FOR MONITORING TIME-DEPENDENT, HIGH-DIMENSIONAL DATA

B. De Ketelaere¹

¹Department of Biosystems – Division MeBioS, Belgium
Leuven Statistics Research Centre, KU Leuven, Belgium
bart.deketelaere@biw.kuleuven.be

Abstract

Modern processes are typically highly automated and equipped with in-line sensor technologies that produce vast amounts of data in a short time. The result is the availability of large process streams that often display autocorrelation because of the fast sampling schemes. Additionally, in a substantial part of real-life processes nonstationarity is introduced because of warmup/cooldown, machine wear and variability in incoming material. This scenario of multivariate, time-dependent data is one of the most challenging encountered in statistical process monitoring (SPM), but it is often overlooked, although the separate fields of multivariate SPM and SPM for autocorrelated data have received more attention during the last decade. Approaches which are based on latent variables are a valuable direction for handling the multivariate nature, but need to be extended to cope with the time-dependent behaviour. Dynamic principal-component analysis, recursive principal-component analysis, and moving-window principal-component analysis are such extensions to cope with high-dimensional and time-dependent features. We present a short review of these methods and will provide real-data examples to help draw connections between the methods and the behaviour they display. As parameter selection for those methods is a challenging aspect for which literature is very limited, we will present possible routes for choosing them.

Keywords: *Time-dependent data, Process monitoring, Principal Component Analysis*

ICISE-065

**SCALAR-ON-IMAGE REGRESSION VIA SOFT-THRESHOLDED GAUSSIAN
PROCESSES**

Jian Kang¹

¹University of Michigan
jiankang@umich.edu

Abstract

Scalar-on-image regression is a useful model to study the association between the scalar response and a large number of imaging predictors. The focus of this work is on spatial variable selection for scalar-on-image regression, for which a new class of Bayesian nonparametric models, soft-thresholded Gaussian processes are proposed and the efficient posterior computation algorithms are also developed. Theoretically, soft-thresholded Gaussian processes provide the large prior support for the spatially varying coefficients that enjoy piecewise smoothness, sparsity and continuity characterizing the important features of imaging data. Also, the soft-thresholded Gaussian process can lead to the posterior consistency for both parameter estimation and variable selection for scalar-on-image regression. That is, under some mild regularity conditions, the proposed approach can consistently select all true spatially-dependent imaging predictors and accurately estimate their effects on the response variable, even when the number of true predictors is larger than the sample size. The proposed method is illustrated by extensive simulation studies compared with existing approaches and an analysis of Electroencephalography (EEG) data in the alcoholism study.

Keywords: *Image analysis, Gaussian processes*

ICISE-066

**REGULARIZED OUTCOME WEIGHTED SUBGROUP IDENTIFICATION FOR
DIFFERENTIAL TREATMENT EFFECTS**

Sijian Wang¹

¹ University of Wisconsin
swang@biostat.wisc.edu

Abstract

To facilitate comparative treatment selection when there is substantial heterogeneity of treatment effectiveness, it is important to identify subgroups that exhibit differential treatment effects. We propose a method that approximates a target function whose value directly reflects correct treatment assignment for patients. The function uses patient outcomes as weights instead as modeling targets. Consequently, our method can deal with binary, continuous and time-to-event outcomes in the same fashion. We first focus on identifying only directional estimates from linear rules that characterize important subgroups. We further consider estimation of differential comparative treatment effects for identified subgroups. We demonstrate the advantages of our method in simulation studies and in an analysis of a mammography screening study dataset.

Keywords: *Subgroup analysis*

ICISE-068

**BAYESIAN OPTIMALITY OF CHANGE-POINT DETECTION FOR FINITE
OBSERVATION SAMPLES**

Dong Han¹, Fugee Tsung²

¹ Shanghai Jiao Tong University, China
donghan@sjtu.edu.cn

² The Hong Kong University of Science and Technology, Hong Kong
ftsung@gmail.com

Abstract

Consider a series of finite observation samples whose distribution may change at some point in time. Our objective is to raise an alarm as soon as the change occurs, subject to a restriction on the rate of false alarms. The optimality of a change-point detection test for finite observation samples means that the detection test (alarm time) delay in some sense is the smallest among all detection tests (alarm times) with a false alarm rate no less than a given value when the probability distribution of change-point is known and the number of observation samples is limited. By introducing suitable loss random variables of detection, we obtain the optimal test of change-point detection not only for a general prior distribution of change-point but also for the observation samples being a general stochastic process.

Keywords: *Optimality, change-point detection, finite observation samples*

ICISE-069

MEASUREMENT SYSTEMS ANALYSIS OF RAILWAY MEASUREMENT CARS

Bjarne Bergquist¹, Peter Söderholm²

¹ Quality Technology and Management, Luleå University of Technology, Sweden
bjarne@ltu.se

² The Swedish Transport Administration, P.O. Box 809, 971 25 Luleå
peter.soderholm@trafikverket.se

Abstract

Purpose: *The presentation proposes ways to understand and quantify the variation component due to the measurement system of railway track properties using subsequent runs from measurement cars.*

Background: *Railway infrastructure conditions are commonly inspected by using measurement cars. The measurements are performed with some regularity, and the inspection frequencies could for instance be set taking into account the common train axle loads, railway speed or load bearing classification, number of trains passing, the known railway condition, or the availability of the measurement cars. By combining different inspections of the same track section, it is also possible to monitor the degradation of the infrastructure over time. Often, the railway system is inspected by many measurement cars, and for single tracks, measurements can be obtained from the car travelling in different directions. The measurements are performed at different speeds, related to random variation, but also to the maximum speeds at which the measurement cars operate.*

The measurements are also afflicted by external variation sources, some of which are acting with a known direction, such as the wear of the track which increases property variation. Maintenance usually (but not always) result in reduced property variation, whereas other sources such as climate related properties such as spring thaw may induce variation over time, but also induce variation that show a periodic behavior with periods with increasing as well as decreasing property variation. This presentation aims to devise a model for how these variation sources may be separated, with the main aim to classify measurement error, but also to estimate the magnitude of other variation sources.

Method: *No statistically significant differences were found between repeated measurements of cars travelling back and forth on the single track found at the Swedish Iron ore line. These measurements contain measurement error as well as error due to short term degradation and variation due to measurement. As measurement variance is added, it was concluded that the measurement variation could not be larger than the variation shown by repeat measurements. By comparing repeated measurements over time and subtracting variation due to wear, measurement variation for different cars, measurement speeds and measurement directions was estimated using Generalized Linear Models regression analysis. Co-variation between measurement cars and measurement speeds were accounted for using Ridge regression and Elastic Net regression.*

Results: *The regression analysis shows that whereas both measurement speed and the measurement car individuals correlate with the measurement variation obtained, regularized regression points to the measurement cars as the major variation factor and that different measurement cars have different measurement precision.*

Discussion and conclusion: *The study demonstrates how repeated measurements from regular process data and thus not obtained using the regular and systematic experimental procedures of measurement system analysis can be used for estimation of the variation components of the measurement system. As a side effect, the sizes of other variation sources, external to the measurement system, can be estimated.*

Keywords: *Measurement systems analysis, MSA, Repeated measurements, Measurement trains, Railway maintenance, Condition monitoring*

ICISE-070

ROBUST PRODUCT DESIGN VIA COMPUTER EXPERIMENTS

Maria Adamou¹, Dave Woods²

¹ University of Southampton, UK
M.Adamou@soton.ac.uk

² University of Southampton, UK
D.Woods@soton.ac.uk

Abstract

Computer experiments are now widely used in many areas of science and engineering to understand and exploit numerical models or simulators. In this work we focus on experiments with two types of inputs, control and noise variables. Values of the control variables can be set by the experimenter in both the computer experiment and in the physical process; values of the noise variables can be chosen in the computer experiment but are subject to uncontrolled variation in the physical process. The aim of the experiment is to find a robust product design, i.e. settings of the control variables that ensure quality output from the physical process is achieved even in the face of variation in the uncontrollable noise variables. Usually the settings of the control variables are chosen to minimise the variation in the response due to the noise variables whilst achieving a target mean response.

We assume a Bayesian approach to address this problem, using a Gaussian process model as an emulator for the output from a computer simulator. We present a sequential algorithm that (i) selects the values of the control variables for the next design point using a constrained expected improvement criterion, and (ii) selects values of the noise variables for the next design point to minimise the posterior predictive variance. The proposed approach is motivated by examples from the consumer products and pharmaceutical industries, and demonstrated through a number of examples.

Keywords: *Bayesian design, Expected improvement, Gaussian process*

ICISE-072

STATISTICAL COMPUTING IN PROTEIN FOLDING

Samuel Kou¹, Samuel Wong²

¹Harvard University, USA
kou@stat.harvard.edu

²University of Florida, USA
swkwong@stat.ufl.edu

Abstract

Predicting the native structure of a protein from its amino acid sequence is a long standing problem. A significant bottleneck of computational prediction is the lack of efficient sampling methods to explore the configuration space of a protein. In this talk we will introduce a new statistical computing method to address this challenge: PETALS, which stands for Parallely-filtered Energy Targeted All-atom Loop Sampler. PETALS combines statistical learning (namely, learning from the protein data bank) with sequential sampling to guide the computation, resulting in a fast and effective exploration of the configurations. We will illustrate the PETALS method with real protein examples.

Keywords: *Computational biology*

ICISE-073

SPATIAL BAYESIAN VARIABLE SELECTION AND GROUPING FOR HIGH-DIMENSIONAL SCALAR-ON-IMAGE REGRESSION

Fan Li¹, Tingting Zhang²

¹ Duke University

fli@stat.duke.edu

² University of Virginia

tz3b@uva.edu

Abstract

Multi-subject functional magnetic resonance imaging (fMRI) data has been increasingly used to study the population-wide relationship between human brain activity and individual biological or behavioral traits. A common method is to regress the scalar individual response on imaging predictors, known as a scalar-on-image (SI) regression. Analysis and computation of such massive and noisy data with complex spatio-temporal correlation structure is challenging. In this article, motivated by a psychological study on human affective feelings using fMRI, we propose a joint Ising and Dirichlet Process (Ising-DP) prior within the framework of Bayesian stochastic search variable selection for selecting brain voxels in high-dimensional SI regressions. The Ising component of the prior makes use of the spatial information between voxels, and the DP component groups the coefficients of the large number of voxels to a small set of values and thus greatly reduces the posterior computational burden. To address the phase transition phenomenon of the Ising prior, we propose a new analytic approach to derive bounds for the hyperparameters, illustrated on 2- and 3-dimensional lattices. The proposed method is compared with several alternative methods via simulations, and is applied to the fMRI data collected from the KLIFF hand-holding experiment.

Keywords: Bayesian, Dirichlet Process, Ising model, fMRI, variable selection, scalar-on-image regression

ICISE-074

**TUNING PARAMETER SELECTION FOR VOXEL-WISE BRAIN CONNECTIVITY
ESTIMATION VIA LOW DIMENSIONAL SUBMATRICES**

Bin Nan¹, Hai Shu²

¹University of Michigan
bnan@umich.edu

²University of Michigan
haishu@umich.edu

Abstract

The major computing cost for estimating the voxel-wise brain connectivity, especially the precision matrix, is from the tuning parameter selection. Recently we established the convergence rates for thresholding estimation of large covariance matrix and graphic-lasso estimation of large precision matrix for temporally correlated data with temporal correlations bounded by certain polynomial decay rate that can be long-memory. We found that the estimating convergence rates only depend on the temporal correlation decay rate via sample size – the number of images measured over time, which is fixed for a given data set, whereas their relations to the dimension of each image are (almost) independent of the temporal correlation decay rate. This observation motivates us to consider a tuning parameter selection procedure using cross-validation via low dimensional submatrices. Simulation results and a voxel-wise resting state fMRI data analysis will be presented.

Keywords: *Large matrix estimation, Resting state fMRI, Temporal dependence, Voxel-wise connectivity*

ICISE-075

**ANALYZING MULTISTAGE MANUFACTURING PROCESS BY FUNCTIONAL
QUADRATIC REGRESSION AND LASSO**

Shuen-Lin Jeng¹

¹Department of Statistics, National Cheng Kung University
sljeng@mail.ncku.edu.tw

Abstract

In this talk, we will propose a model building procedure for analyzing a multistage manufacturing process of products. The purpose of the analysis is for realizing the temporal influence of the process variables on final quality as well as for finding the root causes of products with abnormal quality. A data alignment technique helps to deal with the problem of uneven process length between the products. The considered model is based on the idea of functional quadratic regression by Yao and Muller (2010). Functional principle components (FPC) are implemented for the reduction of the number of variables. The variables are represented by the truncated Karhunen-Loeve expansions with the FPC scores and the orthonormal eigenfunctions. Then the functional quadratic regression is represented by a regular regression model with the FPC scores as the covariates. The resulting influential covariates of the regular regression model are selected by LASSO. A case study will illustrate the usage of the proposed model building procedure.

Keywords: *Functional principle components, functional quadratic regression, LASSO, multistage manufacturing process*

ICISE-076

BAYESIAN CALIBRATION OF STOCHASTIC SIMULATORS

Matthew Pratola¹, Oksana Chkrebtii²

¹Dept. of Statistics, The Ohio State University
mpratola@stat.osu.edu

²Dept. of Statistics, The Ohio State University
oksana@stat.osu.edu

Abstract

Inference on large-scale models is of great interest in modern science. Examples include deterministic simulators of fluid dynamics to recover the source of a pollutant, or stochastic agent-based simulators to infer features of consumer behaviour. When computational constraints prohibit model evaluation at all but a small ensemble of parameter settings, exact inference becomes infeasible. In such cases, emulation of the simulator allows the interrogation of the approximate model at arbitrary parameter values. This type of approximate inference is referred to as computer model calibration. The choice of the emulator model is a critical aspect of calibration. Existing approaches treat the mathematical model as implemented on computer as an unknown but deterministic response surface. However, in many cases the underlying mathematical model, or the simulator approximating the mathematical model, are not deterministic and in fact have some uncertainty associated with their output. In this paper, we propose a Bayesian statistical calibration model for stochastic simulators. The approach is motivated by two applied problems: a deterministic mathematical model of intra-cellular signalling whose implementation on computer nonetheless has discretization uncertainty, and a stochastic model of river water temperature commonly used in hydrology. We show the proposed approach is able to map the uncertainties of such non-deterministic simulators through to the resulting inference while retaining computational feasibility.

Keywords: *Uncertainty Quantification, Computer Experiments, Differential Equation, Stochastic Simulation, Physical Statistical, Models*

ICISE-077

**BAYESIAN ANALYSIS FOR SOFTWARE RELIABILITY GROWTH MODELS
WITH REPAIR TIME INCORPORATED**

Lujia Wang¹, Qingpei Hu², Jian Liu³, Min Xie⁴

¹ Academy of Mathematics and Systems Science, Chinese Academy of Sciences
lujiawang@amss.ac.cn

² Academy of Mathematics and Systems Science, Chinese Academy of Sciences
qingpeihu@amss.ac.cn

³ Department of Systems and Industrial Engineering, University of Arizona
jianliu@email.arizona.edu

⁴ Department of Systems Engineering and Engineering Management, City University of Hong Kong
minxie@cityu.edu.hk

Abstract

Nearly all conventional Bayesian approaches to estimating software reliability were only for fault detection process which is based on the impractical assumption that the faults are corrected immediately with no debugging time delay. There are few attempts on Bayesian approaches considering fault correction process. In this paper, we derive effective parameter estimation algorithm based on Bayesian framework not only for FDP, but also for combined FDP & FCP. And then model comparison is studied to find the fittest model for combined FDP & FCP. To have a better understanding of the estimation performance, a simulation study is conducted to compare the performance of the proposed Bayesian approach with that of MLE method. Furthermore, it is also illustrated with two practical applications to software fault detection and correction data.

Keywords: *Reliability Growth, Repair Time, Software Reliability*

ICISE-078

**AN EASY-TO-IMPLEMENT VARIABLE SELECTION METHOD FOR MODELS
FOLLOWING HEREDITY**

Kedong Chen¹, William Li², Sijian Wang³

¹ University of Minnesota
chen3548@umn.edu

² University of Minnesota
wli@umn.edu

³ University of Wisconsin, Madison
wangs@stat.wisc.edu

Abstract

In many practical regression problems, it is desirable to select important variables with heredity constraint satisfied. In other words, when an interaction term is selected, it is preferred to select all the corresponding main effects as well. In this paper, we propose a general strategy to maintain heredity in variable selection through a novel heredity-induced data standardization. After the standardization, any variable selection method (including stepwise selection, lasso, SCAD and others) can be applied and the selected model is automatically guaranteed to satisfy the heredity constraint. Furthermore, the same procedure works for all types of regression including linear regression, generalized linear regression and regression with censored outcome. Therefore, our proposed strategy is easy (almost effortless) to implement in practice to maintain the heredity. Simulations and real examples are used to illustrate the merits of the proposed methods.

Keywords: Variable Selection, Heredity, Easy-to-Implement

ICISE-079

A GROUP-SPECIFIC RECOMMENDER SYSTEM

Xuan Bi¹, Annie Qu², Junhui Wang³, Xiaotong Shen⁴

¹ University of Illinois at Urbana-Champaign
xuanbi2@illinois.edu

² University of Illinois at Urbana-Champaign
anniequ@illinois.edu

³ City University of Hong Kong
junhwang@cityu.edu.hk

⁴ University of Minnesota
xshen@umn.edu

Abstract

In recent years, there has been a growing demand to develop efficient recommender systems which track users' preferences and recommend potential items of interest to users. In this paper, we propose a group-specific method to utilize dependency information from users and items which share similar characteristics under the singular value decomposition framework. The new approach is effective for the "cold-start" problem, where, in the testing set, majority responses are obtained from new users or for new items, and their preference information is not available from the training set. One advantage of the proposed model is that we are able to incorporate information from the missing mechanism and group-specific features through clustering based on the numbers of ratings from each user and other variables associated with missing patterns. In addition, since this type of data involves large-scale customer records, traditional algorithms are not computationally scalable. To implement the proposed method, we propose a new algorithm that embeds a back-fitting algorithm into alternating least squares, which avoids large matrices operation and big memory storage, and therefore makes it feasible to achieve scalable computing. Our simulation studies and MovieLens data analysis both indicate that the proposed group-specific method improves prediction accuracy significantly compared to existing competitive recommender system approaches.

Keywords: Cold-start problem, group-specific latent factors, non-random missing observations, personalized prediction

ICISE-080

BAYESIAN CALIBRATION OF INEXACT COMPUTER MODELS

Matthew Plumlee¹

¹University of Michigan
mplumlee@umich.edu

Abstract

Bayesian calibration is used to study computer models in the presence of both a calibration parameter and model bias. Using the predominant methodology, the parameter's posterior can drastically change depending on the bias's prior. This effect can lead to unreasonable inference on the parameter. To date, there has been no generally accepted alternatives. This paper proposes using Bayesian calibration where the prior distribution on the bias is orthogonal to the gradient of the computer model. Problems associated with Bayesian calibration are shown to be mitigated through analytic results and both numerical and real examples.

Keywords: Calibration

ICISE-081

**THE BUEHLER LOWER LIMITS ON SYSTEM RELIABILITY BASED ON THE
EXPERIMENT DATA**

Zhaohui Li¹, Senyao Du², Qingpei Hu³, Dan Yu⁴

¹Institute of Systems Science, Academia Sinica
lizhh@mail.ustc.edu.cn

²Institute of Systems Science, Academia Sinica
disneydsy@gmail.com

³Institute of Systems Science, Academia Sinica
qingpeihu@amss.ac.cn

⁴Institute of Systems Science, Academia Sinica
dyu@amss.ac.cn

Abstract

The confidence interval calculation framework by Buehler provides an exact approach to determine the lower limits for system reliability from components test data. In this study, an approach has been proposed to calculate the lower limits when the system component life times are distributed differently. As the component reliability point estimate only depends on its true value, this Buehler lower limits calculation is actually an optimization issue in a multi-dimensional hypercube. Simulation results show the better performance of the proposed approach over others.

Keywords: *Buehler Limits, System Reliability*

ICISE-084

WIND ENERGY FORECASTING BY DIFFUSION MODELS

Alain BENSOUSSAN¹, Alexandre BROUSTE²

¹ City University Hong Kong and University of Texas at Dallas
abensous@cityu.edu.hk

² Université du Maine
Alexandre.Brouste@univ-lemans.fr

Abstract

We model the evolution of wind speed as a Markov diffusion process, in continuous time. The driving element of the model is the long term (ergodic) behavior of the process. Practitioners assume commonly that the long term distribution is a Weibull distribution. Assuming that property, we derive formulas for the drift and the volatility terms and fix a diffusion process for the transitory behavior.

We then apply this model to forecast wind power energy on a short term horizon.

A similar methodology can be used to model the transitory behavior as a CIR diffusion (Cox-Ingersoll-Ross). The long term behavior is then a Gamma distribution.

We make comparison with more standard methods, like ARIMA, for forecasting purposes. We check that the diffusion models perform correctly, with the additional advantage that there is a knowledge model underneath the method, and not just a time series analysis.

Keywords: *Wind Energy*

ICISE-085

**EXPERIMENTAL DESIGN-BASED SUBAGGING FOR GAUSSIAN PROCESS
MODELS**

Ying Hung¹

¹Rutgers University
yhung@stat.rutgers.edu

Abstract

We study the problem of simultaneous variable selection and parameter estimation in Gaussian process models. Conventional penalized likelihood approaches are attractive but the computational cost of the penalized likelihood estimation (PMLE) or the corresponding one-step sparse estimation (OSE) can be prohibitively high as the sample size becomes large. This is because the likelihood function heavily involves operations of a covariance matrix of the same size as the number of observations. To address this issue, this article proposes an efficient subsample aggregating (subagging) approach with an experimental design-based subsampling scheme. The proposed method is computationally cheaper, yet it can be shown that the resulting subagging estimators achieve the same efficiency as the original PMLE and OSE asymptotically. The finite-sample performance is examined through simulation studies. Application of the proposed methodology to a data center thermal study reveals some interesting information, including identifying an efficient cooling mechanism.

Keywords: *Computer experiment, experimental design, variable selection*

ICISE-086

**ADDRESSING CURRENT CHALLENGES OF COMPUTER EXPERIMENTS IN
INDUSTRY**

William Myers¹

¹The Procter & Gamble Company
myers.wr@pg.com

Abstract

This talk will explore some recent research that addresses both design and modeling challenges for computer experiments in industry. Computer experiments provide a competitive advantage in industry where fast and cost effective product development is critical. In many industrial applications computer experiments are replacing physical experiments because the physical creation and testing of prototypes is very prohibitive in terms of time and cost. Computer experiments typically involve complex systems with numerous input variables. A primary goal in the application of computer experiments is to develop a metamodel – a good empirical approximation to the original complex computer model. This provides an easier and faster approach to sensitivity analysis, prediction and optimization. The talk will discuss actual industry applications.

Keywords: *Computer experiment, Space-filling design, Sliced latin hypercube*

ICISE-087

**DIAGNOSIS OF MULTI-SCALE SPATIAL POINT PATTERN INTERACTION
BASED ON DECOMPOSITION OF THE K FUNCTION-BASED T^2 STATISTIC**

Xiaohu Huang¹, Qiang Zhou²

¹ Department of Systems Engineering and Engineering Management
xhhuang6-c@my.cityu.edu.hk

² Department of Systems Engineering and Engineering Management
q.zhou@city.edu.hk

Abstract

Data in the form of spatial point distribution are commonly encountered in manufacturing processes such as nanoparticles in composite materials. By analyzing their distributional characteristics which are often related to product quality, we can monitor and diagnose their fabrication processes. Based on recent advancement on modeling the K function of point patterns using Gaussian process, this paper proposes to diagnose point patterns through decomposition of a K function-based T^2 statistic. The decomposition provides a novel way for independently analyzing spatial point interactions at multiple scales, which is particularly useful for fault diagnosis when the process is out-of-control. Effectiveness of the proposed method has been verified through several simulated examples and real data.

Keywords: *Spatial point pattern, the K function, Hotelling's T^2 control chart, MYT decomposition*

ICISE-088

**A MODIFIED LAWSON METHOD FOR DETECTING SIGNIFICANT EFFECTS
BASED ON A HALF-NORMAL PROBABILITY PLOT***

Jong Hee Chung¹, Yong Bin Lim²

¹Department of Statistics, Ewha Womans University
stachung@hanmail.net

²Department of Statistics, Ewha Womans University
yblim@ewha.ac.kr

Abstract

In analyzing data from unreplicated factorial designs, the half-normal probability plot is commonly used method to screen the vital few effects. Recently, many formal methods have been proposed to overcome the subjectivity on this plot. Among them, Lawson et al.(1998) suggested a hybrid method based on the half-normal probability plot, which is a blend of Lenth's (1989) and Loh's (1992) methods. The method consists of fitting a simple least squares line to the inliers, which are determined by Lenth. The effects exceeding the prediction limits based on the fitted line are judged to be significant. Improving the accuracy of partitioning the effects into inliers and outliers, we propose a modified method in which more outliers could be classified by using both methods, Carling's (2000) for adjusted boxplot and Lenth's method. If there exists no outlier or a wide range in the inliers determined by Lenth, we could obtain a chance to find more outliers by Carling method. A simulation study was conducted in unreplicated 2^4 designs with number of active effects being 2, 3, 4, 5 and 6. Letting Power I denote the proportion of detecting all active effects and Power II denote proportion of exactly detecting all active effects, the efficiency of the original Lawson method and the proposed method is compared by simulation study. It is shown that the proposed method is better than the original Lawson method. Also, we generated the table of critical values used in the proposed method.

Keywords: *detecting significant effects, half- normal probability plot, modified Lawson method*

**This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A2002032)*

ICISE-089

**A FRONT-END LUMPED-PARAMETER APPROACH TO PLANNING PHYSICAL
ROBUST ENGINEERING DESIGN EXPERIMENTS ON MANUFACTURED
DYNAMIC DEVICES**

Mark Atherton¹

¹ Brunel University London
mark.atherton@brunel.ac.uk

Abstract

This presentation will focus on how to use a lumped-parameter analysis to prioritise the design factors to be included in physical robust engineering design (RED) experiments on dynamic devices, particularly where the design factor values will be dictated by as-built assemblies selected from the production line.

It is shown that planning RED experiments on dynamic devices is helped by combining analysis of invariant groups, system equations and causal insight from bond graph models.

This analytical approach will inform the configuration of simulation models for RED on lumped-parameter dynamic systems in terms of output response and representation of manufacturing variability. In addition, it provides a framework for dealing with design factor levels that are difficult to configure as orthogonal design factor level design experiments due to their manufactured values being dictated by the particular assembly build.

Using invariant groups to identify dynamic similarity in the context of dynamic devices provides the RED practitioner with a means of conducting physical experiments with manufactured assemblies, which will also be of value when validating simulation models in RED experiments.

Bond graphs and dimensional analysis are the main analytical methods used to gain insight into the dynamic device. The system matrix of the device is formulated and used to assess the role of the design factors. Invariant groups are identified from dimensional analysis to determine a suitable output measure(s) for simulation and also to identify pairs of devices that have dynamic similarity. This front-end RED approach is applied to lumped-parameter dynamic systems.

Keywords: *Robust Engineering Design, Lumped-parameter model, Bond graph, System equations, Invariant groups, Parameter estimates*

ICISE-090

**SELECTING APPROPRIATE CONSTRAINTS FOR ALIGNMENT OF BATCH
PROCESS DATA WITH DYNAMIC TIME WARPING**

Max Spooner¹, Murat Kulahci²

¹Technical University of Denmark
mpsp@dtu.dk

²Technical University of Denmark
muku@dtu.dk

Abstract

Batch process data for I batches usually contain J variables collected over K time points. Often, the batches will have different durations, and so K will vary from batch to batch. However most standard methods of analysis assume the data to be a 3-dimensional array, $I \times J \times K$ with a fixed K . Even if all batches have the same duration, it is possible that any given event in the process may take place at varying time points from batch to batch. Therefore, prior to analysis the data for different batches need to be aligned. Many alignment methods have been proposed in the literature to address the issue of varying durations. But aligning the key events in different batches is still a challenging task. Dynamic Time Warping (DTW), a dynamic programming algorithm, can align the key events through stretching and compressing the local time dimension of a batch. This generates a warping function, which shows a progress signature for each batch and can be appended to the initial data for further analysis. Various constraints on the degree of warping can be imposed yet an in depth analysis on properly setting these constraints is missing in the literature. In this work, the various options within the DTW algorithm are examined and an objective method for selecting the most appropriate constraints is proposed. The goal of this method is to align the batches whilst avoiding pathological warpings in local batch time. This method is applied to real data from an industrial bacteria fermentation process.

Keywords: *Batch process, Alignment, Dynamic time warping, Fermentation*

ICISE-091

**FAST FLEXIBLE FILLING (FFF) DESIGNS FOR SPACE FILLING IN
CONSTRAINED REGIONS**

Bradley Jones¹

¹ JMP Division/SAS
bradley.jones@jmp.com

Abstract

FFF designs are space filling designs constructed using a speedy clustering algorithm on a random sample of feasible points. Recent improvements to the original approach are using the MaxPro criterion to choose a point from each cluster and the ability to add both qualitative and discrete numeric factors. This talk will demonstrate the capabilities of this design approach through several examples. One example will show the design points for a non-convex region. Another will compare FFF designs formed using the MaxPro criterion to the original centroid criterion. A final example will show FFF designs formed using a qualitative factor with multiple levels.

Keywords: *Space filling, MaxPro, Constrained region*

ICISE-092

**DESIGN OF CUSUM PROCEDURES WITH PROBABILITY CONTROL LIMITS
UNDER VARIABLE SAMPLE SIZES**

Wenpo Huang¹, Lianjie SHU², William H. Woodall³, Kwok-Leung Tsui⁴

¹Shang Hai Jiao Tong University
bobhuang09@gmail.com

²University of Macau
ljshu@umac.mo

³Virginia Tech
bwoodall@vt.edu

⁴City University of Hong Kong
klttsui@cityu.edu.hk

Abstract

Control charts are usually designed with constant control limits. In this paper, we consider the design of control charts with probability control limits aimed at controlling the conditional false alarm rate at the desired value at each time step. The resulting control limits would be dynamic, which are more general and capable of accommodating more complex situations in practice as compared to the use of constant control limit. We limit the discussion to the situation when the sample sizes are varying over time, with primary focus on the cumulative sum (CUSUM) type control charts. Unlike other methods, no assumptions about future sample sizes are required with our approach. An integral equation approach is developed to facilitate the design and analysis of the CUSUM control chart with probability control limits. The relationship between the CUSUM charts using probability control limits and the CUSUM charts with the fast initial response (FIR) feature is investigated.

Keywords: *Dynamic control limits*

ICISE-093

**MULTISCALE STATISTICS AND STOCHASTICS PROCESSES WITH
APPLICATIONS IN IMAGE ANALYSIS**

Anthony Joseph Yezzi¹, Stefano Barone²

¹ Georgia Institute of Technology, School of Electrical and Computer Engineering, USA
anthony.yezzi@ece.gatech.edu

² University of Palermo, Polytechnic School, Italy
stefano.barone@unipa.it

Abstract

Several image segmentation algorithms are formulated as iterative geometrically constrained pixel classification methods in which image intensity distributions are estimated separately within differently classified regions of the image domain. These regions are subsequently adjusted to increase some measure of difference between their resulting newly estimated intensity distributions. For complex images containing intricate texture, multi-scale structure, and/or noise, the resulting region intensity distribution estimates may be similarly complex. As such, simplification of the image via diffusion processes within each separately classified region may be useful in the context of these algorithms.

In fact, even richer information can be gathered about an image region by tracking the evolution of the empirical pixel intensity distribution during a diffusion process. Taking the linear heat equation as a test example, the resulting diffusion process will gradually simplify the image based on a model that treats the greyscale intensity as the temperature of the pixel, which changes over time due to the physical phenomenon of heat diffusion. At each iteration step, including the starting point (initial image), the empirical distribution of the pixel intensity can give valuable information, although it is a synthesis of the bi-dimensional function of the pixel intensity.

The empirical distribution of the pixel intensity at the generic iteration step can be modeled as a mixture of Gaussian random variables. Such Gaussian mixture evolves over time according to a stochastic process, which can be modeled too. The idea is that by estimating the parameters of this stochastic process, it is possible to attain richer sets of multiscale statistics with more discriminating difference measures between image regions in order to yield better image segmentations, especially in frequently occurring scenarios, where it does not make sense to model pixels within a given segmented region as independently, identically distributed random samples of a single underlying intensity distribution.

We further conjecture that, for a given stochastic process, the initial image is unique apart from domain transformations within the invariance group of the diffusion process. In other words, there is only one image that could generate that stochastic process. This idea will be documented through several case studies and the assumption of uniqueness will be supported by numerical simulations.

Keywords: *Multiscale statistics, Diffusions, Stochastic processes, Image analysis*

ICISE-094

**REAL-TIME PATH PLANNING FOR UNMANNED GROUND VEHICLES USING
MISSION PRIOR KNOWLEDGE**

Amir Sadrpour¹, Jionghua (Judy) Jin and Galip Ulsoy²

¹ Amazon, USA

jhjin@umich.edu

² The University of Michigan, USA

jhjin@umich.edu

Abstract

Surveillance missions that involve unmanned ground vehicles (UGVs) include situations where a UGV has to choose between alternative paths to complete its mission. Currently, UGV missions are often limited by the available on-board energy. Thus, we propose a dynamic most energy reliable path planning algorithm that integrates mission prior knowledge with real-time sensory information to identify the mission's most energy-reliable path. Our proposed approach predicts and updates the distribution of energy requirement of alternative paths using recursive Bayesian estimation through two stages: (1) exploration – road segments are explored to reduce their energy prediction uncertainty; (2) exploitation – the most reliable path is selected using the collected information in the exploration stage and is traversed. Our simulation results show that the proposed approach outperforms offline methods, as well as a method that only relies on exploitation to identify the most energy reliable path.

ICISE-095

**GEOMETRIC UNCERTAINTY IN THE ANALYSIS OF TURBINE BLADE
PERFORMANCE**

Ron Bates¹

¹Rolls-Royce plc
ron.bates@rolls-royce.com

Abstract

It is now possible to obtain very accurate measurements of as-manufactured component geometry using 3D structured light measurement systems. This presentation will explore how these measurements can be combined with finite element analysis models of components for Robust Design. The work will focus on the use of a bespoke mesh morphing technique to represent uncertainty in geometry as one aspect of a validation study for the modal analysis of a gas turbine compressor blade. Results will be presented that show the discrepancy between test and measurement can be explained at least in part by the variation in blade geometry.

Keywords: *Uncertainty, Robust Design, Geometry, Model Validation*

ICISE-096

**ON COMPARING SYSTEMS WITH HETEROGENEOUS COMPONENTS VIA
SURVIVAL SIGNATURES**

Francisco J. Samaniego¹

¹ University of California, Davis
fjsamaniego@ucdavis.edu

Abstract

The focus of this presentation is the comparison of coherent systems whose component lifetimes are independent but may vary in an arbitrary way. This work is based on the new concept of “survival signature” due to Coolen and Coolen-Maturi (2012) which extends the well-known concept of system signatures to the case of systems with exchangeable, heterogeneous components. The more restrictive assumption of the independence of component lifetimes permits certain representations of system reliability that lead to more detailed comparisons. Our results include a characterization of the survival signature of a monotone system of size n that is equivalent to a given system of size $m < n$. This allows for the direct comparison of systems of different sizes with varying component types. A variety of examples show that the approach can be productively applied to system comparisons.

This work is joint with Jorge Navarro, Universidad de Murcia.

ICISE-097

WHERE SHOULD STATISTICAL MONITORING GO?

Geoff Vining¹

¹ Virginia Tech
vining@vt.edu

Abstract

The origins of statistical process monitoring were extremely practical, identifying assignable causes. Banks (1993, Statistical Science, "Is Industrial Statistics Out of Control") was an extremely controversial paper; however, it did challenge whether statistical process monitoring (SPM) was still relevant and practical. The question becomes even more important as engineers develop truly big data approaches for inspection, which then form the opportunities for the next generation of SPM methodologies. This paper asks the question "Are the researchers in SPM ready for the task?" The goal of this talk is to nudge researchers in this area to move away from their comfort zones to address the real pressing questions that SPM now faces.

Keywords: *Big data, Future Directions*

ICISE-098

**PERFORMANCE EVALUATION OF SOCIAL NETWORK ANOMALY
DETECTION USING A MOVING WINDOW BASED SCAN METHOD**

**Meng J. Zhao¹, Anne Ryan Driscoll², Ronald D. Fricker, Jr.³, William H. Woodall⁴, Dan J.
Spitzner⁵**

¹ Virginia Tech
jmz167@vt.edu

² Virginia Tech
adriscoll@vt.edu

³ Virginia Tech
rf@vt.edu

⁴ Virginia Tech
bwoodall@vt.edu

⁵ University of Virginia
djs4y@eservices.virginia.edu

Abstract

Data generated by networks are increasingly readily available from a variety of sources, particularly social media and other types of the Internet-based applications and communications. As a result, there is significant interest in studying these types of network data, often with a focus on detecting anomalous events in social networks. A variety of methods have been proposed for monitoring such networks. However, research assessing the performance of these methods has been sparse. In this talk, I give an overview of basic social network ideas and applications. I describe suggested methods, particularly a popular scan-based method, for monitoring social network data and discuss advantages and limitations of these methods. This will be followed by a discussion of modifications for the methods to improve performance. Finally, the talk will end with ideas for future research in the area of social network monitoring.

Keywords: social network, monitoring, moving window, performance, simulation, sub-network

ICISE-099

**AN MCMC BASED SOLUTION TO THE QCD PROBLEM OF PARTICLE
PHYSICS**

Joshua Landon¹

¹George Washington University
jlandon@gwu.edu

Abstract

In this talk, we will start with an overview of the essentials of particle physics, for a general audience. The overview will be followed by the presentation of a system of equations developed by several Noble Prize winning physicists, known as the Lattice Quantum Chromodynamic Equations, abbreviated QCD. The Lattice QCD Equations endeavor to estimate the mass of a sub-atomic particle. They are notoriously difficult to solve because one encounters here the scenario of a finite number of equations, each equation the result of a physics based code, for estimating an infinite number of parameters. The physics based codes are time consuming and expensive to run; thus the finite number of equations. A simplifying assumption, namely, that of constant spacing, enables us to identify a telescopic pattern to these equations; this partly eases the burden of infinite dimensions. A statistical model can then be endowed on the simplified QCD equations, and a Marko Chain Monte Carlo based method can be invoked to perform the necessary estimation. This is joint work with Nozer Singpurwalla and Frank Lee, a particle physicist, and it is able to produce results which go beyond that which is currently available in the physics literatures. It has appeared in Statistical Science.

ICISE-100

SIMULATING RARE EVENTS AND TOLERANCE REGIONS

David Steinberg¹

¹Tel-Aviv University

dms@post.tau.ac.il

Abstract

There are many applications where it is important to find tolerance regions; i.e. regions that contain, with prescribed statistical confidence $1-\alpha$, a given fraction, say $1-\varepsilon$, of a probability distribution. This talk will describe methods for generating such regions when the probability distribution is the result of a complex computer model with probabilistic inputs. For the practical application that stimulated this work, we are interested in very small values of ε , on the order of 10^{-6} , so that brute force simulation is prohibitively expensive. We show that appropriate modeling of the computer model can dramatically reduce the computational expense involved in simulating these extreme events and in estimating the contour that describes the tolerance region.

This is joint work with Michael Gringauz, an M Sc student at Tel Aviv University.

Keywords: *Computer Experiments, Simulation, Rare Events*

ICISE-101

ACHIEVING NEW PERFECT CLASSIFICATION FOR FUNCTIONAL DATA

Aurore Delaigle¹

¹University of Melbourne
A.Delaigle@ms.unimelb.edu.au

Abstract

We consider the problem of classifying curves, or functions, in several groups, when a training sample of examples from those groups is available. It is standard to project the functions on a finite dimensional space before performing classification. Often this projection is made without having in mind the optimisation of classification performance. We show that, in the functional data context, there exists an optimal projection which often ensures very good, indeed sometimes perfect, classification performance. This is joint work with Peter Hall.

Keywords: *functional data*

ICISE-102

REAL-TIME PROCESS MONITORING USING KERNEL DISTANCES

Wei Jiang¹

¹ Shanghai Jiao Tong University
jiangwei@sjtu.edu.cn

Abstract

Real-time monitoring is an important task in process control. It often relies on estimation of process parameters in Phase I and Phase II and aims to identify significant differences between the estimates when triggering signals. Real-time contrast (RTC) control charts use classification methods to separate the Phase I and Phase II data and monitor the classification probabilities. However, since the classification probability statistics take discretely distributed values, the corresponding RTC charts become less efficient in the detection ability. In this paper, we propose to use distance-based RTC statistics for process monitoring, which are related to the distance from observations to the classification boundary. We illustrate our idea using the kernel linear discriminant analysis (KLDA) method and develop three distance-based KLDA statistics for RTC monitoring. The performance of the KL-DA distance-based charting methods is compared with the classification probability-based control charts. Our results indicate that the distance-based RTC charts are more efficient than the class of probability-based control charts.

ICISE-103

SEQUENTIAL TEST PLANNING FOR POLYMER COMPOSITES

I-Chen Lee¹, Yili Hong², Sheng-Tsaing Tseng³

¹ Institute of Statistics, National Tsing Hua University, Hsinchu, Taiwan
lictof@gmail.com

² Department of Statistics, Virginia Tech, Blacksburg, VA 24061
yilihong@vt.edu

³ Institute of Statistics, National Tsing Hua University, Hsinchu, Taiwan
sttseng@stat.nthu.edu.tw

Abstract

Polymer composite materials are widely used in areas such as aerospace and alternative energy industries, due to their lightweight and comparable levels of strength and endurance. To ensure the material can last long enough in the field, accelerated cyclic fatigue tests are commonly used to collect data and then make prediction for the field performance. Thus, a good testing strategy is desirable for evaluating the property of the polymer composites. While there has been a lot development in optimum test planning, most of the methods assume that the true parameter values are known. However, in reality, the true model parameters may depart from the planning values. In this paper, we propose a sequential strategy for test planning, and use Bayesian framework for the sequential model updating. We also use extensive simulation to evaluate the properties of the proposed sequential test planning strategy. Finally, we compare the proposed method to the traditional optimum design. Our results show that the proposed strategy is more robust and efficient when true values of parameters are unknown.

Keywords: *Sequential test planning, Fatigue testing*

ICISE-105

**A DECISION THEORETIC APPROACH TO STOCHASTIC LINEAR
PROGRAMMING**

Joshua Landon¹

¹George Washington University
jlandon@gwu.edu

Abstract

Linear programs are used to optimize a linear function, known as the objective function, subject to a set of linear constraints, which specify the feasible region. In this presentation we will address the issues encountered when trying to determine the optimal result when some or all of the linear functions are unknown. We will first consider the case when the objective function is unknown but the feasible region is known. This type of program is known as a DCRO (Deterministic Constraints, Random Objective Function) Linear Program. Our approach uses decision making techniques, assigning a loss function to each of the vertices of the feasible region, and defining the optimal choice as the vertex that minimizes the expected loss. We then consider the case when the constraints are unknown and the objective function is known. This type of program is known as a RCDO (Random Constraints, Deterministic Objective Function) Linear Program. With an unknown feasible region, our approach addresses the trade-off between choosing points which give larger values of the objective function, at the expense of a lower probability of being in the feasible region. We will finally address the last scenario, when both the objective function and the constraints are unknown. This type of program is known as a RCRO (Random Constraints, Random Objective Function) Linear Program. We will show how these types of problems can be solved using a combination of our methods for the DCRO and RCDO programs. This is joint work with Nozer Singpurwall.

ICISE-106

**MULTI-RESOLUTION INFERENCE: AN ENGINEERING (ENGINEERED?)
FOUNDATION OF STATISTICAL INFERENCE**

Xiao-Li Meng¹

¹Department of Statistics, Harvard University
meng@stat.harvard.edu

Abstract

One of the Big-Data buzz words is personalized medicine (which sounds heavenly!). But where on earth did they find enough guinea pigs to verify a treatment's efficacy for me? More generally, any (statistical) inference is a process of "transition to similar", namely, transferring our knowledge about a group of entities to a group of similar entities. As pondered by philosophers from Galen to Hume, how similar is similar? Wavelet-inspired Multi-resolution (MR) inference (Meng, 2014, COPSS 50th Anniversary Volume) allows us to theoretically frame this question, with the primary resolution defining the appropriate level of similarity. The search for the appropriate primary resolution is a quest for a sensible bias-variance trade-off: precise estimation of a less relevant treatment effect versus imprecise estimation of a more relevant treatment effect. In other words, which is better: the right answer to the wrong question or the wrong answer to the right one? The formalism afforded us by the MR framework reveals some (initially) counter-intuitive strategies for negotiating a satisfying bias-variance tradeoff (Liu and Meng, 2014, The American Statistician). A real-life Simpson's paradox from comparing kidney stone treatments will be used to engage the audience.

Keywords: *Bias-variance trade-off, Personalized treatment, Robust-Relevance trade-off, Simpson's paradox, Similarity, Wavelets*

ICISE-107

**QUALITY ENGINEERING FACES THE CHALLENGES OF BIG DATA AND
LITTLE DATA**

Fugee Tsung¹

¹The Hong Kong University of Science and Technology
ftsung@gmail.com

Abstract

This talk will present and discuss the challenges and opportunities that quality engineers face in the era of big data. The ability to separate signal and noise in the data-rich-information-poor environment would be the key.

The second part of the talk will present and discuss the challenges and opportunities that quality engineers face in the era of additive manufacturing (i.e., 3D printing), where there is little data due to its one-of-a-kind nature. For example, statistical quality control (SPC) originated from mass production cannot be applied directly because such a small or single lot production does not have repeated measures of the same kind.