TABLE OF CONTENTS

INTRODUCTION

Organizer and Host Organization

Department of Industrial Engineering, Hanyang University

Co-organizer

School of Data Science, 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 in engineering applications
- Data mining and machining learning
- Design and Analysis of Experiments, robust design
- Quality and reliability engineering, six sigma
- System informatics and control
- System operational control and decision making in data-rich environments
- Statistical data analysis with engineering models
- Smart and connected systems (e.g., manufacturing, healthcare, communities, cities)
- Analytics for advanced manufacturing, e.g, additive manufacturing, nano/bio manufacturing
- Industrial data analytics for Internet of Things
- Big data in industry
- Cyber-physical system security and data analytics
- Statistical risk analysis and engineering

WELCOME ADDRESS FROM CONFERENCE ORGANIZER

We are greatly honored to announce that the fifth International Conference of Interface between Statistics and Engineering (ICISE2019) will be held in Hanyang University, Seoul, during June 25-28, 2019.

After previous ICISE conferences in Beijing, China (2009), Tainan, Taiwan (2012), Hong Kong (2014), and Palermo, Italy (2016), ICISE has become highly recognized conference which provides the broad interface between statistics and engineering in the fields of complex system design and operation, quality and reliability improvement, and optimal proactive decision-making. To memorize 80th anniversary of Hanyang University and 60th anniversary of Department of Industrial Engineering, Hanyang University, we are greatly pleased to host the prestigious conference in our campus.

We hope that this conference will extend our understanding on the most recent innovations, trends, concerns, challenges, and solutions in statistics and engineering fields and provide a good opportunity for academic exchange and strengthening friendships among participants from different countries.

The local organizing committee sincerely welcomes all colleagues and friends to Seoul for this magnificent gathering. We truly hope all participants will have a great time in Seoul with a long history and splendid culture.

Suk Joo Bae
Chair of Local Organizing Committee

Professor
Department of Industrial Engineering

Hanyang University, Seoul, Korea

GREETINGS FROM CONFERENCE CO-CHAIR

On behalf of the School of Data Science (SDSC), I am pleased to extend my warmest greetings to everyone attending the fifth International Conference of Interface between Statistics and Engineering (ICISE2019) in Seoul.

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), Hong Kong (2014), and Palermo, Italy (2016) 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 three influential researchers as our keynote speakers, Professor William Q. Meeker (Iowa State University), Professor Yu Ding (Texas A&M University) and Professor Sung H. Park (Seoul National University). I am sure the conference participants will greatly benefit from their insightful speech and discussions.

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

Kwok L TSUI
Conference Co-Chair
Chair Professor
School of Data Science
City University of Hong Kong

CONFERENCE COMMITTEES

Advisory Committee

Chair:

Jianjun Shi	Georgia Institute of Technology

Member:

Vijay Nair	University of Michigan

Program Committee

Co-chairs:

Stefano Barone	University of Palermo
Kwok L Tsui	City University of Hong Kong
Shiyu Zhou	University of Wisconsin-Madison

Organizing Committee

Chair:

Suk Joo Bae	Hanyang University

Member:

Dong-Ho Lee	Hanyang University
In-Jae Jeong	Hanyang University
Tae-Bok Kim	Hanyang University
Kichun Lee	Hanyang University
Chuljin Park	Hanyang University
Seong Bum Kim	Korea University
Sung Won Han	Korea University
Yong-soo Kim	Kyonggi University
Seong-joon Kim	Chosun University
KS Chin	City University of Hong Kong
Louis Liu	City University of Hong Kong
Lolli Lee	City University of Hong Kong

Program Committee Members

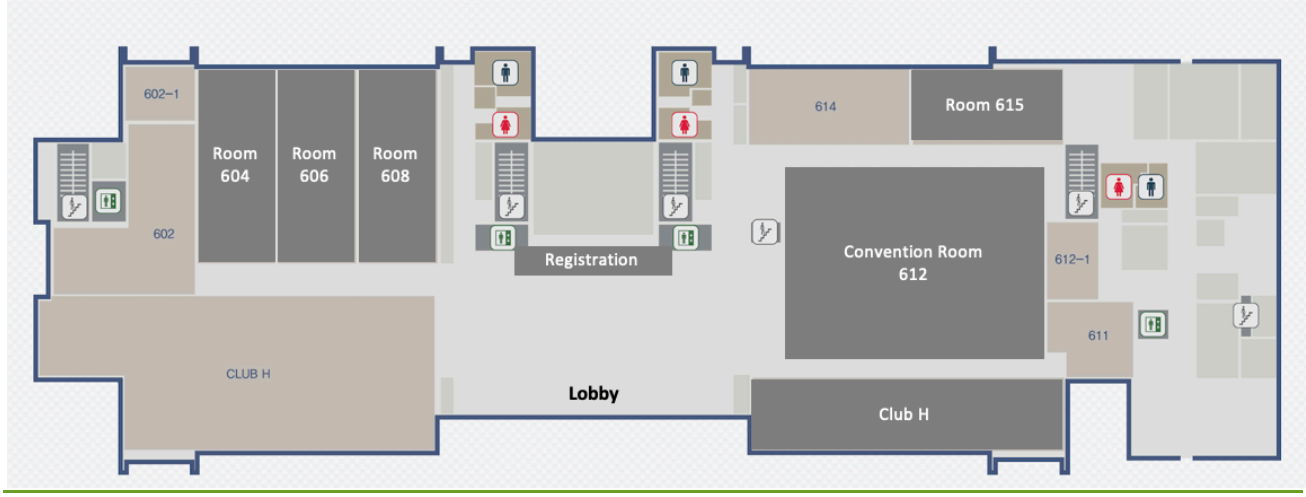
Suk Joo Bae	Hanyang University
Peter Chien	University of Wisconsin-Madison
Joo Ho Choi	Korea Aerospace University
David Coit	Rutgers University
Ronald Does	University of Amsterdam
Ying Hung	Rutgers University
Ching-Kang Ing	National Tsing Hua University
Myong-kee Jeong	Rutgers University
Wei Jiang	Shanghai Jiao Tong University
Seoung Bum Kim	Korea University
William Li	Shanghai Advanced Institute of Finance
Regina Liu	Rutgers University
Changsoon Park	SUNY Korea
Chiwoo Park	Florida State University
ST Tseng	National Tsinghua University
Loon Ching Tang	National University of Singapore
Hiroe Tsubaki	The National Statistics Center, JP
Fugee Tsung	Hong Kong University of Science and Technology
Roshan Vengazhiyil	Georgia Institute of Technology
Geoffrey Vining	Virginia Tech
Dave Woods	University of Southampton
Lifeng Xi	Shanghai Jiao Tong University
Dan Yu	Academy of Mathematics and Systems Science

CONFERENCE VENUES AND MAP

Registration

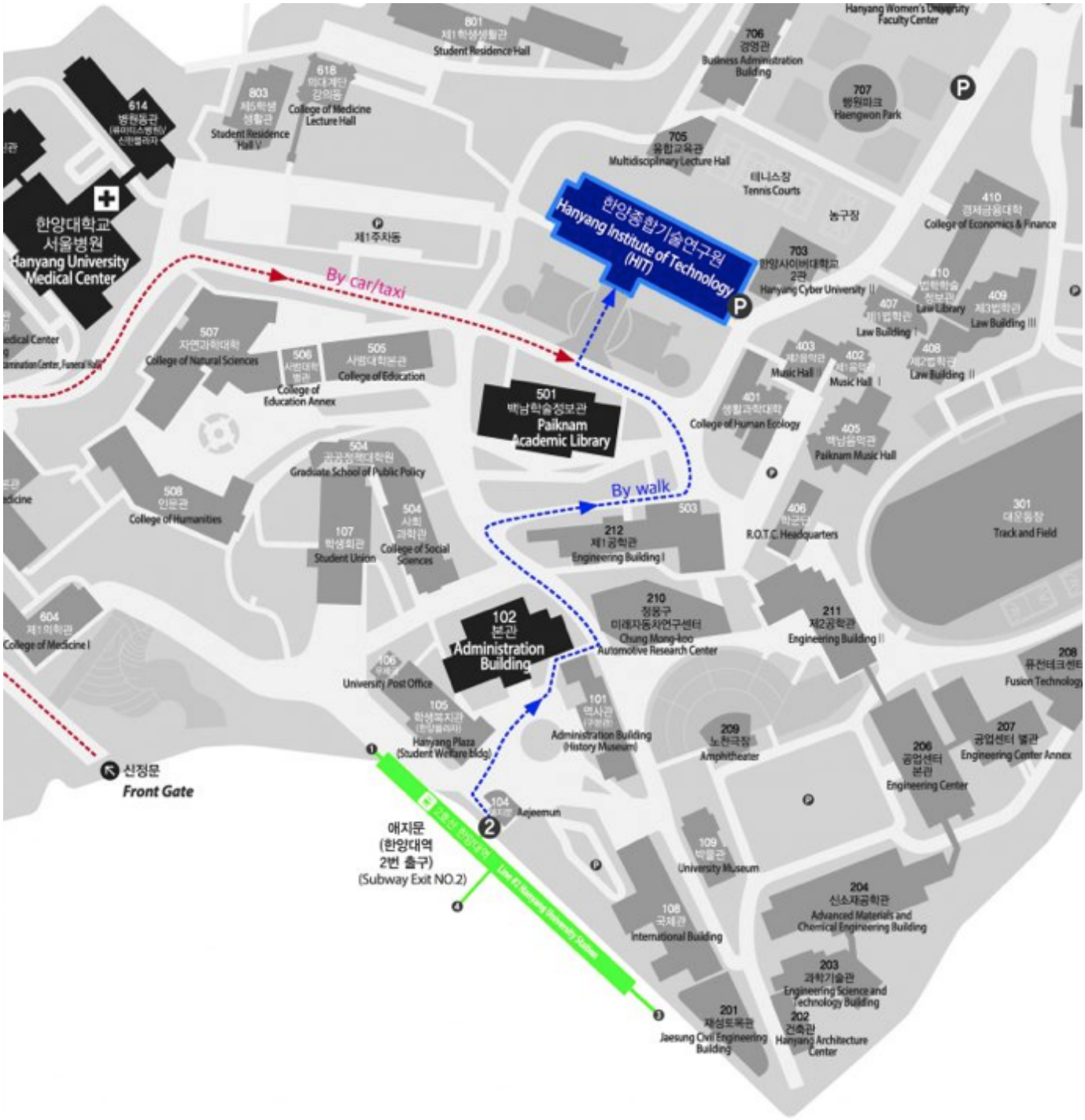
Lobby, 6/F, Hanyang Institute of Technology, Hanyang University, Seoul, South Korea

Conference Venues



Room No.	Conference Program
Room 615	Pre-Workshop(June 25) / Parallel Session(June 26~28)
Room 612(Convention Room)	Conference Opening/Keynote Speech
Room 604/606/608	Parallel Session
Club H	Welcome Reception / Lunch

Campus Map


한양대학교
서울병원
Hanyang University
Medical Center
801
제1학생생활관
Student Residence Hall
803
제5학생
생활관
Student Residence
Hall V
618
의대계단
강의동
College of Medicine
Lecture Hall
706
경영관
Business Administration
Building
707
행원파크
Haengwon Park
705
융합교육관
Multidisciplinary Lecture Hall
테니스장
Tennis Courts
농구장
410
경제금융대학
College of Economics & Finance
한양종합기술연구원
Hanyang Institute of Technology
(HIT)
703
한양사이버대학교
2관
Hanyang Cyber University II
제1주차동
By car/taxi
507
자연과학대학
College of Natural Sciences
506
사범대학
별관
College of
Education Annex
505
사범대학본관
College of Education
501
백남학술정보관
Paiknam
Academic Library
407
제1법학관
Law Building I
410
법학학술
정보관
Law Library
409
제3법학관
Law Building III
403
제2음악관
Music Hall II
402
제1음악관
Music Hall I
408
제2법학관
Law Building II
401
생활과학대학
College of Human Ecology
405
백남음악관
Paiknam Music Hall
504
공공정책대학원
Graduate School of Public Policy
508
인문관
College of Humanities
By walk
503
406
학군단
R.O.T.C Headquarters
301
대운동장
Track and Field
107
학생회관
Student Union
504
사회
과학관
College of Social
Sciences
212
제1공학관
Engineering Building I
210
정몽구
미래자동차연구센터
Chung Mong-koo
Automotive Research Center
211
제2공학관
Engineering Building II
604
제1의학관
College of Medicine I
102
본관
Administration
Building
106
우체
University Post Office
105
학생복지관
(한양플라자)
Hanyang Plaza
(Student Welfare Bldg)
101
역사관
(박물관)
Administration Building
(History Museum)
209
노천극장
Amphitheater
208
퓨전테크센터
Fusion Technology
207
공업센터 별관
Engineering Center Annex
206
공업센터
본관
Engineering Center
신정문
Front Gate
104
Aejeemun
애지문
(한양대역
2번 출구)
(Subway Exit NO.2)
109
박물관
University Museum
204
신소재공학관
Advanced Materials and
Chemical Engineering Building
108
국제관
International Building
203
과학기술관
Engineering Science and
Technology Building
201
재성토목관
Jaesung Civil Engineering
Building
202
건축관
Hanyang Architecture
Center

CONFERENCE PROGRAM

Time	Pre-workshop (June 25)	Day 1 (June 26)	Day 2 (June 27)	Day 3 (June 28)
08:30 – 09:00		**Registration**		
09:00 – 09:30		**Conference Opening**[3]	**Keynote Speech III**[3] *Prof. Sung H. Park*	**Parallel Session VI**
09:30 – 10:00		**Keynote Speech I**[3] *Prof. William Q. Meeker*		
10:00 – 10:30			Refreshment Break	
10:30 – 11:00		Refreshment Break	**Parallel Session III**	Refreshment Break
11:00 – 11:30		**Parallel Session I**		**Parallel Session VII**
11:30 – 12:00				
12:00 – 12:30				
12:30 – 13:00		**Lunch**[2]		
13:00 – 13:30				**Lunch**[2]
13:30 – 14:00				
14:00 – 14:30		**Keynote Speech II**[3] *Prof. Yu Ding*	**Parallel Session IV**	**City Tour** (Full Delegates Only)
14:30 – 15:00				
15:00 – 15:30	**Pre-workshop**[1] *Prof. Jianjun Shi* (Registered Only)			
15:30 – 16:00		**Refreshment Break**		
16:00 – 16:30		**Parallel Session II**	**Parallel Session V**	
16:30 – 17:00				
17:00 – 17:30				
17:30 – 18:00	**Welcome Reception**[2] (Full Delegates Only)			
18:00 – 18:30		Shuttle to Banquet		
18:30 – 19:00		**Banquet**[4] (Full Delegates Only)		
19:00 – 19:30				
19:30 – 20:00				

[1] Room 615, 6/F, Hanyang Institute of Technology, Hanyang University
[2] Club H, 6/F, Hanyang Institute of Technology, Hanyang University
[3] Convention Room, 6/F, Hanyang Institute of Technology, Hanyang University
[4] Brahms Hall, 19/F, Hotel President

KEYNOTE SPEECH I

TITLE: RELIABILITY IN THE 21ST CENTURY

Professor William Q. Meeker

Department of Statistics
Center for Nondestructive Evaluation
Iowa State University

About the Speaker

William Q. Meeker is Professor of Statistics and Distinguished Professor of Liberal Arts and Sciences at Iowa State University. He has more than 40 years of experience working in the application of statistical methods to engineering applications including reliability and nondestructive evaluation. He has done research and consulted extensively on problems in reliability data analysis, warranty analysis, experimental design, accelerated testing, nondestructive evaluation, and statistical computing. His practical experience includes numerous long-term visits to AT&T Bell Laboritories, General Electric Global Research, and Los Alamos National Laboratory. He is a Fellow of the American Statistical Association (ASA) the American Society for Quality (ASQ), and the American Association for the Advancement of Science, and a past Editor of Technometrics. He is co-author of the books Statistical Methods for Reliability Data with Luis Escobar (1998), the second edition of Statistical Intervals with Luis Escobar and Gerald Hahn (2017), 14 book chapters, and many publications in the engineering and statistical literature. He has won numerous awards for his research and contributions to the statistical and engineering professions including the ASQ Shewhart medal. ASA's 2015 Deming Lecture Award.

Abstract

Reliability is an engineering discipline that relies heavily on the application of probability and statistics. Changes in sensor, communications, and storage technologies are changing the nature of reliability field data. An increasing number of modern systems are being outfitted with sensors that capture information about how and when and under what environmental and operating conditions individual systems are being used. In some cases, the physical/chemical state of critical system components can also be quantified and reported. For many systems such information is being downloaded continuously into data farms. In addition, improvements in computing capabilities and investment in developing physics-based models for failure provide another important dimension of reliability information. There are many potential applications for using such data to improve safety and reduce costs but existing statistical methods for reliability assessment and prediction are inadequate for the tasks. This talk reviews some particular applications where the modern field reliability data are used and explores some of the opportunities to use modern reliability data to provide stronger statistical/physical methods that can be used to operate and predict the performance of systems in the field. We also provide some examples of recent technical developments designed to be used in such applications and outline remaining challenges.

Keywords: Condition-based maintenance, Dynamic covariates, Materials state awareness, Prognostics, Structural health monitoring

KEYNOTE SPEECH II

TITLE: CONSTRAINED STATISTICAL LEARNING FOR WIND TURBINE PERFORMANCE ASSESSMENT

Professor Yu Ding

Mike and Sugar Barnes Professor of Industrial & Systems Engineering
Professor of Electrical & Computer Engineering
Texas A&M University

About the Speaker

Dr. Yu Ding is the Mike and Sugar Barnes Professor of Industrial & Systems Engineering, Professor of Electrical & Computer Engineering, and a member of Texas A&M Institute of Data Science, Texas A&M Energy Institute, and TEES Institute of Manufacturing Systems. Dr. Ding received his Ph.D. degree from the University of Michigan in 2001. Dr. Ding's research interest is in the area of system informatics, and data and quality science. Dr. Ding is a recipient of the 2018 Texas A&M Engineering Research Impact Award, the recipient of the 2019 IISE Technical Innovation Award, and a Fellow of IISE and ASME.

Abstract

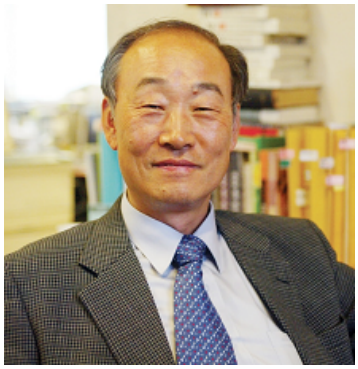
Data science has demonstrated its strength in the evaluation and prediction of system behavior and becomes an essential tool in engineering decision making. The success in system analysis, however, requires a deep understanding of engineering systems under consideration, in addition to the knowledge derived from data. This is because a certain system property, be it mechanical, chemical, or biological, could restrict a system's behavior but a purely data-driven approach overlooks such engineering constraints. In our research undertaking, we incorporate the pertinent yet imprecise engineering knowledge or physical constraints into a statistical learning process via imposing relevant shape/sign constraints on a system response function; the constraints include the requirements of positivity or negativity, monotonicity, convexity, or smoothness of the response function. Such constrained statistical learning, when applied to wind power production data, not only improves the estimation and prediction accuracy of the learning process, but enhances the interpretability of the resulting models and facilitates decisions for operations.

KEYNOTE SPEECH III

TITLE: THE QUALITY MANAGEMENT AND LEAN SIX SIGMA IN KOREA; WHERE WE TODAY AND WHERE ARE WE GOING?

Professor Sung H. Park
Emeritus Professor in Statistics
Seoul National University

President of Social Responsibility and
Management Quality Institute



About the Speaker

Professor Sung H. Park is an Emeritus Professor in Statistics at Seoul National University, and the president of Social Responsibility and Management Quality Institute. He was the president of Korean Academy of Science and Technology during 2013-2016, and a member of Presidential Advisory Council of Science and Technology in Korea during 2013-2015. He served as the director in the Directorate for Basic Research in Science and Engineering, National Research Foundation of Korea during 2010-2012.

He graduated from Seoul National University, Korea, in 1968 with a Bachelor of Science in Chemical Engineering. In 1970 he went to the USA to study Operations Research for his Master of Science Degree, and Statistics for his Ph.D. degree at North Carolina State University (NCSU). After graduating from NCSU in 1975, he went to Mississippi State University as an assistant professor in the business college, and then returned to his country, Korea, in 1977. Since 1977 he has served as an associate professor and then as a professor in statistics at Seoul National University (SNU). He has retired from SNU in 2010. He is the only one academician of IAQ (International Academy for Quality) from Korea since 2007.

He was the president of the Korean Society for Quality Management, the Korean Statistical Society as well as the Korean Society for Sustainability Science. He received two prestigious medals from Korean Government; one is 'Order of Service Merit, Red Stripes Medal' for his contribution to quality management in 2000, and another is 'Order of Science and Technology Merit, Hyeoksin Medal' for his contribution of Science and Technology for Korea in 2015. He also served as Dean of College of Natural Sciences, SNU, during 2000-2002, and as the chairman of Faculty Council of SNU during 2005-2007. He is a member of the National Academy of Sciences, Republic of Korea, which is very prestigious as a scholar.

Abstract

Korea was divided into two, South Korea and North Korea, when it became independent from Japan in 1945. In 1950's, (South) Korea was one of the poorest countries in the world. However, Korea now becomes one of the well living and advanced countries. There is no doubt that quality management has played an important role for Korea's development. In this paper, the past and present quality management policies and practices in Korea during the last 50 years are summarized and explained in detail. Also the introduction and development of Six Sigma, Lean and Lean Six Sigma is discussed.

However, recently Korea has encountered some challenging issues such as low economic growth, low industrial innovation, tough global competition, low corporate social responsibility practice, and rapid change of company environment. Now we are in the middle of the 4th Industrial Revolution, and

Korea should seek some breakthrough to solve the challenging issues and should establish a new paradigm of quality management and Lean Six Sigma. Some possible solutions for the issues are suggested in this paper. There is no doubt that this rapidly changing society needs a new paradigm for overall quality management as well as Lean Six Sigma.

PRE-WORKSHOP SEMINAR

TITLE: HOW TO PUBLISH IN TOP JOURNALS?

Professor Jianjun Shi
Editor in Chief, IISE Transactions
Member, US National Academy of Engineering
The Carolyn J. Stewart Chair and Professor
School of Industrial and Systems Engineering
Georgia Institute of Technology

About the Speaker

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

Dr. Shi's research is in the area of data enabled manufacturing. His methodologies integrate system informatics, advanced statistics, and control theory, and fuse engineering systems models with data science methods for design and operational improvements of manufacturing and service systems. The technologies developed by Dr. Shi's research group have been widely implemented in a wide variety of production systems and produced significant economic impacts. Dr. Shi is the founding chair of the Quality, Statistics and Reliability (QSR) Subdivision at INFORMS, and currently serving as the Editor-in-Chief of the IISE Transactions, the flagship journal of the Institute of Industrial and Systems Engineers (IISE). He is a Fellow of IISE, ASME, and INFORMS, an Academician of the International Academy for Quality, and a member of National of Academy of Engineering (NAE).

Abstract

This presentation will discuss how to identify promising research topics, and how to publish papers in top Journals. Publishing in top international journals is important to many researchers in educational and research institutions. This presentation will discuss various topics important for research and paper writing, including

- *How to judge if a journal is among the top ranking?*
- *How to select and define a research topic?*
- *How to organize and write a paper?*
- *What is the reviewing and publishing process for a paper?*

The presentation will highlight the "dos and don'ts". In addition, a "methodology tree" strategy will be introduced based on my experience of research and supervision of Ph.D. students. The "methodology tree" is a systematic strategy that can be used to organize the literature reviews, evaluate contributions of a research topic, facilitate communications among researchers, and provide a big picture of research problems involved. The workshop should help scholars to reflect on their own experiences as researchers, authors, and reviewers. It will be interactive and tailored to the interests of those attending. Do bring along your questions and comments.

PARALLEL SESSION SCHEDULE

WEDNESDAY, JUNE 26(DAY 1)

Session Room	**Room 604**	**Room 606**	**Room 608**	**Room 615**
08:00 –	Registration			
09:15 –09:30	Conference Opening Remarks			
09:30 – 10:30	Keynote Speech I **"Reliability in the 21st Century"** Professor William Q. Meeker(Iowa State University)			
10:30 – 11:00	Refreshment Break			
11:00 – 12:30	*S1* *Recent Developments in Statistical Process Monitoring [Ronald Does]*	*S4* *Reliability and Monitoring of Quality [Changsoon Park]*	*S5* *Data Science for Engineering Applications [Chiwoo Park]*	*S6* *Statistical Inferences for Reliability Data [Loon Ching Tang]*
11:00 – 11:30	ICISE-139	ICISE-111	ICISE-121	ICISE-202
11:30 – 12:00	ICISE-115	ICISE-131	ICISE-108	ICISE-169
12:00 –12:30	ICISE-138	ICISE-128	ICISE-114	ICISE-194
12:30 – 14:00	Lunch			
14:00 – 15:00	Keynote Speech II **"Constrained Statistical Learning for Wind Turbine Performance Assessment"** Professor Yu Ding(Texas A&M University)			
15:00 – 15:30	Refreshment Break			
15:30 – 17:30	*S2* *Advanced Statistical Methods for Process Improvement [Byunghoon Kim/ Youngseon Jeong/ Myong-kee Jeong*	*S16* *Bayesian Approaches in the Engineering Reliability Problems [Joo Ho Choi]*	*S18* *Application of Deep Learning and Network Analysis [Sung Won Han]*	*S19* *Prognostics and Health Management (PHM) for Industrial Applications [Daeil Kwon]*
15:30 – 16:00	*ICISE-187*	*ICISE-192*	*ICISE-171*	*ICISE-185*
16:00 – 16:30	*ICISE-208*	*ICISE-197*	*ICISE-137*	*ICISE-198*
16:30 – 17:00	*ICISE-183*	*ICISE-174*	*ICISE-155*	*ICISE-200*
17:00 – 17:30	*ICISE-113*	*ICISE-173*	*ICISE-196*	*ICISE-178*
17:30	Shuttle Bus to Banquet			
18:30 – 20:30	Banquet			

JUNE 27, 2019 (DAY 2)

Session Room	**Room 604**	**Room 606**	**Room 608**	**Room 615**
08:00 –	Registration			
09:00 –10:00	Keynote Speech III **"The Quality Management and Lean Six Sigma in Korea; Where We Today and Where Are We Going?"** Professor Sung H. Park(Seoul National University)			
10:00 –10:30	Refreshment Break			
10:30 – 12:30	*S12 Statistical Learning and Inference for Complex Industrial Systems I [Qingpei Hu]*	*S14 Inference in Complex Data Settings [Matthias Tan/ Regina Liu]*	*S17 Statistical Modeling for Complex Systems [Ying Hung]*	*S24 Matrix Completion Application [Zhensong Chen]*
10:30 – 11:00	ICISE-120	ICISE-162	ICISE-188	ICISE-151
11:00 – 11:30	ICISE-166	ICISE-161	ICISE-189	ICISE-176
11:30 – 12:00	ICISE-142	ICISE-136	ICISE-207	ICISE-190
11:00 – 12:30	ICISE-167	----	ICISE-170	ICISE-206
12:30 – 14:00	Lunch			
14:00– 15:30	S3 SPC Theory and Applications [Inez Zwetsloot/ Wei Jiang]	S7 Interface between Statistics and New Technology Development [Hiroe Tsubaki]	S9 Computer Experiments [Tony Ng/ Roshan Vengazhiyil]	S11 Design of Experiments [Yang Zhao/ Dave Woods]
14:00 – 14:30	ICISE-130	ICISE-165	ICISE-179	ICISE-203
14:30 – 15:00	ICISE-117	ICISE-119	ICISE-153	ICISE-109
15:00 – 15:30	ICISE-110	ICISE-112	ICISE-180	ICISE-133
15:30 – 16:00	Refreshment Break			
16:00– 18:00	S10 Current Issues in the Design of Experiments [Geoffrey Vining]	S13 Statistical Learning and Inference for Complex Industrial Systems II [Jian Liu]	S15 New Development in Design and Analysis of Computer Experiments [Dong Wang/ Peter Chien]	S22 Predictive Model and Machine Learning [Fangfang Yang]
16:00 – 16:30	ICISE-181	ICISE-152	ICISE-201	ICISE-157
16:30 – 17:00	ICISE-140	ICISE-126	ICISE-149	ICISE-160
17:00 – 17:30	ICISE-141	ICISE-168	ICISE-177	ICISE-193
17:30 – 18:00	----	ICISE-132	ICISE-163	

JUNE 28, 2019 (DAY 3)

Session Room	**Room 604**	**Room 606**	**Room 608**	
08:00 –	Registration			
09:00 – 10:30	*S21 Bayesian Inference Applications [Seong-joon Kim]*	*S23 Statistics in General Engineering Applications [Ji Zhu]*	*S25 Applications of Data Driven Modeling [Si-il Sung]*	
09:00 – 09:30	ICISE-122	ICISE-205	ICISE-125	
09:30 – 10:00	ICISE-127	ICISE-158	ICISE-164	
10:00 – 10:30	ICISE-172	ICISE-191	ICISE-184	
10:30 – 11:00	Refreshment Break			
11:00 – 13:00	*S8 Recent Advances in Time Series and High-dimensional Data Analysis [Ching-Kang Ing]*	*S20 Statistics and Analytics [Sungil Kim]*		
11:00 –11:30	ICISE-144	ICISE-143		
11:30 –12:00	ICISE-146	ICISE-195		
12:00 –12:30	ICISE-148	ICISE-186		
12:30 –13:00	ICISE-147			
13:00 –14:00	Lunch			
14:00 –18:00	City Tour			

TECHNICAL PROGRAM

Wednesday, June 26(Day 1)

Convention Room

09:15 –09:30	Conference Opening Remarks
09:30 –10:30	Keynote Speech I **"Reliability in the 21st Century"** Professor William Q. Meeker(Iowa State University)
14:00 –15:00	Keynote Speech II **"Constrained Statistical Learning for Wind Turbine Performance Assessment"** Professor Yu Ding(Texas A&M University)

Room 604

S1 Recent Developments in Statistical Process Monitoring [Ronald Does] 11:00-12:30

11:00 –11:30	ICISE-139	NONPARAMETRIC CONTROL OF THE CONDITIONAL PERFORMANCE IN STATISTICAL PROCESS MONITORING	Rob Goedhart	University of Amsterdam
11:30 –12:00	ICISE-115	AGGREGATION STRATEGIES IN STATISTICAL PROCESS MONITORING	Inez M. Zwetsloot	City University of Hong Kong
12:00 –12:30	ICISE-138	THE EFFECT OF CONTINUOUSLY UPDATING CONTROL CHART LIMITS ON CONTROL CHART PERFORMANCE	Leo C.E. Huberts	University of Amsterdam

S2 Advanced Statistical Methods for Process Improvement 15:30-17:30
[Byunghoon Kim/ Youngseon Jeong/ Myong-kee Jeong]

15:30 –16:00	ICISE-187	CLASS DESCRIPTION NETWORK	Hyungrok Do	Korea University
16:00 –16:30	ICISE-208	SEQUENTIAL DESIGN OF EXPERIMENTS FOR MULTIRESPONSE SURFACE OPTIMIZATION	Donghee Lee	Hanyang University
16:30 –17:00	ICISE-183	A NEW UNCERTAIN DECISION TREE METHOD FOR CLASSIFYING SEMICONDUCTOR WAFER DEFECT BASED ON MULTIPLE WAFER MAPS	Byunghoon Kim	Hanyang University
17:00 –17:30	ICISE-113	TRANSFER LEARNING OF GRAPHICAL STRUCTURES FOR CONNECT MANUFACTURING SYSTEMS	Shiyu Zhou	University of Wisconsin Madison

Room 606

S4 Reliability and Monitoring of Quality [Changsoon Park] 11:00-12:30

11:00 –11:30	ICISE-111	QUALITY BIG DATA	Fugee Tsung	Hong Kong University of Science and Technology

11:30 –12:00	ICISE-131	GENERAL PATH MODELS FOR DEGRADATION DATA WITH MULTIPLE CHARACTERISTICS AND COVARIATES	Yili Hong	Virginia Tech
12:00 –12:30	ICISE-128	MODIFICATION OF THE GLR CHART STATISTICS FOR MONITORING COUNT PROCESSES	Jaeheon Lee	Chung-Ang University

S16 Bayesian Approaches in the Engineering Reliability Problems [Joo Ho Choi] 15:30-17:30

15:30 –16:00	ICISE-192	BAYESIAN NETWORK FOR PROGNOSTICS OF MULTI-COMPONENT SYSTEMS	Seokgoo Kim	Korea Aerospace University
16:00 –16:30	ICISE-197	BAYESIAN UPDATING METHODS FOR LIFE PREDICTION USING FIELD FAILURES DATA	Joo Ho Choi	Korea Aerospace University
16:30 –17:00	ICISE-174	STATISTICAL UNCERTAINTY MODELING AND RELIABILITY ASSESSMENT USING BAYESIAN NEURAL NETWORKS IN RELIABILITY-BASED DESIGN OPTIMIZATION	Juhee Lim	Yonsei University
17:00 –17:30	ICISE-173	PROBABILISTIC WEAR LIFE-CENTERED MAINTENANCE DESIGN OF AUTOMATION SYSTEM WHEELS	Yeongmin Yoo	Yonsei University

Room 608

S5 Data Science for Engineering Applications [Chiwoo Park] 11:00-12:30

11:00 –11:30	ICISE-121	ELECTRICITY DEMAND FORECASTING WITH DETECTING AND PREDICTING CHANGES	Young Myoung Ko	Pohang University of Science and Technology
11:30 –12:00	ICISE-108	A STATISTICS-GUIDED APPROACH TO QUALITY CHARACTERIZATION AND COMPENSATORY DESIGN FOR 3D PRINTING	Kaibo Wang	Tsinghua University
12:00 –12:30	ICISE-114	DESIGN OPTIMIZATION FOR ESTIMATING DISCONTINUOUS REGRESSION FUNCTIONS	Chiwoo Park	Florida State University

S18 Application of Deep Learning and Network Analysis [Sung Won Han] 15:30-17:30

15:30 –16:00	ICISE-171	INTEGRATIVE DEEP LEARNING FOR IDENTIFYING DIFFERENTIALLY EXPRESSED(DE) BIOMARKERS	SungHwan Kim	Konkuk University
16:00 –16:30	ICISE-137	THE EFFECT OF AMGRU ON THE TIME-SERIES ANALYSIS	Joo Heui Son	Korea University
16:30 –17:00	ICISE-155	THE ESTIMATION OF RISK-GENE NETWORKS AND THE PREDICTION OF SURVIVAL RISKS	Minhyeok Lee	Korea University

17:00 –17:30	ICISE-196	MACHINE LEARNING ALGORITHM FOR SUBSTATION OF CHEMICAL REACTION MODEL	Jong Ho Shin	Chosun University

Room 615

S6 Statistical Inferences for Reliability Data [Loon Ching Tang] 11:00-12:30

11:00 –11:30	ICISE-202	RESILIENCE ASSESSMENT OF POWER SYSTEMS USING TREND TESTS	Lijuan Shen	National University of Singapore
11:30 –12:00	ICISE-169	DEGRADATION IN COMMON DYNAMIC ENVIRONMENTS	Zhisheng Ye	National University of Singapore
12:00 –12:30	ICISE-194	ON A ROBUST STATISTIC FOR SCALE AND ITS APPLICATION IN ENGINEERING	Hideki Nagatsuka	Chuo University

S19 Prognostics and Health Management (PHM) for Industrial Applications [Daeil Kwon] 15:30-17:30

15:30 –16:00	ICISE-185	A CNN-BASED FAULT DETECTION METHOD USING VIBRATION VIDEO	Bayu Adhi Tama	Pohang University of Science and Technology
16:00 –16:30	ICISE-198	SOH-FLUCTUATION ANALYSIS FOR EARLY DETECTION OF UNHEALTHY LI-ION BATTERIES: A RANDOM FOREST APPROACH	Changyong Lee	Ulsan National Institute of Science and Technology
16:30 –17:00	ICISE-200	FATIGUE LIFE SIMULATION OF PTH SOLDER JOINTS CONSIDERING ALEATORY UNCERTAINTY	Hyunseok Oh	Gwangju Institute of Science and Technology
17:00 –17:30	ICISE-178	PHM FRAMEWORK FOR SMART MANUFACTURING SYSTEMS	Daeil Kwon	Konkuk University

Thursday, June 27(Day 2)

Convention Room

09:30 –10:30 Keynote Speech III

"The Quality Management and Lean Six Sigma in Korea; Where We Today and Where Are We Going?"
Professor Sung H. Park(Seoul National University)

Room 604

S12 Statistical Learning and Inference for Complex Industrial Systems I [Qingpei Hu] 10:30-12:30

10:30 –11:00	ICISE-120	STATISTICAL INFERENCE FOR Mt/G/INFINITY QUEUEING SYSTEMS WITH INCOMPLETE INFORMATION	Dongmin Li	Chinese Academy of Sciences

11:00 –11:30	ICISE-166	DEGRADATION MODELING BASED ON A WIENER PROCESS WITH RANDOM DEGRADATION RATE AND INITIAL VALUE	Hongyu Wang	Beihang University
11:30 –12:00	ICISE-142	PERSONALIZED LEARNING SYSTEM FOR SMART TRANSPORTATION DEMAND MANAGEMENT (TDM)	Jingshuo Feng/ Shuai Huang	University of Washington
11:00 –12:30	ICISE-167	THE FOURTH INDUSTRIAL REVOLUTION AND A NEW PARADIGM: "RESILIENCE" TO A HYPER CONNECTED SOCIETY	Eui Hoon Lee	Chungbuk National University

S3 SPC Theory and Applications [Inez Zwetsloot/ Wei Jiang] 14:00-15:30

14:00 –14:30	ICISE-130	SPATIO-TEMPORAL SEQUENCE PREDICTION BASED ON SINGULAR VALUE DECOMPOSITION AND ARIMA	Lining Yang/ Yanting Li	Shanghai Jiao Tong University
14:30 –15:00	ICISE-117	A SPARSE LEADING-EIGENVALUE-DRIVEN CONTROL CHART FOR PHASE I ANALYSIS OF HIGH-DIMENSIONAL COVARIANCE MATRICES	Lianjie Shu	University of Macau
15:00 –15:30	ICISE-110	INDIVIDUALIZED DEGRADATION MODELING AND PROGNOSTICS IN A HETEROGENEOUS GROUP VIA MODELING COVARIATE INTERRELATIONS AMONG UNITS	Minhee Kim	University of Wisconsin Madison

S10 Current Issues in the Design of Experiments [Geoffrey Vining] 16:00-18:30

16:00 –16:30	ICISE-181	SIMULTANEOUS OPTIMIZATION OF CORRELATED QUALITY AND LIFETIME CHARACTERISTICS IN CONSTRAINED RANDOMIZATION EXPERIMENTS	Shanshan Lv	Hebei University of Technology
16:30 –17:00	ICISE-140	DATA TRANSFORMATIONS OR THE GENERALIZED LINEAR MODEL? SOME INSIGHTS AND RECOMMENDATIONS	Marcus Perry	University of Alabama
17:00 –17:30	ICISE-141	BALANCING ROBUST CLASSICAL DESIGNS WITH DESIGN CONSTRAINTS – APPLICATIONS IN TESTING DEFENSE SYSTEMS	Laura Freeman	Virginia Tech

Room 606

S14 Inference in Complex Data Settings [Matthias Tan/ Regina Liu] 10:30-12:30

10:30 –11:00	ICISE-162	ADDITIVE FUNCTIONAL REGRESSION WITH DENSITIES AS RESPONSES	Byeong Uk Park	Seoul National University

11:00 –11:30	ICISE-161	ARE REPORTED LIKELIHOOD RATIOS WELL CALIBRATED?	Jan Hannig	University of North Carolina at Chapel Hill
11:30 –12:00	ICISE-136	STATISTICAL INFERENCE COMBINED WITH PERSISTENT HOMOLOGY FOR PREDICTING FLUID FLOW IN POROUS MEDIA	Chul Moon	Southern Methodist University

S7 Interface between Statistics and New Technology Development [Hiroe Tsubaki] 14:00-15:30

14:00 –14:30	ICISE-165	HYPER DESIGN AND COMPREHENSIVE DESIGN METHOD - DESIGN CONCEPT AND APPLICATION TO TWIN ROTOR PAPER HELICOPTER	Takenori Takahashi	Keio University
14:30 –15:00	ICISE-119	ISO TC69/SC8: STANDARDIZING METHODOLOGIES FOR NEW TECHNOLOGY AND PRODUCT	Watalu Yamamoto	University of Electro-Communications
15:00 –15:30	ICISE-112	THE MAHALANOBIS-TAGUCHI SYSTEM BASED ON STATISTICAL MODELING	Masato Ohkubo	Toyo University

S13 Statistical Learning and Inference for Complex Industrial Systems II [Jian Liu] 16:00-18:00

16:00 –16:30	ICISE-152	DETECTING BURSTS IN WATER PIPE SYSTEM USING FOURIER BASIS EXPANSION	Jian Liu	The University of Arizona
16:30 –17:00	ICISE-126	OPPORTUNISTIC MAINTENANCE POLICY FOR MULTI-UNIT SYSTEM UNDER MARKOVIAN ENVIRONMENT CHANGES	Yu Odajima	Univerisity of Electro-Communications
17:00 –17:30	ICISE-168	MULTIOBJECTIVE METER PLACEMENT MODEL FOR ROBUST WATER DISTRIBUTION SYSTEM PIPE BURST DETECTION	Donghwi Jung	Keimyung University
17:30 –18:00	ICISE-132	PHASE-TYPE DISTRIBUTIONS FOR PRODUCT RETURN DATA WITH TWO-LAYER CENSORING	Yudong Wang	National University of Singapore

Room 608

S17 Statistical Modeling for Complex Systems [Ying Hung] 10:30-12:30

10:30 –11:00	ICISE-188	GAUSSIAN PROCESS MODELS FOR COMPUTER EXPERIMENTS WITH NON-QUANTITATIVE INPUTS	Xinwei Deng	Virginia Tech
11:00 –11:30	ICISE-189	MODELING IN-PLANE DEVIATIONS OF SHAPES TO COME BASED ON PRIOR DEVIATION FEATURES IN ADDITIVE MANUFACTURING SYSTEMS	Arman Sabbaghi	Purdue University
11:30 –12:00	ICISE-180	MULTI-RESOLUTION FUNCTIONAL ANOVA FOR LARGE-SCALE, MANY-INPUT COMPUTER EXPERIMENTS	Chih Li Sung	Michigan State University
11:00 –12:30	ICISE-170	ROBUST ESTIMATION OF COMPONENT RELIABILITY BASED ON SYSTEM LIFETIME DATA	Hon Keung Tony Ng	Southern Methodist University

S9 Computer Experiments [Tony Ng/ Roshan Vengazhiyil] 14:00-15:30

14:00 –14:30	ICISE-179	COMPUTER EXPERIMENTS WITH BINARY TIME SERIES WITH APPLICATIONS IN CELL BIOLOGY: MODELING, ESTIMATION, AND CALIBRATION	Ying Hung	Rutgers University
14:30 –15:00	ICISE-153	SLICED LATIN HYPERCUBE DESIGNS WITH ARBITRARY RUN SIZES	Xu He	Chinese Academy of Sciences
15:00 –15:30	ICISE-207	POWER LAW PROCESS WITH LEFT-TRUNCATED AND RIGHT-TRUNCATED DATA FOR MULTIPLE REPAIRABLE SYSTEMS	Byeong Min Mun	Hanyang University

S15 New Development in Design and Analysis of Computer Experiments [Dong Wang/ Peter Chien] 16:00-18:00

16:00 –16:30	ICISE-201	ESTIMATION OF HEALTHCARE ACCESSIBILITY USING AN ONLINE EXPERIMENT	Youngdeok Hwang	Sungkyunkwan University
16:30 –17:00	ICISE-149	EFFICIENT INPUT UNCERTAINTY QUANTIFICATION VIA GREEN SIMULATION USING SAMPLE-PATH LIKELIHOOD RATIOS	Eunhye Song	Pennsylvania State University
17:00 –17:30	ICISE-177	THE RECONSTRUCTION APPROACH: FROM INTERPOLATION TO REGRESSION, CLASSIFICATION, AND NUMERICAL COMPUTING	Shifeng Xiong	Chinese Academy of Sciences
17:30 –18:00	ICISE-163	ASSESSING ELDERLY'S FUNCTIONAL BALANCE AND MOBILITY VIA ANALYZING DATA FROM WEARABLE ACCELEROMETER IN TIME UP AND GO TESTS	Lisha Yu	City University of Hong Kong

Room 615

S24 Matrix Completion Application [Zhensong Chen] 10:30-12:30

10:30 –11:00	ICISE-151	TRADE-IN STRATEGIES IN DUAL-CHANNEL CLOSED-LOOP SUPPLY CHAIN WITH REMANUFACTURING	Lu Xiao	Guangdong University of Technology
11:00 –11:30	ICISE-176	REPEATED PARAMETER ESTIMATION WITH CONTROLLED RANDOM SEARCH FOR AN EXPENSIVE AND HIGHLY NONLINEAR MODEL	Hyungjin Kim	Hanyang University
11:30 –12:00	ICISE-190	FEASBLE DESIGN REGION IDENTIFICATION IN ADDITIVE MANUFACTURING VIA SURROGATE MODELLING OF DESIGN RULES	SungKu Kang	Virginia Polytechnic Institute and State University
11:00 –12:30	ICISE-206	AN ITERATIVE APPROACH FOR FIXTURE ALLOCATION AND SCHEDULING FOR A FLEXIBLE MACHINING SYSTEM WITH MULTI-FIXTURING PALLETS AND CONTROLLABLE PROCESSING TIMES	Dong-Ho Lee	Hanyang University

S11 Design of Experiments [Yang Zhao/ Dave Woods] 14:00-15:30

14:00 –14:30	ICISE-203	A CONSTRUCTION OF COST-EFFICIENT DESIGNS WITH GUARANTEED REPEATED MEASUREMENTS ON INTERACTION EFFECTS	Frederick Phoa	Academia Sinica
14:30 –15:00	ICISE-109	GAUSSIAN PROCESS MODELING USING INFORMATION FROM PARTIAL DIFFERENTIAL EQUATION MODELS	Matthias Tan	City University of Hong Kong
15:00 –15:30	ICISE-133	ACCELERATED DEGRADATION RELIABILITY DEMONSTRATION TEST PLAN USING DIRECT PREDICTION METHOD	Seongjoon Kim	Chosun University

S22 Predictive Model and Machine Learning [Fangfang Yang] 16:00-18:00

16:00 –16:30	ICISE-157	PREDICTIVE MODELS FOR SHORT-TERM PASSENGER FLOW FORECASTING IN METRO	Yang Zhao	City University of Hong Kong
16:30 –17:00	ICISE-160	A MODIFIED LATIN HYPERCUBE SAMPLING BASED ON PRIOR INFORMATION	Xueqing Wang	Tianjin University
17:00 –17:30	ICISE-193	COMBINING EXTREME LEARNING MACHINE, SPARSE AUTO-ENCODER AND BACK PROPAGATION FOR HEALTH INDICATOR EXTRACTION	Taiwo Joel Omoleye	City University of Hong Kong

Friday, June 28(Day 3)

Room 604

S21 Bayesian Inference Applications [Seong-joon Kim] 09:00-10:30

09:00 –09:30	ICISE-122	BAYESIAN INFERENCE FOR THE EXPONENTIAL-LAW PROCESS	Van Cuong Do	University of South Britany
09:30 –10:00	ICISE-127	A SEQUENTIAL BAYESIAN PARTITIONING APPROACH FOR ONLINE STEADY STATE DETECTION OF MULTIVARIATE SYSTEMS	Jianguo Wu	Peking University
10:00 –10:30	ICISE-172	REVISIT TO BAYESIAN MODEL AVERAGING IN FORECASTING INFLUENZA WITH GOOGLE SEARCH QUERIES	Hao Pan	City University of Hong Kong

S8 Recent Advances in Time Series and High-dimensional Data Analysis [Ching-Kang Ing] 11:00-13:00

11:00 –11:30	ICISE-144	HIGH-DIMENSIONAL MODEL SELECTION VIA CHEBYSHEV GREEDY ALGORITHMS	Ching Kang Ing	National Tsing Hua University
11:30 –12:00	ICISE-146	ON ASYMPTOTIC RISK OF SELECTING MODELS FOR POSSIBLY NON-STATIONARY TIME-SERIES	Shu Hui Yu	National University of Kaohsiung
12:00 –12:30	ICISE-148	MODEL SELECTION FOR HIGH-DIMENSIONAL MISSPECIFIED TIME SERIES MODELS	Hsueh Han Huang	University of National Tsing Hua University
12:30 –13:00	ICISE-147	MODELLING FINANCIAL INTERVAL TIMES SERIES	Li Hsien Sun	National Central University

Room 606

S23 Statistics in General Engineering Applications [Ji Zhu] 09:00-10:30

09:00 –09:30	ICISE-205	MATRIX COMPLETION FOR NETWORK ANALYSIS	Ji Zhu	University of Michigan
09:30 –10:00	ICISE-158	DYNAMIC FIELD MONITORING BASED ON MULTITASK LEARNING IN SENSOR NETWORKS	Xi Zhang	Peking University
10:00 –10:30	ICISE-191	PRIME: A PERSONALIZED RECOMMENDATION FOR INFORMATION VISUALIZATION METHODS VIA EXTENDED MATRIX COMPLETION	Ran Jin	Virginia Polytechnic Institute and State University (Virginia Tech)

S20 Statistics and Analytics [Sungil Kim] 11:00-13:00

11:00 –11:30	ICISE-143	DETECTION AND CLUSTERING OF MIXED-TYPE DEFECT PATTERNS IN WAFER BIN MAPS	Heeyoung Kim	KAIST
11:30 –12:00	ICISE-195	PRIVACY PRESERVING AND SECURE LOGISTIC REGRESSION FOR HORIZONTALLY DISTRIBUTED DATA	Junghye Lee	Ulsan National Institute of Science and Technology
12:00 –12:30	ICISE-186	REVEALING HOUSEHOLD CHARACTERISTICS USING CONNECTED HOME PRODUCTS	Sungil Kim	Ulsan National Institute of Science and Technology

Room 608

S25 Applications of Data Driven Modeling [Si-il Sung] 09:00-10:30

09:00 –09:30	ICISE-125	A NEUTRALLY OPTIMIZED METHOD OF AHP FOR DECISION-MAKING WITH EXPERTISE DATA	Jicang Xu	Ren Min University of China
09:30 –10:00	ICISE-164	STATE-OF-CHARGE ESTIMATION OF LITHIUM-ION BATTERIES VIA LONG SHORT-TERM MEMORY	Fangfang Yang	City University of Hong Kong
10:00 –10:30	ICISE-184	A GENERALIZED STATISTICAL PARAMETER FOR SYSTEM HEALTH MONITORING AND ITS PARALLEL TO DEEP LEARNING	Dong Wang	Shanghai Jiao Tong University

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-108

A STATISTICS-GUIDED APPROACH TO QUALITY CHARACTERIZATION AND COMPENSATORY DESIGN FOR 3D PRINTING

Hao Wang[1], Qiong Zhang[2], Xinwei Deng[3], Kaibo Wang[4]

[1] Department of Industrial Engineering, Tsinghua University and E-mail: w-h14@mails.tsinghua.edu.cn
[2] School of Mathematical and Statistical Sciences, Clemson University and E-mail: qiongz@clemson.edu
[3] Department of Statistics, Virginia Tech and E-mail: xdeng@vt.edu
[4] Department of Industrial Engineering, Tsinghua University and E-mail: kbwang@tsinghua.edu.cn

Abstract

Three-Dimensional (3D) printing plays an important role in customized production of manufacturing. A typical quality problem of 3D printing products is the geometric deformation during the solidification process. In this paper, we propose a statistics-guided approach to quality characterizations of the customized 3D printed surfaces. Based on the quality characterizations, a compensation approach is developed to adjust the original 3D printing design to reduce the deformation. We use two case studies to show that, the proposed quality characterization approach can provide effective quality measures for both simple and complex 3D printing surfaces, and the proposed compensation approach can significantly improve the accuracy of 3D printing by adjusting the original 3D design.

Keywords: *Quality assessment, Gaussian process interpolation, NURBS model*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-109

GAUSSIAN PROCESS MODELING USING INFORMATION FROM PARTIAL DIFFERENTIAL EQUATION MODELS

Matthias HY Tan[1]

[1] School of Data Science, City University of Hong Kong and E-mail: matthtan@cityu.edu.hk

Abstract

Gaussian process (GP) emulators of computer models are typically constructed based purely on data from a computer experiment using a standard stationary GP prior with product Matérn or Gaussian correlation function. This often ignores valuable engineering and mathematical knowledge about the behavior of the computer model. In this talk, I will present my research on the use of known behavior/properties of partial differential equation models solved numerically by computer codes to improve construction of GP emulators for this type of computer models.

Keywords: *Uncertainty quantification, Computer experiments, Emulator*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-110

INDIVIDUALIZED DEGRADATION MODELING AND PROGNOSTICS IN A HETEROGENEOUS GROUP VIA MODELING COVARIATE INTERRELATIONS AMONG UNITS

Minhee Kim[1], Changyue Song[2], Kaibo Liu[3]

[1] University of Wisconsin, Madison and E-mail: mkim555@wisc.edu
[2] University of Wisconsin, Madison and E-mail: csong39@wisc.edu
[3] University of Wisconsin, Madison and E-mail: kliu8@wisc.edu

Abstract

This study focuses on individualized degradation modeling and prognostics for a heterogeneous group, where each individual unit shows a distinct degradation process. Existing degradation models usually treat each unit as independent and do not provide a systematic approach to measure interrelations among different units. In this study, we propose a generic framework which handles the heterogeneity across units by effectively leveraging the covariates of each unit, i.e., the distinct individual characteristics, which are closely related to each individual unit's degradation process. Specifically, we employ a multivariate Gaussian process to model the interrelations among units according to their covariates similarities. The significant advantage of this approach is to enable us to efficiently transfer the information among units according to their interrelations. In other words, the collected degradation signals from one unit can be shared with the entire group, enhancing degradation modeling and prognostics of other units. A theoretical justification of the proposed model is also investigated. Simulation studies are presented to evaluate the estimation accuracy and test the sensitivity of the proposed method. A case study on the Alzheimer's Disease Neuroimaging Initiative dataset is further conducted, which shows the advantage of the proposed method over existing benchmark approaches.

Keywords: *Remaining Useful Lifetime, Transfer Learning, Multivariate Gaussian Process*

ICISE-111

QUALITY BIG DATA

Fugee Tsung[1]

[1] HKUST and E-mail: season@ust.hk

Abstract

This talk will present and discuss the challenges and opportunities that quality engineers and managers 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, especially for industrial big data. Emerging research issues include data fusing with heterogeneous data sources, statistical transfer learning, and statistical process control and monitoring for big data streams.

Keywords: *Quality analytics*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-112

THE MAHALANOBIS-TAGUCHI SYSTEM BASED ON STATISTICAL MODELING

Masato Ohkubo[1]

[1] Faculty of Business Administration, Department of Business Administration, Toyo University and E-mail: ohkubo@toyo.jp

Abstract

Dr. Genichi Taguchi proposed the Mahalanobis-Taguchi (MT) system as a practical methodology for anomaly detection. The MT system developed by adding a unique idea to Hotelling's theory, starting with Hotelling's multivariate control chart. The MT system is widely used in the Japanese manufacturing industry and plays an important role in various fields such as quality control, medical care, and business. However, when the MT system is applied to a real-world problem, it is not always possible to perform appropriate analysis. That is, even when using the MT system, as with the general methods of multivariate analysis, the analyst needs to carefully consider an appropriate analysis procedure based on the purpose of the analysis and the features of the data.

Therefore, in this study, we propose a novel analysis framework for an MT system, based on statistical modeling, to deal with real-world problems adequately and flexibly. Specifically, after using statistical modeling to estimate the model, the Mahalanobis distance is calculated using the parameter estimate of the statistical model. By introducing this process, it is possible to execute an analysis procedure that considers the purpose of the analysis and the features of the data. Consequently, an improvement of both the estimation accuracy of the Mahalanobis distance on the population as well as the anomaly detection performance can be expected.

Keywords: *Anomaly detection, Mahalanobis distance, Mahalanobis-Taguchi method, Recognition Taguchi method, Taguchi method*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-113

TRANSFER LEARNING OF GRAPHICAL STRUCTURES FOR CONNECT MANUFACTURING SYSTEMS

Chao Wang[1], Xiaojin Zhu[2], Shiyu Zhou[3], Yingqing Zhou[4]

[1] University of Wisconsin - Madison and E-mail: cwang436@wisc.edu
[2] University of Wisconsin - Madison and E-mail: jerryzhu@cs.wisc.edu
[3] University of Wisconsin - Madison and E-mail: shiyuzhou@wisc.edu
[4] Dimensional Control Systems, Inc. and E-mail: yzhou@3dcs.com

Abstract

Information revolution is turning modern manufacturing systems into a data-rich environment. This work focuses on developing data-driven methods to identify the dependency relationships among system variables such as quality measures at different manufacturing operations. Such relationship will be very useful in system modeling, monitoring, and control. In this work, graphical modeling approach is employed to characterize such relationship and a transfer learning technique is established for the structure learning through sharing the information from different processes. The effectiveness of the proposed method is demonstrated with both simulations and real case studies.

Keywords: *Graphical models, Informative Prior, Transfer learning*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-114

DESIGN OPTIMIZATION FOR ESTIMATING DISCONTINUOUS REGRESSION FUNCTIONS

Chiwoo Park[1], Peihua Qiu[2]

[1] Florida State University and E-mail: cpark5@fsu.edu
[2] University of Florida and E-mail: pqiu@ufl.edu

Abstract

Selecting input data or design points for statistical models has been of great interest in advanced sensing and active learning. In this talk, we present a new strategy of selecting the design points for a regression model when the underlying regression function is discontinuous. Two main motivating examples are (1) adaptive and sparse imaging with the purpose of accelerating the imaging speed and (2) optimal design for regression analysis over a phase diagram in chemistry. In both of the examples, the underlying regression functions have discontinuities, so many of the existing design optimization approaches cannot be applied for the two examples because they mostly assume a continuous regression function. There are some studies for estimating a discontinuous regression function from its noisy observations, but the noisy observations are typically given and usually densely and are uniformly located over the design space. We extend the study by developing a design strategy of selecting the design points for regression analysis with discontinuities. In this talk, we first review the existing approaches relevant to design optimization and active learning for regression analysis and discuss their limitation for a discontinuous regression function. We then present our novel design strategy for a discontinuous regression function: regression error analysis with a fixed design will be first presented, and the analysis result is used to propose a new criterion of selecting the design points for the regression analysis. Sequential design of experiments with the new criterion will be presented with statistical properties and numerical examples.

Keywords: *Sequential design of experiment, Active learning for regression, Discontinuous regression function*

ICISE-115

AGGREGATION STRATEGIES IN STATISTICAL PROCESS MONITORING

Inez Zwetsloot[1], W.H. Woodall[2]

[1] Department of Systems Engineering and Engineering Management, City University of Hong Kong and E-mail: i.m.zwetsloot@cityu.edu.hk
[2] Virginia Tech and E-mail: bwoodall@vt.edu

Abstract

In this talk, I will first review issues related to aggregation of data and its effect on process monitoring performance, based on Zwetsloot and Woodall (2018). The aggregation of data into samples is ubiquitous in statistical applications. It has always been a common practice in process monitoring, but it is becoming a more important issue as data are collected at an increasing frequency. We review the literature and offer some practical advice and some directions for future research.

In the second part of this talk, I will dive into the effect of temporal aggregation on multivariate SPM. In this part, we study the effect of temporal aggregation MV SPC and also discuss appropriate performance metrics to use. Some case studies will illustrate these issues.

Keywords: *Temporal aggregation, Statistical process monitoring, average time to signal, multivariate control chart, Timestamp, Time-between-events*

References: *I.M. Zwetsloot and W.H. Woodall (2019). "A Review of Some Sampling and Aggregation Strategies for Basic Statistical Process Monitoring" – with discussions. Accepted for publication in the Journal of Quality Technology.*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-117

A SPARSE LEADING-EIGENVALUE-DRIVEN CONTROL CHART FOR PHASE I ANALYSIS OF HIGH-DIMENSIONAL COVARIANCE MATRICES

Jinyu Fan[1], Lianjie Shu[2]

[1] School of Mathematics and Statistics, Guang Dong University of Finance and Economics and E-mail: yb47001@connect.um.edu.mo
[2] Faculty of Business, University of Macau and E-mail: ljshu@um.edu.mo

Abstract

In statistical process control (SPC), a proper Phase I analysis is essential to the success of Phase II monitoring. With recent advances in sensing technology and data acquisition systems, Phase I analysis of high-dimensional data is increasingly encountered. However, the high dimensionality presents a new challenge to the traditional Phase I techniques. A literature review reveals nearly no Phase I techniques in existence for analyzing high-dimensional process variability. Motivated by this, this paper develops a sparse-leading-eigenvalue-driven control chart for retrospectively monitoring high-dimensional covariance matrices in Phase I, denoted as the SLED control chart. The key idea of it is to track changes in the sparse leading eigenvalue between two covariance matrices. It is shown that the proposed method can gain high detection power, especially when the shift is weak and is not very dense, which is often the case in practical applications.

Keywords: *Eigenvalue, Two-sample test*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-119

ISO TC69/SC8: STANDARDIZING METHODOLOGIES FOR NEW TECHNOLOGY AND PRODUCT

Watalu Yamamoto[1], Hiroe Tsubaki[2]

[1] University of Electro-Communications and E-mail: watalu@inf.uec.ac.jp
[2] The Institute of Statistical Mathematics and E-mail: tsubaki@ism.ac.jp

Abstract

ISO TC69/SC8, application of statistical and related methodology for new technology and product development, has been established in 2009 and is the newest SC of ISO/TC69, Applications of statistical methods. The purpose of this subcommittee is standardizing a process for new technology and product development based on a model of value chain cycle among voice of customers, quality, parameters of systems, good design, and good products. Primary activities of SC8 covers three areas, i) collection of voices of customers, ii) transformation of voice of customers into engineering models or systems, and iii) system design and optimization. So far an international standard on robust tolerance design has been developed for optimization and a series of international standards using quality function deployment have been developed for transformation. In this talk, the scope of activities of ISO/TC69/SC8 is introduced and the developed standards, ISO standards 16336 on robust parameter design, 16355 on quality function deployment, and a standard under development 16337 on robust tolerance design is explained in details.

Keywords: *Value chain cycle model, Taguchi method, robust parameter design, robust tolerance design, quality function deployment*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-120

STATISTICAL INFERENCE FOR M_t/G/INFINITY QUEUEING SYSTEMS WITH INCOMPLETE INFORMATION

Dongmin Li[1], Qingpei Hu[2], Dan Yu[3]

[1] Center of Quality and Data Science, Academy of Mathematics and Systems Science; School of Mathematical Sciences, University of Chinese Academy of Sciences and E-mail: lidongmin17@mails.ucas.ac.cn
[2] Center of Quality and Data Science, Academy of Mathematics and Systems Science; School of Mathematical Sciences, University of Chinese Academy of Sciences and E-mail: qingpeihu@amss.ac.cn
[3] Center of Quality and Data Science, Academy of Mathematics and Systems Science; School of Mathematical Sciences, University of Chinese Academy of Sciences and E-mail: dyu@amss.ac.cn

Abstract

M_t/G/Infinity queueing systems are widely used models to analyze real-life complex systems such as software testing systems, telecommunication systems, and so on. Given information of the queueing system, the statistical inference on performance measures, the expected cumulative number of arrivals and departures, is of great importance for decision makers to analyze current situation, predict future situation, and make cost-effective decisions accordingly. In most real-life cases, only incomplete information is available because it is often difficult to keep track of each item from arrival to departure. In this paper, we provide a general framework, inferring on model parameters and the expected cumulative number of arrivals and departures given interval censored data, which are given as counts that occur in fixed time intervals as opposed to the exact times. We propose a Maximum Likelihood Estimation method to infer on the arrival rate and service duration and the method is applicable to general service duration distribution. More importantly, a combination of bootstrap method and delta method is proposed to infer on the expected cumulative number of arrivals and departures. Simulation study shows that the point and interval estimation of the proposed MLE method are overall satisfying. As the number of intervals increases, the results of the proposed MLE are toward the results of MLE given complete data. Our approach enables one to obtain estimates without having to keep track of each item, saving a great deal of resource monitoring items and storing data. Application on a software testing system shows that the goodness-of-fit performance and predictive performance of the proposed MLE method are satisfying.

Keywords: *queueing, interval censored data, Maximum Likelihood Estimation (MLE), parametric bootstrap, delta method, predictive performance*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-121

ELECTRICITY DEMAND FORECASTING WITH DETECTING AND PREDICTING CHANGES

Dongyeon Jeong[1], Young Myoung Ko[2]

[1] Pohang University of Science and Technology and E-mail: dongyeon.jeong@lstlab.org
[2] Pohang University of Science and Technology and E-mail: youngko@postech.ac.kr

Abstract

In this talk, we present an algorithm for predicting changepoints in energy demand with periodic trend changes. Energy demand profiles in buildings or factories often show sudden increases and decreases during day and night times. Such abrupt changes may significantly affect the performance of a prediction algorithm since a slight error in changepoints results in a considerable error between real and predicted values. Changepoint detection algorithms aim to find abrupt changes in time series data. Previous studies focused on either detecting change points from the observed data or determining if the new data is a change point based on the observed data. There, however, are not many studies on predicting future changepoints. If we know when the change will take place in the future, it will help reduce the prediction error of energy demand. Online changepoint detection is known to be adequate for evaluating the very next data, but it has a limitation in predicting when changes occur again. Therefore, we propose a changepoint prediction algorithm to reduce the prediction error. We build a deep learning model to predict changepoints and add a probabilistic approach to obtaining confidence intervals that changepoints may fall. Numerical experiments show the effectiveness of our algorithm.

Keywords: *Changepoint, Prediction, Deep learning, Confidence interval*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-122

BAYESIAN INFERENCE FOR THE EXPONENTIAL-LAW PROCESS

Van Cuong Do[1], Evans Gouno[2]

[1] University of South Britany, Campus of Tohannic, 56000 Vannes, France and E-mail: van-cuong.do@univ-ubs.fr
[2] University of South Britany, Campus of Tohannic, 56000 Vannes, France and E-mail: evans.gouno@univ-ubs.fr

Abstract

*In this work we develop Bayesian analysis for the exponential-law process (ELP). The ELP was first introduced by Goel and Okumoto in 1979 for modelling software reliability. Since then, it is also referred as Goel-Okumoto model. This is a nonhomogeneous Poisson process with exponential intensity so we use the terminology "**exponential-law process**" in our work. It is considered as an alternative model for the power-law process (PLP) when the grow rate changes very fast according to an exponential-law. The likelihood is tractable but it requires numerical methods to get maximum likelihood estimates.*
*For Bayesian inference on the ELP, we consider different choices for priors. Firstly, we consider a noninformative prior using Jeffreys'rule. That leads us to introduce the **modified-Gumbel** distribution. This is a unimodal distribution having finite second moment but the expectation and the variance of this distribution need numerical methods to be obtained. Secondly, we propose a reparametrization in order to define independent conjugate priors. Thirdly, we suggest a joint conjugate prior for the ELP as a combination of gamma distributions and modified-Gumbel distribution. We name that natural conjugate prior the **Gamma-modified-Gumbel** distribution. Properties of this bivariate distribution are investigated in details. We then introduce some elicitation strategies with trials and errors procedures to obtain values for the hyperparameters. A simulation study is conducted to compare Bayesian and maximum likelihood estimates. We employ both ELP and PLP to a real dataset and use AIC and BIC to compare the two models.*

Keywords: *Reliability, Stochastic process, Bayesian inference, Noninformative prior, Conjugate prior, Simulation*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-125

A NEUTRALLY OPTIMIZED METHOD OF AHP FOR DECISION-MAKING WITH EXPERTISE DATA

Jicang Xu[1], Xiaomi An[2], Ming Ren[3]

[1] Ren Min University of China and E-mail: xujicang@ruc.edu.cn
[2] Ren Min University of China and E-mail: anxiaomi@ruc.edu.cn
[3] Ren Min University of China and E-mail: renm@ruc.edu.cn

Abstract

In data-rich world, structure and unstructured data, real-time and historic data continues to relating the physical world to the theoretical world. Using regression analysis and machine learning, we easily build relationship among variables to help us make decisions in physical world. However, those methods are based on prior data. What if there is no mapping between the physical and theoretical world before in certain condition, like a social management mechanism needing assessment and comparison. This time, we use expertise data which is based on observation to make decision under evaluation index system. And then the expertise data becomes prior data. AHP is such a milestone for expertise data decision-making. But it's been long time since its birth.

***This research aims at** (1) reducing the inaccuracy in the process of index weight calculation, considering the interaction effects among the index. Different from AHP, ANP and DEMATEL method, this method aims to (2) balance the impacts from the subjective empowerment of experts and the objective interrelationship in index system.*

*According to the research, **we find this method can** apparently (1) enhance the rationality measurement value of index weight which is based on the information entropy formula. Compared with traditional AHP, this method (2) transfers the abundant weight that is formed for the overall influence reason to the exactly important index. This method also (3) eliminates the shortcomings of inconsideration of the expertise empowerment for each index while ANP and DEMATEL only get results from the interrelationship analysis. Finally, on aspect of operation, this method contributes to the (4) efficiency of doing expert review in this field, supporting the correct and efficient decision-making.*

*This new method can be put forward to make further discussion among scholars whose **research field is decision-making system or evaluation index system** like risk assessment. We believe this method is a comparatively great modification of AHP.*

We mainly achieve the method by using modified PageRank algorithm to train expertise data and apply it into a case of management evaluation. More details on the presentation.

***Keywords:** AHP, decision-making system, evaluation index system*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-126

OPPORTUNISTIC MAINTENANCE POLICY FOR MULTI-UNIT SYSTEM UNDER MARKOVIAN ENVIRONMENT CHANGES

Yu Odajima[1], Lu Jin[2]

[1] Department of Informatics niverisity of Electro-Communications, Tokyo and E-mail: kotashimashu@gmail.com
[2] Department of Informatics Univerisity of Electro-Communications, Tokyo and E-mail: jinlu@inf.uec.ac.jp

Abstract

Opportunistic maintenance (OM) is a form of maintenance planning based on convenient replacement by taking advantage of maintenance of other units. This research studies an opportunistic maintenance policy for a multi-unit system under a varying environment. The environment conditions are assumed to affect the deterioration process of the units directly. The transitions of both the deterioration state and environment condition follow a discrete-time Markov chain. The occurrence probability of a maintenance opportunity comes from other units following a Bernoulli process. Since the true deterioration state cannot be known exactly, the decision-maker obtains the information on the deterioration states probabilistically from monitoring results at every time period. This research focuses on one of the units and formulates the decision-making problem as a partially observable Markov decision process (POMDP). The cost structure is composed of repair, operating, and two kinds of setup cost. This research investigates the properties of an optimal opportunistic maintenance policy minimizing the expected total discounted cost over an infinite horizon and provides sufficient conditions under which the objective function for optimizing maintenance policy is non-decreasing in a deterioration state and environment condition. Moreover, regarding the presence or absence of maintenance opportunities, this research obtains some properties on the threshold for optimal maintenance actions within a set of a deterioration state and environment. Finally, a numerical example is provided to illustrate our obtained results.

Keywords: *Decision-making, Deteriorating System, Partially Observable Markov Decision Process, Preventive Maintenance, Optimal Maintenance Policy*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-127

A SEQUENTIAL BAYESIAN PARTITIONING APPROACH FOR ONLINE STEADY STATE DETECTION OF MULTIVARIATE SYSTEMS

Jianguo Wu[1]

[1] College of Engineering, Peking University and E-mail: j.wu@pku.edu.cn

Abstract

Steady state detection is critically important in many engineering fields, such as fault detection and diagnosis, process monitoring and control. However, most of the existing methods were designed for univariate signals. In practice, most of the systems or processes are inherently multivariate. With the rapid development of sensing technology, multiple sensor signals have become unprecedentedly available to better capture the system conditions. Therefore, multivariate steady state detection algorithms are highly desirable. In this paper, we propose an efficient online steady state detection method for multivariate systems through a sequential Bayesian partitioning approach. The multivariate signal is sequentially segmented into phases of constant mean and covariance matrix under the Bayesian framework, and the posterior distribution of the phase duration is used to test the steady state. Once the duration is sufficiently large, the signal is claimed steady. The main challenges of this method are how to sequentially find the change-point regarding the mean and covariance, and how to select appropriate hyperparameters. To overcome these challenges, we develop an efficient recursive method to calculate the posterior distributions analytically, and then provide several insightful guidelines on hyperparameter selection. The effectiveness of the proposed method is demonstrated through thorough numerical and real case studies.

Keywords: *Steady state detection, multivariate system, change-point detection*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-128

MODIFICATION OF THE GLR CHART STATISTICS FOR MONITORING COUNT PROCESSES

Jaeheon Lee[1]

[1] Chung-Ang University and E-mail: jaeheon@cau.ac.kr

Abstract

In this paper we consider the generalized likelihood ratio (GLR) charts for monitoring count processes, such as binomial, Bernoulli, Poisson, and multinomial processes. The chart statistic in these GLR charts contains the maximum likelihood estimator (MLE) of the out-of-control value of a process parameter. In some special cases of the observations, however, the value of MLE has been said to make the GLR statistic undefined. Several approaches to avoiding this problem have been proposed, such as putting an upper and/or a lower bound on the MLE, and using enough observations to get a reasonable, non-problematic value of the MLE. The objective of this paper is to modify the GLR chart statistic in order to show that it is well-defined in any situation. The advantage of this modification is that the GLR chart does not require practitioners to use arbitrary bounds on the MLE to avoid an undefined GLR chart statistic.

Keywords: *generalized likelihood ratio chart, count process, statistical process control*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-130

SPATIO-TEMPORAL SEQUENCE PREDICTION BASED ON SINGULAR VALUE DECOMPOSITION AND ARIMA

Lining Yang[1], Yanting LI[2], Lianjie Shu[3]
[1] Shanghai Jiao Tong University and E-mail: yanglining@sjtu.edu.cn
[2] Shanghai Jiao Tong University and E-mail: ytli@sjtu.edu.cn
[3] University of Macau and E-mail: ljshu@um.edu.mo

Abstract

Real-time short-term and long-term prediction based on spatio-temporal data now becomes feasible and necessary. Spatial-temporal models which take spatial information into account tend to provide superior forecast than time series models. Motivated by this, this paper proposes a spatio-temporal prediction model, ST-SVD, which com bines singular value decomposition (SVD) and traditional autoregressive integrated moving average model (ARIMA) together. The performance of the new model was compared with Lasso vector autoregressive (Lasso-VAR), autoregressive-integrated moving average (ARIMA) and multivariate long short term memory (LSTM). The results show that the new method outperforms other models in term of both computational e ciency and accuracy.

Keywords: *Spatio-Temporal Sequence, Singular Value Decomposition, ARIMA*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-131

GENERAL PATH MODELS FOR DEGRADATION DATA WITH MULTIPLE CHARACTERISTICS AND COVARIATES

Lu Lu[1], Bing Xing Wang[2], Yili Hong[3], Zhisheng Ye[4]

[1] Department of Mathematics and Statistics, University of South Florida and E-mail: lulu1@usf.edu
[2] School of Statistics, Zhejiang Gongshang University and E-mail: wangbingxing@163.com
[3] Department of Statistics, Virginia Tech and E-mail: yilihong@vt.edu
[4] Dept. of Industrial Systems Eng. and Management, National University of Singapore and E-mail: yez@nus.edu.sg

Abstract

Degradation data have been broadly used for assessing product and system reliability. Most existing work focuses on modeling and analysis of degradation data with a single characteristic. In some degradation tests, interest lies in measuring multiple characteristics of the product degradation to understand different aspects of the reliability performance, resulting in degradation data with multiple characteristics. The literature in modeling such data is scarce. Motivated by the photodegradation process of polymeric materials, we propose a multivariate general path model for degradation data with multiple degradation characteristics (DCs). The proposed model incorporates covariates for modeling the nonlinear degradation path, and it also includes random effects that are correlated among the multiple DCs to capture the unit-to-unit variation in the individual degradation paths and model the dependence between the multivariate measurements. An expectation-maximization algorithm combined with the Markov Chain Monte Carlo simulation is developed for estimating the model parameters and quantifying their associated uncertainty, which are then used for predicting the product reliability. The performance of the developed method is evaluated and compared with existing methods through a simulation study. The proposed method is also illustrated with an example for analyzing the NIST coating degradation test data.

Keywords: *photodegradation, nonlinear general path model, multivariate degradation data, MCEM algorithm, simulation*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-132

PHASE-TYPE DISTRIBUTIONS FOR PRODUCT RETURN DATA WITH TWO-LAYER CENSORING

Wang Yudong[1], Ye Zhi-Sheng[2]

[1] Department of Industrial Systems Engineering & Management, National University of Singapore and E-mail: yudongw@u.nus.edu
[2] Department of Industrial Systems Engineering & Management, National University of Singapore and E-mail: yez@nus.edu.sg

Abstract

Product return data, such as warranty claims, are commonly subject to two layers of right censoring. The first layer, which is called warranty censoring, applies to the product lifetime due to a fixed warranty limit. The second layer, which is called end-of-study censoring, applies to the sum of the sales lag and the lifetime due to the end-of-study date for the data collection. An unreturned item would either have a lifetime longer than the warranty limit or the sum of the sales lag and the lifetime longer than the end-of-study date. The two-layer censoring in the product return data renders traditional nonparametric methods for right-censored data inapplicable. This study develops a generic method for the two-layer censored data using acyclic phase-type distributions (APHDs) in the canonical form. The APHD estimators can be regarded as nonparametric sieve estimators since the family of APHDs is dense in the field of all positive-valued distributions. Based on the property that the class of APHDs is closed under convolution, a dedicated expectation-maximization algorithm is proposed for parameter estimation. Comprehensive simulations are conducted to evaluate the performance and compare with the inverse probability of censoring weighted approach, which is applicable in the absence of warranty censoring. Two real examples are used to illustrate the proposed method.

Keywords: *Warranty censoring, end-of-study censoring, sales lag, sieve estimator*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-133

ACCELERATED DEGRADATION RELIABILITY DEMONSTRATION TEST PLAN USING DIRECT PREDICTION METHOD

Seongjoon Kim[1], Suk Joo Bae[2], Byeongmin Mun[3]
[1] Chosun University and E-mail: seongjoon.kim@chosun.ac.kr
[2] Hanyang University and E-mail: sjbae@hanyang.ac.kr
[3] Hanyang University and E-mail: stat@psm.hanyang.ac.kr

Abstract

The reliability demonstration test (RDT) is widely adopted to make a decision of either accepting or rejecting the product's reliability. This study proposes a method for RDT plan by ADT using the lifetime distribution based direct prediction method. The proposed method is useful in two aspects: (1) the functional link between degradation model and stress-acceleration model is not required. (2) it can be easily applied by practitioners since the design factors and the decision rule is based on the conventional lifetime distribution.

Keywords: *Accelerated Degradation Test, Reliability Demonstration Test Plan*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-136

STATISTICAL INFERENCE COMBINED WITH PERSISTENT HOMOLOGY FOR PREDICTING FLUID FLOW IN POROUS MEDIA

Chul Moon[1], Scott A. Mitchell[2], Jason E. Heath[3], Andrew Matthew[4]

[1] Southern Methodist University and E-mail: chulm@smu.edu
[2] Sandia National Laboratories and E-mail: samitch@sandia.gov
[3] Sandia National Laboratories and E-mail: jeheath@sandia.gov
[4] Carl Zeiss X-ray Microscopy and E-mail: matthew.andrew@zeiss.com

Abstract

We propose a porous materials analysis pipeline using persistent homology. We first compute persistent homology of binarized 3D images of sampled material subvolumes. For each image we compute sets of homology intervals, which are represented as summary graphics called persistence diagrams. We convert persistence diagrams into image vectors in order to analyze the similarity of the homology of the material images using the methods developed for image analysis. Each image is treated as a vector and we compute its principal components to extract features. We fit a statistical model using the loadings of principal components to estimate material porosity, permeability, anisotropy, and tortuosity. We also propose an adaptive version of the Structural SIMilarity index (SSIM), a similarity metric for images, as a measure to determine the Statistical Representative Elementary Volumes (sREV) for persistence homology. Thus we provide a capability for making a statistical inference of the fluid flow and transport properties of porous materials based on their geometry and connectivity.

Keywords: *Topological Data Analysis, Persistent Homology, Porous material*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-137

THE EFFECT OF AMGRU ON THE TIME-SERIES ANALYSIS

Joo Heui Son[1], Jee Whan Yoon[2], Sung Won Han[3]

[1] School of Industrial Management Engineering, Korea University and E-mail: nanashon@korea.ac.kr
[2] School of Management of Technology, Korea University and E-mail: towny@korea.ac.kr
[3] School of Industrial Management Engineering, Korea University and E-mail: swhan@korea.ac.kr

Abstract

Time series prediction tasks have been challenged and there are still many problems to solve. Although many researches have been proposed to improve its performance successfully via deep learning methods, there is still room for further improvement. The dominant time-series prediction models are based on RNN(Recurrent Neural Network) which has a limitation of long-term memory dependency due to gradient vanishing problem. To overcome this, we propose extended network model of existing GRU(Gated Recurrent Unit), so called, AMGRU(Attention on Multivariate GRU) model, to predict time series data further better. In this model, GRU and self-attention mechanism are integrated to reflect the characteristics of features in the previous layers well. To show that the proposed model is universally effective, experiments are conducted using six different data sets. The results showed that the proposed model achieved the highest performance of MSE(Mean squared error), MAPE(Mean absolute percentage error), MSLE(Mean squared logarithmic error) metrics for all experimented datasets compared with other existing deep learning models.

Keywords: *Time-series analysis, Deep learning, Multivariate time series prediction, self-attention, GRU, forecasting*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-138

THE EFFECT OF CONTINUOUSLY UPDATING CONTROL CHART LIMITS ON CONTROL CHART PERFORMANCE

Leo C.E. Huberts[1], Marit Schoonhoven[2], Ronald J.M.M. Does[3]

[1] Department of Operations Management, University of Amsterdam and E-mail: L.C.E.Huberts@uva.nl
[2] University of Amsterdam
[3] University of Amsterdam

Abstract

An open topic within Statistical Process Monitoring is the effect on control chart properties of updating the control chart limits during the monitoring period. The challenge is to use the correct data for updating the control limits as in-control data could be incorrectly classified as out-of-control and therefore not used for re-estimating the parameters, and out-of-control data could be classified as in control and therefore used for re-estimation. In the present article we study the effect of updating the Shewhart, Cumulative Sum and Exponentially Weighted Moving Average control chart limits. We simulate different scenarios: the monitoring data could be in or out-of-control and the practitioner may or may not be able to find out whether the process is indeed out-of-control when the control chart gives a signal. The results reveal that the variation in the performance of the conditional control charts decreases significantly as a result of updating the control chart limits when the updating data are in control and also when the updating data are out-of-control and the practitioner is able to classify correctly data samples that produce a signal. However, when a practitioner is not able to classify a signal correctly, the advisability of updating depends on the type of control chart and the level of data contamination.

Keywords: *big data, conditional, CUSUM, EWMA, Shewhart, unconditional*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-139

NONPARAMETRIC CONTROL OF THE CONDITIONAL PERFORMANCE IN STATISTICAL PROCESS MONITORING

Rob Goedhart[1]

[1] IBIS UvA, Department of Operations Management, University of Amsterdam and E-mail: r.goedhart2@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. This holds even when the actual distribution of the data is known. When the distribution of the data is unknown, it has to be estimated as well, along with its parameters. This yields even more uncertainty in control chart performance when parametric models are applied. Thus, in order to deal with this variation, nonparametric control limits for Shewhart control charts are proposed that guarantee a minimum in-control performance with a specified probability. Other existing methods, such as data-transformations or a bootstrap procedure, are also discussed.

A minimum in-control performance guarantee generally lowers the out-of-control performance. In order to balance the trade-off between in-control and out-of-control performance, the minimum performance threshold and specified probability can be adjusted as desired. Next to that, a major advantage of nonparametric control limits is that they can be applied to any monitoring statistic of interest, such as the mean or the standard deviation.

Keywords: *Control charts, Parameter estimation, Nonparametric*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-140

DATA TRANSFORMATIONS OR THE GENERALIZED LINEAR MODEL? SOME INSIGHTS AND RECOMMENDATIONS

Marcus Perry[1]

[1] Department of Info Systems, Statistics, and Management Science, University of Alabama and E-mail: mperry@cba.ua.edu

Abstract

In today's Industry 4.0, industrial processes are increasing in complexity, presenting significant challenges to the industrial experimenter. Nonstandard experimental data is quite common with such processes, and is typically analyzed in practice by either: 1) applying a data transformation to the response, i.e., transform to normality, and proceeding with an analysis developed under normal theory models, or 2) analyzing the experimental data via the class of generalized linear models. In this talk, we consider the process of machining copper bars and subsequently using electron microscopy to statistically examine the resulting surface microstructure, namely the distribution of grain sizes. The resulting experiments lead to two common nonstandard situations, i.e., correlations amongst observations, and non-normal residual error distributions. Using this dataset as motivation, we discuss the advantages and disadvantages, both theoretical and practical, to the use of data transformations versus the generalized linear model (and its variants, e.g., GEEs, GLMMs) when the experimental response data are nonstandard. We also discuss their implications from an experimental design perspective. We analyze the electron microscopy experimental data under both analysis frameworks and highlight some of the major differences observed from the results of the analyses. We close with some general insights and recommendations for the modern data modeler.

Keywords: *Generalized linear model, Data transformations, Design of experiments*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-141

BALANCING ROBUST CLASSICAL DESIGNS WITH DESIGN CONSTRAINTS – APPLICATIONS IN TESTING DEFENSE SYSTEMS

Laura Freeman[1]

[1] Virginia Tech and E-mail: laura.freeman@vt.edu

Abstract

Historical approaches for statistically planned experiments emphasize "classical" designs such as 2^k full factorial designs, 2^{k-p} fractional factorial designs, Plackett-Burman designs, central composite designs (CCDs), Box-Behnken designs, etc. These classical designs emphasizes design robustness over optimality. Robustness, here, refers to the capability of the experimental design and its associated analysis to withstand complications that arise in physical testing, which may include outliers, missing data, nuisance errors, and violations to the statistical assumptions. However, in many engineering applications, including defense system testing, constraints limit the direct applications of classical designs. For example, defense systems contain cutting edge technologies, often engineered to work in new ways that need more testing in some regions than others. Additionally, testing is expensive testing and often constrained to non-regular design regions due to safety or logical considerations. This talk will show several examples of how "classical" experimental designs principles have improved defense system assessments ranging from body armor to fighter aircraft using optimal-designs and sequential Bayesian approaches. I will showcase how classical design of experiments principles can be used to test systems sequentially, and provides a defensible trade-space tool for decision makers. Finally, I will discuss how design of experiments is being applied to computer simulations of defense systems to help shape future test designs and provide a basis for validating computer experiments.

Keywords: *Experimental Design, Defense Testing, Constrained Design, Optimal Design, Classical Design*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-142

PERSONALIZED LEARNING SYSTEM FOR SMART TRANSPORTATION DEMAND MANAGEMENT (TDM)

Jingshuo Feng[1], Xi Zhu[2], Feilong Wang[3], Shuai Huang[4], Cynthia Chen[5]

[1] Industrial & Systems Engineering, University of Washington and E-mail: jingsf@uw.edu
[2] Civil & Environmental Engineering, University of Washington and E-mail: zhux2012@uw.edu
[3] Civil & Environmental Engineering, University of Washington and E-mail: flwang@uw.edu
[4] Industrial & Systems Engineering, University of Washington and E-mail: shuaih@uw.edu
[5] Civil & Environmental Engineering, University of Washington and E-mail: qzchen@uw.edu

Abstract

Much research attention has been focused on the Transportation demand management (TDM) in the past decades due to the rising high demand in driving which is closely related to a number of urban issues, such as congestion, air pollution, and public health. While being able to provide accurate personalized incentive holds great value for the success of the TDM systems, the rapid proliferation of smart personal technologies in recent years make it possible to offer incentives individually. In this presentation, we will introduce our ongoing works in developing new analytic methods that can statistically accurately to estimate individual's travel preference from their behavior data (sample size limited and fragmented), computationally efficient to implement, and easy to be integrated with dynamic incentive optimization methods to modify individuals' travel behaviors to achieve optimal solutions in both system-level and individual-level goals.

Keywords: *Smart transportation demand management, personalized behavior modeling, machine learning*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-143

DETECTION AND CLUSTERING OF MIXED-TYPE DEFECT PATTERNS IN WAFER BIN MAPS

Heeyoung Kim[1]

[1] KAIST and E-mail: heeyoungkim@kaist.ac.kr

Abstract

In semiconductor manufacturing, a wafer bin map (WBM) is a map that consists of assigned bin values for dies based on wafer test results (e.g., value 1 for good dies and value 0 for defective dies). The bin values of adjacent dies are often spatially correlated, forming some systematic defect patterns. These non-random defect patterns occur due to assignable causes; therefore, it is important to identify these systematic defect patterns in order to know the root causes of failure and to take actions for quality management and yield enhancement. In particular, as wafer fabrication processes have become more complicated, mixed-type defect patterns (two or more different types of defect patterns occur simultaneously in a single wafer) occur more frequently than in the past. For more effective classification of wafers based on their defect patterns, mixed-type defect patterns need to be detected and separated into several clusters of different patterns; subsequently, each cluster of a single pattern can be matched to a well-known defect type (e.g., scratch, ring) or it may indicate the emergence of a new defect pattern. There are several challenges to be overcome in the detection and clustering of mixed-type defect patterns. These include (i) the separation of random defects from systematic defect patterns; (ii) determining the number of clusters; and (iii) the clustering of defect patterns of complex shapes. To address these challenges, in this article, we propose a new framework for detecting and clustering mixed-type defect patterns. First, we propose a new filtering method, called the connected-path filtering method, to denoise WBMs. Subsequently,we adopt the infinite warped mixture model for the clustering of mixed-type defect patterns; this model is flexible in its ability to deal with complex shapes of defect patterns; furthermore, the number of clusters does not need to be specified in advance but is automatically determined simultaneously during the clustering procedure. We validate the proposed method using real data from a semiconductor company. The experimental results demonstrate the effectiveness of the proposed method in estimating the number of underlying clusters as well as in the clustering of mixed-type defect patterns.

Keywords: *semiconductor manufacturing*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-144

HIGH-DIMENSIONAL MODEL SELECTION VIA CHEBYSHEV GREEDY ALGORITHMS

Ching-Kang Ing[1]

[1] Institute of Statistics, National Tsing Hua University and E-mail: cking@stat.nthu.edu.tw

Abstract

We are concerned with model selection problems in high-dimensional sparse nonlinear models. We first use the Chebyshev greedy algorithm (CGA) to perform variable screening and derive, under a fairly general sparsity condition, its rate of convergence in terms of the number of iterations and the approximation error. We then introduce a high-dimensional information criterion (HDIC) to determine the number of CGA iterations, and show that CGA used in conjunction with HDIC achieves the optimal rate of convergence. Finally, the proposed method is applied to the analysis of high-dimensional logistic and Cox regressions.

Keywords: *high-dimensional sparse nonlinear models, Chebyshev greedy algorithm, high-dimensional information criterion*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-146

ON ASYMPTOTIC RISK OF SELECTING MODELS FOR POSSIBLY NON-STATIONARY TIME-SERIES

Shu-Hui Yu[1]

[1] Institute of Statistics, National University of Kaohsiung and E-mail: shuhui@nuk.edu.tw

Abstract

Model selection criteria are often assessed by the so-called asymptotic risk. Asymptotic risk is defined either with the mean-squared error of estimated parameters (James and Stein, 1961, Stein, 1981, Hansen, 2014); or with the mean-squared error of prediction (Shibata, 1976, Li, 1987, Hansen, 2007, Leeb, 2009). The literature focuses on i.i.d. or stationary time-series data though. Using the latter definition of asymptotic risk, this paper assesses the conventional AIC-type and BIC-type information criteria, which are arguably most suitable for univariate time series in which the lags are naturally ordered. Throughout we consider a univariate AR process in which the AR order and the order of integratedness are finite but unknown. We prove the BIC-type information criterion, which penalty goes to infinity, attains zero asymptotic excess risk. In contrast, the AIC-type information criterion, which penalty goes to a finite number, renders a strictly positive asymptotic excess risk. Further, the asymptotic excess risk increases with the admissible number of lags Kn. The last result gives a warning about certain high-dimensional analyses on time series data, should the underlying data generating process is low-dimensional. In sum, we extend the existing asymptotic risk results in threefold: (i) a general I(d) process; (ii) same realization prediction; and (iii) an information criterion more general than AIC. Some simulation study compares the excess risk of AIC with those of AIC3, HQIC, BIC, Lasso as well as adaptive Lasso.

Keywords: *AIC-type information criteria, BIC-type information criteria, high-dimensional analyses, integrated processes, Lasso, same-realization prediction*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-147

MODELLING FINANCIAL INTERVAL TIMES SERIES

Liang-Ching Lin[1], Li-Hsien Sun[2]

[1] National Cheng Kung University and E-mail: lclin@mail.ncku.edu.tw
[2] National Central University and E-mail: lihsiensun@ncu.edu.tw

Abstract

In financial economics, a large number of models are developed based on the daily closing price. When using only the daily closing price to model the time series, we may discard valuable intra-daily information, such as maximum and minimum prices. In this study, we propose an interval time series model, including the daily maximum, minimum, and closing prices, and then apply the proposed model to forecast the entire interval. The likelihood function and the corresponding maximum likelihood estimates (MLEs) are obtained by stochastic differential equation and the Girsanov theorem. To capture the heteroscedasticity of volatility, we consider a stochastic volatility model. The efficiency of the proposed estimators is illustrated by a simulation study. Finally, based on real data for S&P 500 index, the proposed method outperforms several alternatives in terms of the accurate forecast.

Keywords: *Interval time series, stochastic differential equation*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-148

MODEL SELECTION FOR HIGH-DIMENSIONAL MISSPECIFIED TIME SERIES MODELS

Hsueh-Han Huang[1], Ching-Kang Ing[2], Wei-Cheng Hsiao[3]

[1] University of National Tsing Hua University, Taiwan and E-mail: s106024801@m106.nthu.edu.tw
[2] University of National Tsing Hua University, Taiwan and E-mail: cking@stat.nthu.edu.tw
[3] Soochow University, Taiwan and E-mail: hsiaowc@scu.edu.tw

Abstract

In the current literature, little attention has been paid to the model selection for high-dimensional misspecified models. However, methods for model selection in misspecified models can be applied to many useful models, such as interaction models or measurement error models. In this article, we investigate the behavior of the orthogonal greedy algorithm (OGA) in high-dimensional misspecified time series models. Under a weak sparsity condition, we derive the convergence rate of OGA. By further assuming the structure of true models, we show that OGA, used in conjunction with a high-dimensional information criteria (HDIC), can achieve the sure screening property (in the sense of misspecified models) under a strong sparsity condition. We propose a concept called the "degrees of misspecification" to illustrate the effects of model misspecification on variable screening. We introduce an algorithm generalized from OGA, called Multi-step OGA (MOGA), and show that it shares similar theoretical properties as OGA but has better performance than OGA in some finite sample cases. Two special cases, interaction models and measurement error models, of misspecified models are studied. We propose a new two-stage model selection procedure for interaction models, called MOHIT-2, under hierarchical model structure. We also propose a novel method named MOHITbc to tackle model selection for measurement error models. Both MOHIT-2 and MOHITbc are shown to possess model selection consistency. Simulation studies and real data analysis are given to demonstrate the advantages of our methods.

Keywords: *Measurement error models, Model misspecifi cation, Interaction models, Sure screening, High-dimensional, Time series*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-149

EFFICIENT INPUT UNCERTAINTY QUANTIFICATION VIA GREEN SIMULATION USING SAMPLE-PATH LIKELIHOOD RATIOS

Eunhye Song[1], Ben Feng[2]

[1] Pennsylvania State University and E-mail: eus358@psu.edu
[2] University of Waterloo and E-mail: ben.feng@uwaterloo.ca

Abstract

Bootstrap is a popular tool for quantifying input uncertainty, variance of simulation output caused by finite-sample estimation error in the input model used to run the simulation. A typical bootstrap-based procedure requires R replications at each of B bootstrapped input distributions, totaling BR simulation effort. In this work, we extend the Green Simulation (GS) to quantify input uncertainty with only R replications by reweighting the R simulation outputs with the sample-path likelihood ratios (SPLRs). For each bootstrapped input distribution, SPLR is the ratio of the product of probability distribution functions at the inputs generated within the replication given the bootstrapped distribution and the distribution used to run the simulation. We propose two asymptotically valid GS-based confidence intervals for the expected simulation output under the true input distribution and analytically and empirically show they have efficiency gains compared to the nominal bootstrap-based approach.

Keywords: *Likelihood ratio, Input uncertainty, Sample path*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-151

TRADE-IN STRATEGIES IN DUAL-CHANNEL CLOSED-LOOP SUPPLY CHAIN WITH REMANUFACTURING

Lu Xiao[1], Kwai-Sang Chin[2], Xian-jia Wang[3], Zhen-song Chen[4]

[1] School of Management, Guangdong University of Technology and E-mail: xiaolu2008happy@gdut.edu.cn
[2] Department of Systems Engineering and Engineering Management, City University of Hong Kong and E-mail: mekschin@cityu.edu.hk
[3] School of Economics and Management, Wuhan University and E-mail: wangxj@whu.edu.cn
[4] School of Civil Engineering, Wuhan University and E-mail: zschen@whu.edu.cn

Abstract

As a tool for price discrimination and used products collection, trade-in is widely used in durable goods market. In this paper, we aim to derive conditions for manufacturer and retailer to implement trade-in policies voluntarily in traditional retail channel and offline/online dual-channel closed-loop supply chain with remanufacturing, respectively.

We apply game theory to construct traditional retail channel model and dual-channel model under two scenarios: without trade-in and with trade-in. We derive the equilibrium prices and optimal trade-in rebates for the manufacturer and retailer, along with the conditions to be met by the manufacturer and retailer to implement trade-in policies voluntarily in the two channel models. We also compare the two channel formats from the perspectives of trade-in quantity, retailer's profit and manufacturer's profit to identify which channel structure is optimal for trade-in policy.

Based on the results found in this study, we identify two areas for future research. The first extension of this research could be that the remanufactured products are different from new products in product quality and sale price. The second potential extension of this paper should investigate that the collected used products have different quality levels and different remanufacturing costs.

Keywords: *Closed-Loop Supply Chain, Remanufacturing, Dual-Channel, Trade-in*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-152

DETECTING BURSTS IN WATER PIPE SYSTEM USING FOURIER BASIS EXPANSION

Yinwei Zhang[1], Kevin Lansey[2], Jian Liu[3]

[1] Department of Systems & Industrial Engineering, the University of Arizona and E-mail: zhangyinwei@email.arizona.edu
[2] Department of Civil Engineering and Engineering Mechanics, the University of Arizona and E-mail: lansey@email.arizona.edu
[3] Department of Systems & Industrial Engineering, the University of Arizona and E-mail: jianliu@email.arizona.edu

Abstract

Bursts in water pipe systems (WPSs) are a special type of short-term, high-flow water loss that can be a significant component of a system's water balance. Since WPSs are usually deployed underground, bursts are difficult to be detected before their catastrophic results are observed on the ground surface. Continuous hydraulic data streams collected from automatic meter reading and advanced metering infrastructure systems make it possible to detect bursts in WDS based on data analytics. Thus, it is of great importance to develop analytical algorithms to detect bursts effectively. Researchers adopted conventional methods, such as EWMA and CUSUM control charts, to detect bursts based on the water flow or pressure data, and their statistical difference between normal conditions, for WPS without bursts, and anomaly conditions that with bursts. However, these methods are ineffective with unacceptable false alarm rate. This is because the continuously collected hydraulic data are highly autocorrelated and the "normal conditions" from historical data are neither reliable nor justifiable. In this research, the authors propose a burst detection method based on the functional analysis of hydraulic data with Fourier basis expansion. Fourier basis expansion is preferred due to its periodicity and the coefficients of different frequencies can be obtained by fast Fourier transform with small computation complexity. Furthermore, the identification of burst-related frequencies can be justified by incorporating domain hydraulic knowledge. The model coefficients from Fourier basis expansion are used to build a Hotelling-T^2 control chart to detect the bursts. The proposed method is evaluated in case study with high-fidelity simulation of Austin WPS and the result shows that it outperforms traditional control charts. However, this method is not able to estimate the starting time of the detected burst and the authors will solve this problem in the future work.

Keywords: *Control chart, Hydraulic data, Model coefficient, False alarm, Average run length*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-153

SLICED LATIN HYPERCUBE DESIGNS WITH ARBITRARY RUN SIZES

Jin Xu[1], Xu He[2], Xiaojun Duan[3], Zhengming Wang[4]

[1] College of Science, National University of Defense Technology and E-mail: xujin_nudt@163.com
[2] Academy of Mathematics and Systems Science, Chinese Academy of Sciences and E-mail: hexu@amss.ac.cn
[3] College of Liberal Arts and Sciences, National University of Defense Technology and E-mail: xj_duan@163.com
[4] College of Advanced Interdisciplinary Research, National University of Defense Technology and E-mail: wzm@nudt.edu.cn

Abstract

Latin hypercube designs achieve optimal univariate stratifications and are useful for computer experiments. Sliced Latin hypercube designs are Latin hypercube designs that can be partitioned into smaller Latin hypercube designs. In this work, we give the first construction of sliced Latin hypercube designs that allow arbitrarily chosen run sizes for the slices. we also provide an algorithm to reduce correlations of our proposed designs.

Keywords: *Computer experiment, Numerical integration, Variance reduction*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-155

THE ESTIMATION OF RISK-GENE NETWORKS AND THE PREDICTION OF SURVIVAL RISKS

Minhyeok Lee[1], Sung Won Han[2], Junhee Seok[3]

[1] School of Electrical Engineering, Korea University and E-mail: suam6409@korea.ac.kr
[2] School of Industrial Management Engineering, Korea University and E-mail: swhan@korea.ac.kr
[3] School of Electrical Engineering, Korea University and E-mail: jseok14@korea.ac.kr

Abstract

Over the last decades, many attempts have been made to apply network-based approaches to biomedical data. Gene networks, which are a successful example of such attempts, have shown numerous applications to illustrate relations between genes and phenotypes. However, there are few studies that apply the gene networks to survival analysis, which is one of the main interests in biomedical research. Furthermore, the existing studies using gene networks have limitations since they employ prior knowledge for network construction. In such approaches, selecting proper prior knowledge can be another problem and applying the approaches to rare diseases is not straightforward due to the lack of prior knowledge for the diseases. In this study, we propose a data-driven risk-gene network, and a survival risk prediction method through the network. In the proposed method, observed data points are partially projected to the risk-gene network in order to reduce variances intrinsic in datasets. The risk-gene network is applied to a low-grade glioma dataset and produces a hypothesis of a relationship between genetic biomarkers of low-and high-grade glioma. In addition, it is demonstrated that the proposed method shows superior prediction performance compared to conventional methods.

Keywords: *network estimation*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-157

PREDICTIVE MODELS FOR SHORT-TERM PASSENGER FLOW FORECASTING IN METRO

Yang Zhao[1], Yuxin He[2], Kwok Leung Tsui[3]

[1] School of Data Science, City University of Hong Kong and E-mail: yangzhao9-c@my.cityu.edu.hk
[2] School of Data Science, City University of Hong Kong and E-mail: yuxinhe2-c@my.cityu.edu.hk
[3] School of Data Science, City University of Hong Kong and E-mail: kltsui@cityu.edu.hk

Abstract

Forecasting short-term traffic flow has been a critical topic in transportation research for decades, which aims to facilitate dynamic traffic control proactively by monitoring the present traffic and foreseeing its immediate future. In this paper, we focus on forecasting short-term passenger flow at subway stations by utilizing the data collected through Automatic Fare Collection (AFC) system along with various external factors, where passenger flow refers to the volume of arrivals at stations during a given period of time. Along this line, we propose a data-driven three-stage framework for short-term passenger flow forecasting, consisting of traffic data profiling, feature extraction, and predictive modelling. We investigate the effect of temporal and spatial features as well as external weather influence on passenger flow forecasting. Moreover, using a real data set collected from Shenzhen AFC system, we conduct extensive experiments for methods validation, feature evaluation and data resolution demonstration.

Keywords: *Short-term passenger flow forecasting, multivariate linear regression, SVR (Support Vector Regression), feature extraction, time series*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-158

DYNAMIC FIELD MONITORING BASED ON MULTITASK LEARNING IN SENSOR NETWORKS

Di Wang[1], Xi Zhang[2]

[1] Peking University and E-mail: di.wang@pku.edu.cn
[2] Peking University and E-mail: xi.zhang@pku.edu.cn

Abstract

Field monitoring serves as an important supervision tool in a variety of engineering domains. An efficient monitoring would trigger an alarm timely once it detects an out-of-control event by learning the state change from the collected sensor data. However, in practice, multiple sensor data may not be gathered appropriately into a database for some unexpected reasons, such as sensor aging, wireless communication failures, and data reading errors, which leads to a large number of missing data as well as inaccurate or delayed detection, and poses a great challenge for field monitoring in sensor networks. This fact motivates us to develop a multitask-learning based field monitoring method in order to achieve an efficient detection when considerable missing data exist during data acquisition. Specifically, we adopt a log likelihood ratio (LR)-based MCUSUM control chart given spatial correlation among neighboring regions within the monitored field. To deal with the missing data problem, we integrate a multitask learning model into the LR-based MCUSUM control chart in the sensor network. Both simulation and real case studies are conducted to validate our proposed approach and the results show that our approach can achieve an accurate and timely detection for an out-of-control state when a large number of missing data exist in the sensor database. Our model provides an effective field monitoring strategy for engineering applications to accurately and timely detect the products with abnormal quality during production and reduce product losses.

Keywords: *multitask learning, field monitoring*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-160

A MODIFIED LATIN HYPERCUBE SAMPLING BASED ON PRIOR INFORMATION

Xueqing Wang[1]

[1] College of Management and Economics, Tianjin University and E-mail: xueqing1012@tju.edu.cn

Abstract

Latin hypercube sampling (LHS) is widely used in industrial engineering. It is an important tool for product quality improvement. In the traditional Latin square sampling, engineers often arrange sample points in the feasible domain uniformly. However, in some complex conditions, the LHS may generate insufficient sample point in some subdomains where this response volatility is relatively large. In practice, engineers may have some prior information about the sub-domains where the response volatility is relatively large, named the interesting sub-domains. In order to make full use of these information, this paper employed the D-S evidence theory to fuse prior information from different sources/fields. Then we divide the feasible domain into different sub-domains and indicate the interesting sub-domains. For the sample placement, we put more points in these interesting sub-domains and less points in other sub-domains. Finally, we construct the model with the proposed sample points placement approach based on prior information. To compared with the traditional Latin square sampling, we can obtain more useful information with the same number of sample points. A simulation study is conducted to illustrate the proposed method (Modified Latin Hypercube Sampling, MLHS). The simulation study shows that the proposed method performs better than the traditional model in MSE, MaxE and StdE.

Keywords: *sample placement, the interesting sub-domains, D-S evidence theory, prior information, Latin hypercube sampling*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-161

ARE REPORTED LIKELIHOOD RATIOS WELL CALIBRATED?

Jan Hannig[1]

[1] University of North Carolina at Chapel Hill and E-mail: jan.hannig@unc.edu

Abstract

Many machine learning algorithms used in forensic practice have as their output Bayes factor sometimes called likelihood ratios. For example, it is not unusual to see a report that claims the DNA found at the crime scene is 1,000,000 times more likely under the assumption that the defendant is the source than under the assumption that someone other than the defendant is the source. In this talk we summarize existing approaches for examining the validity of likelihood ratio systems and discuss a new statistical methodology based on generalized fiducial inference for empirically examining the validity of such likelihood ratio claims.

Joint work with Hari Iyer at National Institute of Standards and Technology

Keywords: *Likelihood ratio, Callibration, Generalized fiducial inference*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-162

ADDITIVE FUNCTIONAL REGRESSION WITH DENSITIES AS RESPONSES

Kyunghee Han[1], Hans-Georg Mueller[2], Byeong Uk Park[3]

[1] Pennsylvania State University and E-mail: kyunghee.stat@gmail.com
[2] University of California, Davis and E-mail: mueller2433@gmail.com
[3] Seoul National University and E-mail: bupark@stats.snu.ac.kr

Abstract

We propose and investigate additive density regression, a novel additive functional regression model for situations where the responses are random distributions that can be viewed as random densities and the predictors are vectors. Data in the form of samples of densities or distributions are increasingly encountered in statistical analysis and there is a need for flexible regression models that accommodate random densities as responses. Such models are of special interest for multivariate continuous predictors, where unrestricted nonparametric regression approaches are subject to the curse of dimensionality. Additive models can be expected to maintain one-dimensional rates of convergence while permitting a substantial degree of flexibility. This motivates the development of additive regression models for situations where multivariate continuous predictors are coupled with densities as responses. To overcome the problem that distributions do not form a vector space, we utilize a class of transformations that map densities to unrestricted square integrable functions and then deploy an additive functional regression model to fit the responses in the unrestricted space, finally transforming back to density space. We implement the proposed additive model with an extended version of smooth backfitting and establish the consistency of this approach, including rates of convergence. The proposed method is illustrated with an application to the distributions of baby names in the United States.

Keywords: *Additive models, Functional data analysis, Random densities, Smooth backfitting*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-163

ASSESSING ELDERLY'S FUNCTIONAL BALANCE AND MOBILITY VIA ANALYZING DATA FROM WEARABLE ACCELEROMETER IN TIME UP AND GO TESTS

Lisha Yu[1], Kwok Leung Tsui[2]

[1] School of Data Science, City University of Hong Kong and E-mail: lishayu2-c@my.cityu.edu.hk
[2] School of Data Science, City University of Hong Kong and E-mail: kltsui@cityu.edu.hk

Abstract

Falls among elderly can lead to injuries, morbidity, and mortality. Poor balance has been cited as one of the key causal factors of such falls. Timely detection of balance impairment can help identify the elderly prone to falls and also trigger early interventions to prevent them. However, continuous balance monitoring requires extensive healthcare and clinical resources. In this paper, we developed a surrogate approach for assessing elderly's functional balance based on Short Form Berg Balance Scale (SFBBS) score. Eighty-five community-dwelling older adults (72.12±6.99 year) recruited from Taiwan participated in our study. Accelerometer data were collected from a waist-mounted tri-axial accelerometer while participants performed a timed up and go test. Clinically relevant features were extracted from the segmented accelerometer signals for fitting SFBBS predictive models. Regularized regression together with random-shuffle-split cross-validation was adapted to facilitate the development of the predictive models for automatic balance estimation. Our results demonstrated that combined clinical and sensor-based variables, together with proper modeling and cross-validation, achieved moderate-high predictive accuracy of SFBBS scores (mean MAE = 2.01 and mean RMSE = 2.55). The approach proposed herein may represent a more sensitive, specific and responsive means of automating the measurement of functional balance as an enhancement to the clinical care of elderly.

Keywords: *balance and mobility, gait instability, fall, elderly care, data mining, accelerometer*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-164

STATE-OF-CHARGE ESTIMATION OF LITHIUM-ION BATTERIES VIA LONG SHORT-TERM MEMORY

Fangfang Yang[1], Kwok-Leung Tsui[2]

[1] School of Data Science, City University of Hong Kong and E-mail: fangfyang2-c@my.cityu.edu.hk
[2] School of Data Science, City University of Hong Kong and E-mail: kltsui@cityu.edu.hk

Abstract

Accurate state-of-charge (SOC) estimation is critical for driving range prediction of electric vehicles and optimal charge control of batteries. In this paper, a stacked long short-term memory network is proposed to model the complex dynamics of lithium iron phosphate batteries and infer battery SOC from current, voltage, and temperature measurements. The proposed network is trained and tested using data collected from the dynamic stress test, US06 test, and federal urban driving schedule. The performance on SOC estimation is evaluated regarding tracking accuracy, computation time, robustness against unknown initial states and compared with results from the model-based filtering approach (unscented Kalman filter). Moreover, different training and testing data sets are constructed to test its robustness against varying loading profiles. Experimental results show that the proposed network well captures the nonlinear correlation between SOC and measurable signals and provides better tracking performance than the unscented Kalman filter. In case of inaccurate initial SOCs, the proposed network presents quick convergence to the true SOC, with root mean square errors within 2% and mean average errors within 1%. Moreover, the estimation time at each time step is sub-millisecond, making it appropriate for real-time applications.

Keywords: *state-of-charge estimation, lithium iron phosphate batteries, long short-term memory, recurrent neural network*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-165

HYPER DESIGN AND COMPREHENSIVE DESIGN METHOD - DESIGN CONCEPT AND APPLICATION TO TWIN ROTOR PAPER HELICOPTER –

Takenori Takahashi[1]

[1] Keio University and E-mail: takenori@sfc.keio.ac.jp

Abstract

Purpose of this Presentation
This paper proposes HD (Hyper Design) concept to discuss various types of designs integrally and suggests a concrete design method based on it. Design needs long time, high cost and many efforts for its success. Therefore, it is necessary to reach the goal when design starts. CDM (Comprehensive Design Method) is proposed as a method to reach the goal of design surely. Case study of Twin Rotor Paper Helicopter developed by author is introduced as application example of HD.

Findings
Robust Parameter Design is placed as a part in HD, and HD can handle many other designs such as Input-Output Design, Morphology Design and Negotiation Design. New design method HD visualizes the process of design clearly so that it enables the agreement of persons concerned reasonably.

Originality/Value of Presentation
MI (Model Identification) and RA (Regression Adjustment) play very important roles in CDM. MI is the approach that evolves the function of Screening Experiment highly. In other words, it chooses active factors and moreover clarifies the model structure. It chooses one among five candidates about model structure. The five candidates are 1st order model, product/interaction model, 2nd order model, 2nd order model without product/interaction and higher order model. RA is the last card to let design reach the goal surely when design doesn't achieve the target.

Design/Methodology/ Approach
MI clarifies the structure of the model using plural statistical inference such as estimation, test, regression analysis and ANOVA first. Detailed model is estimated by Augment Design or Optimal Design based on identified model structure obtained by previous procedure next. Finally, the design gets the optimal solution using Mathematical Programing based on the detailed model.
The solution is confirmed about whether it is achieved after design. When the solution isn't achieved, RA is used to achieve the targeted value certainly.

Keywords: *hyper design, comprehensive design method, model identification, regression adjustment, twin rotor paper helicopter*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-166

DEGRADATION MODELING BASED ON A WIENER PROCESS WITH RANDOM DEGRADATION RATE AND INITIAL VALUE

Hongyu Wang[1], Xiaobing Ma[2], Yu Zhao[3]

[1] School of Reliability and Systems Engineering, Beihang University and E-mail: wanghongyu28@126.com
[2] The Key Laboratory on Reliability and Environmental Engineering Technology; School of Reliability and Systems Engineering, Beihang University and E-mail: maxiaobing@buaa.edu.cn
[3] The Key Laboratory on Reliability and Environmental Engineering Technology; School of Reliability and Systems Engineering, Beihang University and E-mail: zhaoyu@buaa.edu.cn

Abstract

The Wiener process is a popular tool for degradation modeling. In previous studies on the Wiener process, random effects are often incorporated into the degradation rate to characterize the unit heterogeneity in a population. However, in some practical situations, due to the variation of raw materials and manufacturing processes, the initial degradation value besides the degradation rate also varies from unit to unit, which introduces an additional uncertainty to the degradation model. Moreover, it's often observed that the initial degradation value has a relationship with the degradation rate. Motivated by this practical issue, we propose a novel random-effects Wiener process model wherein the degradation rate and the initial value are treated as correlated random variables following a bivariate normal distribution. This new model not only inherits the idea of random degradation rates in existing studies, but also takes the uncertainty of the initial value and the correlation between the degradation rate and the initial value into consideration. Based on the proposed model, a parameter estimation procedure is developed by using the expectation maximization (EM) algorithm, and the lifetime distribution is derived subsequently. It's worth emphasizing that we give explicit expressions of estimated parameters in each iteration of the EM algorithm, and thus the calculation is fast. Finally, a practical study is provided to demonstrate the effectiveness of the proposed model and method, and the result shows that our model fits the degradation data better than the existing model which ignores the randomness of the initial degradation value.

Keywords: *Wiener process, random initial value, random degradation rate, expectation maximization (EM) algorithm*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-167

THE FOURTH INDUSTRIAL REVOLUTION AND A NEW PARADIGM: "RESILIENCE" TO A HYPER CONNECTED SOCIETY

Eui Hoon Lee[1]

[1] School of Civil Engineering, Chungbuk National University and E-mail: hydrohydro@chungbuk.ac.kr

Abstract

The fourth industrial revolution is based on cyber physical systems with big data, cloud computing, system integration, augmented reality, 3D printing and simulation. The network in the first, second and third industrial revolution is unidirectional top-down communication. In contrast, the network of the fourth industrial revolution is interconnection communication including bottom-up approach. It can be called hyper-connected network and a society with the hyper-connected network is called a hyper connected society. The hyper connected society with the fourth industrial revolution is coming. In the hyper connected society, it is very convenient because every component is connected, but there is a possibility of cascading failure. The application of resilience is essential for preventing cascading failure because the concept of resilience includes preparation, failure and recovery procedures. The concept of resilience is required prepare for future changes in the hyper connected society. Various fundamental researches for the application of resilience in urban drainage systems have been conducted.

Keywords: *Fourth industrial revolution, Hyper connected society, Cascading failure, Resilience*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-168

MULTIOBJECTIVE METER PLACEMENT MODEL FOR ROBUST WATER DISTRIBUTION SYSTEM PIPE BURST DETECTION

Donghwi Jung[1]

[1] Department of Civil Engineering, Keimyung University and E-mail: donghwiku@gmail.com

Abstract

Water distribution system (WDS) is to provide the required quantity of water to users with a satisfactory pressure and water quality. Because pipe in WDS is laid underground, its burst is often not detected until a collateral damage (a sinkhole) occurs. Pipe burst is the rupture of pipe wall from pipe deterioration, excessive pressure, and ground shifts from temperature change and earthquake. Therefore, prompt and accurate detection of pipe bursts improves WDS serviceability and help avoid potential water service outage.

During the last two decades, many meter-data-driven methods have been applied to detect WDS pipe bursts. A meter is to remotely measure and send pressure and pipe flow at a location in WDS. In order to provide informative data and finally increase pipe burst detectability, meters should be positioned at optimal locations. In addition, the optimal meter set should be composed to even detect them well under meter's mechanical failure conditions (e.g., battery out – sending no data). While several meter placement models have been proposed, few efforts have been devoted to consider the meter network's failures.

In this study, a multiobjective optimal meter placement (MOMP) model is introduced for robust WDS pipe burst detection. The proposed MOMP model is to (1) minimize the normalized total meter cost, (2) maximize the detection probability (DP), (3) minimize the rate of false alarm (RF), and (4) maximize the meter set's mechanical reliability. The optimal ratio between the pressure and pipe flow meters is also determined from the model given a predefined number of meters. The proposed model is demonstrated on the Austin network. It was confirmed from the results that a nonlinear trade-off relationship exists between DP and RF, and high mechanical reliability could be achieved by including pressure meters only in the optimal meter set.

Keywords: *Water distribution system, Pipe burst detection, Meter placement, Mechanical failure, Multiobjective model, Pressure and pipe flow*

Acknowledgment: *This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2018R1C1B5045011)*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-169

DEGRADATION IN COMMON DYNAMIC ENVIRONMENTS

Ye Zhisheng[1]

[1] ISEM, National University of Singapore and E-mail: yez@nus.edu.sg

Abstract

Degradation studies are often used to assess reliability of products subject to degradation-induced soft failures. Because of limited test resources, several test subjects may have to share a test rig and have their degradation measured by the same operator. The common environments experienced by subjects in the same group introduce significant inter-individual correlations in their degradation, which is known as the block effect. In the present paper, the Wiener process is used to model product degradation, and the group-specific random environments are captured using a stochastic time scale. Both semiparametric and parametric estimation procedures are developed for the model. Maximum likelihood estimations of the model parameters for both the semiparametric and parametric models are obtained using an inexact block coordinate descent algorithm. Performance of the maximum likelihood estimators is validated through large sample asymptotics and small sample simulations. The proposed models are illustrated by an application to lumen maintenance data of blue light-emitting diodes.

Keywords: *Wiener process*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-170

ROBUST ESTIMATION OF COMPONENT RELIABILITY BASED ON SYSTEM LIFETIME DATA

Xiaojie Zhu[1], Hon Keung Tony Ng[2], Ping Shing Chan[3]

[1] Department of Statistical Science, Southern Methodist University and E-mail: xiaojiez@mail.smu.edu
[2] Department of Statistical Science, Southern Methodist University and E-mail: ngh@mail.smu.edu
[3] Department of Statistics, The Chinese University of Hong Kong and E-mail: benchan@cuk.edu.hk

Abstract

In this paper, we consider the estimation of model parameters based on system lifetime data with known system signature using the minimum density divergence estimation method. Different estimation procedures based on the minimum density divergence estimation method using are proposed. A Monte Carlo simulation study is used to evaluate the performance of those proposed procedures and compare with the maximum likelihood estimation method under different contaminated models. We have shown that the proposed estimation procedures are robust to contamination and model misspecification. A numerical example is presented to illustrate the proposed methodology.

Keywords: *Maximum likelhood estimation, Minimum density divergence estimation, Monte Carlo simulation, System signature*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-171

INTEGRATIVE DEEP LEARNING FOR IDENTIFYING DIFFERENTIALLY EXPRESSED(DE) BIOMARKERS

Jayeon Lim[1], Sunghwan Kim[2], Jiyeon Kim[3]

[1] Department of Applied Statistics, Konkuk University, Seoul, South Korea and E-mail: jayun7433@konkuk.ac.kr
[2] Department of Applied Statistics, Konkuk University, Seoul, South Korea and E-mail: shkim1213@konkuk.ac.kr
[3] Department of Statistics, Keimyung University, Daegu, South Korea

Abstract

As a large amount of genetic data were accumulated, an effective analytical method and a significant interpretation are required. Recently, various methods of machine learning have emerged to process genetic data. In addition, machine learning analysis tools with statistical models were proposed. In this study, we propose adding an integrated layer to the deep learning structure, which would allow for effective analysis of genetic data as well as discovery of significant biomarkers of diseases. We conducted a simulation study to compare the proposed method with meta-logistic regression and meta-SVM methods. The objective function with Lasso penalty is used for parameter estimation and Youden J index is used for model comparison. Simulation results show that proposed method is more robust for variance of the data than meta-logistic regression and meta-SVM methods Also, we conducted real data (breast cancer data(TCGA)) analysis. From gene set enrichment analysis, we obtained that TCGA multiple omics data is significantly enriched pathways which contain information related to breast cancer. Therefore, it is expected that the proposed method will help discovering biomarkers in the future.

Keywords: *DE biomarkers, Deep-learning, Backpropagation, genetic data, enrichment analysis, breast cancer*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-172

REVISIT TO BAYESIAN MODEL AVERAGING IN FORECASTING INFLUENZA WITH GOOGLE SEARCH QUERIES

Hao Pan[1], Yang Zhao[2], Kwok Leung Tsui[3]

[1] School of Data Science, City University of Hong Kong, China and E-mail: haopan2-c@cityu.edu.hk
[2] School of Data Science, City University of Hong Kong, China and E-mail: yang.zhao@my.cityu.edu.hk
[3] School of Data Science, City University of Hong Kong, China and E-mail: kltsui@cityu.edu.hk

Abstract

Forecasting problem is one of the crucial considerations in the subject of epidemiology, especially in dealing with the influenza data with internet search queries. Individual learning approaches, such as time series model, e.g., ARIMA, and regression methods, e.g. PCR, LASSO and Elastic-net play essential roles in solving forecasting problems. It's known that there is no single method perform best on all type of data. Those all basic learning models are bearing their own unique advantages and limitations. Under this circumstance, there has been an increasing interest in ensemble learning approaches to data fusion and model assimilation for achieving better forecasting performance, typically by the Bayesian model averaging. In this work, we focus on the statistical properties and theoretical issues Bayesian model averaging around the forecasting problem. Different prior information on the basic learning predictors leads to various type of information criterion. Additionally, we conduct a comparative study among various forecasting methods using the influenza-like-illness in general outpatient clinics (ILI-GOPC) data in the U.S. and Hong Kong.

Keywords: *Bayesian Model Averaging, Ensemble Learning, Forecasting, Influenza*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-173

PROBABILISTIC WEAR LIFE-CENTERED MAINTENANCE DESIGN OF AUTOMATION SYSTEM WHEELS

Mingu Chang[1], Yeongmin Yoo[2], Jongsoo Lee[3]

[1] Department of Mechanical Engineering, Yonsei University and E-mail: mingu.chang@yonsei.ac.kr
[2] Department of Mechanical Engineering, Yonsei University and E-mail: yym9514@yonsei.ac.kr
[3] Department of Mechanical Engineering, Yonsei University and E-mail: jleej@yonsei.ac.kr

Abstract

Modern factory automation is efficiently managed according to the precise plan of the central control system. Especially, in case of the transport equipment which plays the role of connecting the facilities, it is necessary to reach the destination at the correct time and supply the material. If there is a problem with such transportation equipment, the supply of the material will be stopped and the productivity will be lowered. The transport equipment is driven by the rolling of the Driving wheels, and when the wear of the wheels occurs over the limit, it causes an error in the control. Since the control problems caused by abrasion are caused by gradual wear without showing any pre-failure symptoms, it is very important to prevent wear problems by maintenance through prediction of wear life.

In this paper, we propose a Bayesian methodology for estimating the remaining useful life of an automation equipment operating under different environmental conditions without reference and full degradation data. We used the Meta model to estimate the Prior distributions, which are essential for applying the Bayesian method. For this purpose, the wear model was improved and wear data were generated with time according to the shape of the wheel due to wear. In addition, the uncertainty propagation data is generated considering the distribution of effective components and the prior distribution of the degradation model parameters is estimated with this data. Based on the prior distribution and the real wheel wear data, the Bayesian inference is used for statistical wear life prediction and probability distribution of EOL(End Of Life) is derived. Using the predicted wear reliability, a multi - objective function for the number of maintenance times and cost was specified, and the optimal maintenance cycle was derived by applying the Non Sorting Genetic Algorithm.

Keywords: *Remaining Useful life, Wheel wear, Automation System, Bayesian Inference, Non Sorting Genetic Algorithm, Maintenance Interval*

Acknowledgement: *This research is supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (2017R1A2B4009606).*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-174

STATISTICAL UNCERTAINTY MODELING AND RELIABILITY ASSESSMENT USING BAYESIAN NEURAL NETWORKS IN RELIABILITY-BASED DESIGN OPTIMIZATION

Juhee Lim[1], Jongsoo Lee[2]

[1] School of Mechanical Engineering, Yonsei University, Korea and E-mail: juheelim@yonsei.ac.kr
[2] School of Mechanical Engineering, Yonsei University, Korea and E-mail: jleej@yonsei.ac.kr

Abstract

For approximate optimization process, meta-modeling methods are used to reduce the computational cost and to increase the efficiency of the optimization process in complex engineering systems. Metaheuristic optimization is generally performed with various types of surrogate models such as polynomials, kriging, and neural networks. The most commonly used response surface method based metamodel has difficulties in that a large number of data is required depending on the linearity, the accuracy of the data varies depending on the degree, and over-fitting problems can occur. On the other hand, one of the main advantages of the neural network metamodel is that it can handle non-linearity and complex interactions between variables when using a sufficient number of hidden layers. Therefore, a neural network metamodel can be used instead of a simulation in a decision-making process where the results must be obtained very quickly and in computationally intensive reliability-based design process. From a probability theory perspective, it is unjustifiable to use single point-estimates as weights to base any classification on. In this respect, Bayesian inference for neural networks is characterized by providing uncertainty estimates through parameters of a probability distribution type with statistical modeling and calculating the posterior distribution of weights given training data. The Bayesian neural network is more robust to over-fitting problems and is easy to learn from a small number of data sets. In this study, we perform high-fidelity Bayesian neural networks including statistical modeling based reliability-based optimization. After validating the simulation results for the numerical test problem, it was applied to the engineering problem that optimizes the objective function according to the reliability constraint.

Keywords: *uncertainty, statistical modelling, metamodel, reliability based design optimization, Bayesian neural network*

Acknowledgement: *This research is supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (2017R1A2B4009606)*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-176

REPEATED PARAMETER ESTIMATION WITH CONTROLLED RANDOM SEARCH FOR AN EXPENSIVE AND HIGHLY NONLINEAR MODEL

Hyungjin Kim[1], Chuljin Park[2]

[1] Department of Industrial Engineering Hanyang University and E-mail: jins731@naver.com
[2] Department of Industrial Engineering Hanyang University and E-mail: parkcj@hanyang.ac.kr

Abstract

The problem of this study is repeatedly estimating parameters of a highly nonlinear model which interprets the controllable (or independent) variables and the observable (or dependent) variables. It is noted that the model has no explicit form and the values from the model are often obtained by intensive numerical calculations or complicated simulation. In the problem, we assume that the model parameters are correlated and sampled from a multivariate normal distribution and the mean vector and covariance matrix of the parameter distribution are supposable (e.g., pre-known or estimated). Such parameter distribution is referred to as the supposed underlying distribution. In order to solve the problem, we develop a new framework, repeated parameter estimation with controlled random search, which primarily adapts two key concepts, (i) change of coordinates and (ii) solution sampling based on the supposed underlying distribution of the parameters, to the structure of the controlled random search. Specifically, we evaluate and optimize the objective function not over the original coordinate system of the parameters but over a linearly transformed coordinate system. Moreover, the algorithm filters a solution sampled by a heuristic sampling strategy based on the acceptance-rejection method, whereby the function is evaluated at a sampled solution only when the solution is accepted according to the supposed underlying distribution. As a result, the proposed algorithm efficiently achieves small absolute errors regardless of the type of the model and the scale of each parameter. Numerical results show that algorithm performs well when it is applied to simulation studies with two nonlinear regression models and to a real example of measuring the critical dimensions of a 2-dimensional high aspect ratio structure of a wafer in semiconductor manufacturing.

Keywords: *nonlinear parameter estimation, parameter correlation, controlled random search, optical critical dimension, semiconductor manufacturing*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-177

THE RECONSTRUCTION APPROACH: FROM INTERPOLATION TO REGRESSION, CLASSIFICATION, AND NUMERICAL COMPUTING

Shifeng Xiong[1]

[1] Academy of Mathematics and Systems Science, Chinese Academy of Sciences and E-mail: xiong@amss.ac.cn

Abstract

We introduce an interpolation-based method, called the reconstruction approach, for function parameterization and estimation. This approach uses an interpolator to parameterize the unknown function with its values at finite knots, and then estimates these values through optimizing some criteria. In statistical regression and classification problems, the criteria can be selected as regularized empirical risk functions, and the reconstruction approach can include popular methods such as kernel ridge regression and kernel support vector machines as its special cases. It is shown that, the reconstruction idea not only provides different angles to look into existing methods, but also produces new effective experimental design and estimation methods for nonparametric models. We also show that the proposed method can be used in numerical computing, including solving PDE's.

Keywords: *Gaussian process, kernel method, Kriging, nonparametric estimation, space-filling design*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-178

PHM FRAMEWORK FOR SMART MANUFACTURING SYSTEMS

Daeil Kwon[1]

[1] Konkuk University and E-mail: dkwon@konkuk.ac.kr

Abstract

This paper presents a PHM framework for manufacturing systems. While PHM is emerging as an enabling technology to avoid system failure, its practical use for industrial applications is challenging especially for small and medium-sized enterprises due to the lack of internal expertise, time and resources for research and development. The PHM framework provides a readily usable and accessible guideline for application to manufacturing systems. Across manufacturing systems and their kinds, six common core modules were identified; and appropriate PHM approaches for each core module were investigated with respect to the amount of available data and domain knowledge. As a case study, fault prognosis was conducted on a manufacturing system by following the PHM framework.

Keywords: *PHM, framework, smart manufacturing*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-179

COMPUTER EXPERIMENTS WITH BINARY TIME SERIES WITH APPLICATIONS IN CELL BIOLOGY: MODELING, ESTIMATION, AND CALIBRATION

Ying Hung[1]

[1] Rutgers University and E-mail: yhung@stat.rutgers.edu

Abstract

Computer experiments have become ubiquitous in various applications from rocket injector designs to weather forecasts. Although extensive research has been devoted in the literature, computer experiments with binary time-series outputs have received scant attention. Motivated by the analysis of a class of cell adhesion experiments, we introduce a new emulator, as well as a new calibration framework for binary time-series outputs. More importantly, we provide their theoretical properties to ensure the estimation performance in an asymptotic setting. The application to the cell adhesion experiments illustrates that the proposed emulator and calibration framework not only provide an efficient alternative for the computer simulation, but also reveal important insight on the underlying adhesion mechanism, which cannot be directly observed through existing methods.

Keywords: *Computer experiments*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-180

MULTI-RESOLUTION FUNCTIONAL ANOVA FOR LARGE-SCALE, MANY-INPUT COMPUTER EXPERIMENTS

Chih-Li Sung[1], Wenjia Wang[2], Matthew Plumlee[3], Benjamin Haaland[4]

[1] Michigan State University and E-mail: sungchih@msu.edu
[2] Statistical and Applied Mathematical Sciences Institute and E-mail: 1992wangwenjia@gmail.com
[3] Northwestern University and E-mail: mplumlee@northwestern.edu
[4] University of Utah and E-mail: Benjamin.Haaland@hci.utah.edu

Abstract

The Gaussian process is a standard tool for building emulators for both deterministic and stochastic computer experiments. However, application of Gaussian process models is greatly limited in practice, particularly for large-scale and many-input computer experiments that have become typical. We propose a multi-resolution functional ANOVA model as a computationally feasible emulation alternative. More generally, this model can be used for large-scale and many-input non-linear regression problems.

An overlapping group lasso approach is used for estimation, ensuring computational feasibility in a large-scale and many-input setting. New results on consistency and inference for the (potentially overlapping) group lasso in a high-dimensional setting are developed and applied to the proposed multi-resolution functional ANOVA model. Importantly, these results allow us to quantify the uncertainty in our predictions.

Numerical examples demonstrate that the proposed model enjoys marked computational advantages. Data capabilities, both in terms of sample size and dimension, meet or exceed best available emulation tools while meeting or exceeding emulation accuracy.

Keywords: *computer experiments, non-linear regression, large-scale, many-input, overlapping group lasso*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-181

SIMULTANEOUS OPTIMIZATION OF CORRELATED QUALITY AND LIFETIME CHARACTERISTICS IN CONSTRAINED RANDOMIZATION EXPERIMENTS

Shanshan Lv[1], Xueqing Wang[2]

[1] School of Economics and Management, Hebei University of Technology and E-mail: shanshanlv@hebut.edu.cn
[2] College of Management and Economics, Tianjin University and E-mail: xueqing1012@tju.edu.cn

Abstract

Reliability is usually defined as quality over time, and it is one of the most important aspects of quality. In many industrial experiments, it is common to observe both a static quality response and the corresponding reliability response that need to be optimized simultaneously. It is well recognized that, in practice, the quality and reliability characteristics are more likely to be correlated. Furthermore, the quality and reliability experiments are usually not completely randomized design due to the choosing of some experimental protocols, e.g., subsampling, blocking, split - plot, etc. This paper proposes a new three stage method to optimize correlated quality and lifetime responses simultaneously. First, in order to consider the effects of experimental protocols, the mean of quality response and the pth quantile of lifetime response are regarded as new responses, respectively. Then the paper adopts the seemingly unrelated regression (SUR) method to simultaneously consider the correlation between the mean quality characteristic and quantile lifetime characteristics. Finally, the mean quality response and quantile lifetime are optimized simultaneously. We illustrate the implementation of the methodology through a case study. The proposed approach in this paper gives practitioners an easy way to simultaneously optimize correlated quality and reliability characteristics in constrained randomized experiments.

Keywords: *multi-response optimization, quality characteristic, lifetime characteristic, correlation, random effects*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-183

A NEW UNCERTAIN DECISION TREE METHOD FOR CLASSIFYING SEMICONDUCTOR WAFER DEFECT BASED ON MULTIPLE WAFER MAPS

Byunghoon Kim[1], Young-Seon Jeong[2], Seung Hoon Tong[3], Myong-Kee Jeong[4]

[1] Hanyang University and E-mail: byungkim@hanyang.ac.kr
[2] Chonnam National University and E-mail: young.jeong@jnu.ac.kr
[3] Samsung Electronics Company Ltd. and E-mail: shtong@samsung.com
[4] Rutgers University and E-mail: mjeong@rutgers.edu

Abstract

Spatial defect patterns on semiconductor wafers include useful information about problems during fabrication of integrated circuits. Therefore, human being inspectors manually classify and summarize the defect chip information of the wafers to help the root cause analysis of the failure. To automate the inspection, many approaches have been developed to classify defect patterns in wafers. Most of the existing studies, however, focused on a single wafer bin map. We propose a novel approach to classify defect patterns in dynamic random-access memory (DRAM) wafers based on multiple wafer maps. To classify distinct defect patterns of multiple wafer maps, we propose a generalized uncertain decision tree model considering correlations between uncertain features. The experimental results show that the proposed approach performs better than the existing methods reported in the literature.

Keywords: *Data mining, Semiconductor manufacturing*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-184

A GENERALIZED STATISTICAL PARAMETER FOR SYSTEM HEALTH MONITORING AND ITS PARALLEL TO DEEP LEARNING

Dong Wang[1], Lifeng Xi[2]

[1] Department of Industrial Engineering and Management; The State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University and E-mail: dongwang4-c@sjtu.edu.cn
[2] Department of Industrial Engineering and Management; The State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University and E-mail: lfxi@sjtu.edu.cn

Abstract

System health monitoring aims to use monitoring data collected from a system to assess system health conditions over time and prevent unexpected system failures. Some famous statistical parameters including kurtosis, negative entropy, Gini index and smoothness index are able to characterize the impulsiveness of repetitive transients caused by early faults and thus they are handcraft statistical parameters for system health monitoring. In this presentation, we will show our new progresses toward system health monitoring and some new findings as follows: (1) it is discovered that all of the aforementioned statistical parameters fall into a generalized statistical parameter called the sum of weighed normalized square envelope (SWNSE). The main difference among these statistical parameters is that different weights are applied to normalized square envelope; (2) the main difference between kurtosis and negative entropy is whether the logarithm operator is adopted to weight normalized square envelope; (3) SWNSE can inspire readers to design more and more advanced statistical parameters for system health monitoring if new weights are properly designed; (4) a data-driven method based on convex optimization is proposed to optimize the weights of SWNSE so as to get an optimized statistical parameter for system condition monitoring; (5) bearing run to failure degradation data collected from a machine are used as a demonstration to verify the effectiveness of SWNSE for bearing health monitoring. Moreover, in view of optimized weights obtained from the bearing run to failure degradation data, a new analytical statistical parameter for bearing condition monitoring is accordingly proposed. Results showed that the optimized statistical parameter and the new analytical parameter have a highly similar bearing degradation trend with the aforementioned famous statistical parameters. Moreover, it should be noted that, based on our new ideas, we can design a unique statistical parameter for health monitoring of a new system/component; (6) a high similarity between SWNSE and neural networks in deep learning is discovered. The findings in this presentation can be easily extended to healthcare monitoring, machine/component monitoring, wearable sensors based action recognition, etc., in which characterizing impulsive transients is of great concern. Consequently, the proposed SWNSE has potentially wide applications in heath monitoring.

Keywords: *handcraft parameters, health monitoring, generalized statistical parameter, deep learning, optimized weights, weighed normalized square envelope*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-185

A CNN-BASED FAULT DETECTION METHOD USING VIBRATION VIDEO

Bayu Adhi Tama[1], Hyunsuk Huh[2], Kyungho Sun[3], Seungchul Lee[4]

[1] POSTECH and E-mail: bayuat2802@postech.ac.kr
[2] POSTECH and E-mail: kingking012@postech.ac.kr
[3] Korea Institute of Machinery and Material (KIMM) and E-mail: sunkh@kimm.re.kr
[4] POSTECH and E-mail: seunglee@postech.ac.kr

Abstract

In the era of the fourth industrial revolution (Industry 4.0) and Internet of Things, a huge amount of data is enormously accumulated and analyzed from mechanical equipment. Most conventional machine learning algorithms rely upon laboriously manual feature engineering, e.g. hand-crafted feature representation and extraction. In contrast to this, CNN deals with a feature learning approach, where feature representations are discovered from raw data in an automatic manner. The presentation will address an automatic machine monitoring technique by utilizing convolutional neural network (CNN) with class activation maps. A class activation map is able to distinguish the fault region precisely, enable us to effectively localize the fault. The objective of this presentation is to seek how CNN can be used to characterize a real-world vibration video data which represents normal and fault condition. Furthermore, this presentation will give a new insight into the development of automatic machine's condition monitoring for industrial applications.

Keywords: *Vibration Video, CNN, CAM, Deep Learning, Fault Detection*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-186

REVEALING HOUSEHOLD CHARACTERISTICS USING CONNECTED HOME PRODUCTS

Sungil Kim[1]

[1] Ulsan National Institute of Science and Technology and E-mail: sungil.kim@unist.ac.kr

Abstract

Recent technological advances have helped make our homes more intelligent and responsive to our needs. In this context, analyzing data from smart home products to gain insights is becoming increasingly important. This paper proposes an analytical framework to reveal the characteristics of households using event logs from smart door lock systems. The analytical framework uses constraint satisfaction problems to enable streaming event log analysis to solve the problems of overlapping classes and a lack of information concerning the truth class. The proposed method was applied to two datasets: one consisting of door-lock log data from 40 households and the other of time-use survey data from more than 10,000 households in South Korea. The results verify its effectiveness in terms of estimating the number of occupants in a household. The performance of the proposed method was compared with that of a naive clustering approach in terms of mean squared error.

Keywords: *internet of things, smart door lock*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-187

CLASS DESCRIPTION NETWORK

Hyungrok Do[1], Youngdoo Son[2], Seoung Bum Kim[3]

[1] School of Industrial Management Engineering, Korea University and E-mail: hyungrok@korea.ac.kr
[2] Department of Industrial and Systems Engineering, Dongguk University-Seoul and E-mail: youngdoo@dongguk.edu
[3] School of Industrial Management Engineering, Korea University and E-mail: sbkim1@korea.ac.kr

Abstract

We propose a neural network-based binary classification algorithm, named class description network (CDN). The CDN constructs two compact hyperspheres in the feature space such that each hypersphere encloses the instances of one class while excluding those of the other class from the hypersphere. Moreover, the CDN simultaneously learns useful feature representations of the data so that the hyperspheres effectively separate two classes. Unlike softmax-based neural network classifiers, the CDN not only classifies two target classes but also recognizes unseen classes that are not included in the training data. To verify the binary classification and unseen class recognition performances, we conducted numerical experiments on MNIST, Fashion MNIST, and GTSRB image benchmark data sets. The results confirmed that the proposed CDN can identify unseen instances while successfully performing the binary classification task.

Keywords: *neural network, deep learning, data description, unseen class, binary classification*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-188

GAUSSIAN PROCESS MODELS FOR COMPUTER EXPERIMENTS WITH NON-QUANTITATIVE INPUTS

Xinwei Deng[1]

[1] Deparment of Statistics, Virginia Tech and E-mail: xdeng@vt.edu

Abstract

Computer experiments with non-quantitative input factors recently attracts more attentions in different engineering applications. Surrogate modeling for such experiments is not yet completely resolved due to the discrete nature of such non-quantitative inputs. In this talk, we present some recent development on Gaussian process to model computer experiments with non-quantitative factors. The proposed methods consider some novel additive Gaussian process models to enable flexible structure of incorporating qualitative factors in modeling the complex systems of computer experiments. The merits of the proposed method are illustrated by both numerical examples and real-data applications.

Keywords: *Computer Experiments, Gaussian Process, Additive Structure*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-189

MODELING IN-PLANE DEVIATIONS OF SHAPES TO COME BASED ON PRIOR DEVIATION FEATURES IN ADDITIVE MANUFACTURING SYSTEMS

Arman Sabbaghi[1]

[1] Purdue University Department of Statistics and E-mail: sabbaghi@purdue.edu

Abstract

Geometric shape deviation models constitute an important component in dimensional accuracy control for additive manufacturing (AM) systems. The low-volume production inherent in AM results in the fundamental challenge of specifying deviation models for new classes of shapes to come in an AM system without data on their deviations. A methodology that can make full use of prior data and knowledge on deviation features for previously manufactured, and different, shapes is necessary to enable deviation modeling for new shapes to come in an AM system. We present a straightforward framework that addresses this requirement. Our framework specifies a new shape's deviation model according to a decomposition of its computer-aided design model into convex and concave blocks, and simple, interpretable combinations of prior deviation feature models from previous shapes for the different blocks. We illustrate our framework for a freeform shape manufactured under stereolithography.

Keywords: *3D printing, Cyber-Physical Systems, Prescriptive deviation modeling*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-190

FEASBLE DESIGN REGION IDENTIFICATION IN ADDITIVE MANUFACTURING VIA SURROGATE MODELLING OF DESIGN RULES

SungKu Kang[1], Xinwei Deng[2], Ran Jin[3]

[1] Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University and E-mail: kangsungku@vt.edu
[2] Department of Statistics, Virginia Polytechnic Institute and State University and E-mail: xdeng@vt.edu
[3] Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University and E-mail: jran5@vt.edu

Abstract

Additive Manufacturing (AM) provides great design flexibilities to enable the fabrication of highly personalized products. However, as the flexibilities has radically expanded the design space, design space exploration becomes very time-consuming to obtain personalized design for AM. While the challenge calls for feasible region identification methods to efficiently search complicated design space, existing efforts usually focus on the implementation of design rules to evaluate the feasibility of a single AM design. In this work, we propose the feasible design region identification method for AM via surrogate modelling of design rules. The proposed method is to first perform a space-filling design with adaptively placing more design points on the region where the evaluation of design rules takes much less time. Then surrogate models based on Gaussian processes are developed to identify feasible regions over the design space. The proposed method can minimize the time of evaluating design rules for accurate estimation of surrogate models, with the accommodation of continuous, categorical, and binary responses by adopting generalized Gaussian process model (GGPM). The performance of the proposed method is elaborated from the design of microbial fuel cell (MFC) anode manufactured with fused deposition modelling process. We expect that the proposed method can significantly improve the efficiency of personalized AM design candidate generation, thus contribute to enabling the personalized product realization cycle adapting to faster-changing markets.

Keywords: *Design for Additive Manufacturing, Design of Experiment, Feasible Design Region Identification, Generalized Gaussian Process Model (GGPM), Surrogate Modelling*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-191

PRIME: A PERSONALIZED RECOMMENDATION FOR INFORMATION VISUALIZATION METHODS VIA EXTENDED MATRIX COMPLETION

Xiaoyu Chen[1], Nathan Lau[2], Ran Jin[3]

[1] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg and E-mail: xiaoyuch@vt.edu
[2] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg and E-mail: nathan.lau@vt.edu
[3] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg and E-mail: jran5@vt.edu

Abstract

Information visualization is an important mode of human-machine collaboration (HMC) to acquire insights of complex datasets gathered from cybermanufacturing systems. However, effectiveness of any specific visualization is moderated by individual differences in knowledge, skills and abilities in perception and cognition in different contexts (e.g., tasks, environments, etc.). This research proposes a Personalized Recommendation for Information visualization Methods via Extended matrix completion model (PRIME), to recommend the optimal visualization designs for individual users in different contexts (e.g., different tasks). PRIME quantitatively models covariates (i.e., psychological and behavioral measurements) and subjective ratings (i.e., perceived complexity ratings) to predict ratings for users in contexts in real-time, and then adapting visualization based on the predicted ratings. It provides recommendations based on the decomposition of implicit similarity contained in covariates and explicit similarity contained among subjective ratings. In this study, Electroencephalography (EEG), eye movements, and user interaction logs with the system were collected when users were performing pre-defined tasks by using three different visualization designs. The results show that the PRIME can achieve satisfactory recommendation accuracy, especially in handling limited historical data and similarity among new users and new visualization designs. This capability should be invaluable for designing new generation of manufacturing visualization systems that adapt to users' mental status in real time. PRIME can support researchers in reducing the sample size requirements to quantify individual differences, and practitioners in adapting visualization designs as well as other adaptive systems to users' preference.

Keywords: *Adaptive systems, Electroencephalography, Eye tracking, Information visualization, Matrix completion, Recommender systems*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-192

BAYESIAN NETWORK FOR PROGNOSTICS OF MULTI-COMPONENT SYSTEMS

Seokgoo Kim[1], Nam Ho Kim[2], Joo-Ho Choi[3]

[1] Korea Aerospace University and E-mail: sgkim@kau.kr
[2] University of Florida and E-mail: nkim@ufl.edu
[3] Korea Aerospace University and E-mail: jhchoi@kau.ac.kr

Abstract

Numerous studies have been made in the field of prognostics and health management (PHM) to prevent unpredicted failure and reduce the unnecessary maintenance cost. Most of the existing studies, however, focus on the component-level diagnostics and prognostics. Although component-level PHM shows successful performance in real application, many operators are interested in system-level PHM because engineering systems usually consist of multiple components in practice. In system-level PHM, several components are connected and their failure have dependencies, which make system-level PHM more than a simple problem. In this paper, bayesian network (BN) based approaches are proposed to combine degradation information of components and construct the conditional probability of failure to estimate the system degradation. To demonstrate the effectiveness of the proposed approach, numerical example is introduced.

Keywords: *Bayesian Network, Causal relationship, Prognostics and Health Management*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-193

COMBINING EXTREME LEARNING MACHINE, SPARSE AUTO-ENCODER AND BACK PROPAGATION FOR HEALTH INDICATOR EXTRACTION

Taiwo Joel Omoleye[1], Kwok Leung Tsui[2]

[1] School of Data Science, City University of Hong Kong, Hong Kong and E-mail: jtomoleye3-c@my.cityu.edu.hk
[2] School of Data Science, City University of Hong Kong, Hong Kong and E-mail: kltsui@cityu.edu.hk

Abstract

The development of health indicators for Prognostics and Health Management (PHM) applications require in-depth domain knowledge to extract useful features under different operating environments. This paper presents a brisk and simplified approach for unsupervised feature extraction in PHM applications. The integrated approach utilizes a combination of Extreme Learning Machine (ELM) to improve training speed, Sparse Auto-Encoder (SAE) to learn sparse features, and back propagation (BP) algorithm to reduce residual error and improve the learned features. Preliminary results of the proposed single-layer unsupervised BP-SAE_ELM approach when tested with IEEE PHM 2012 Data Challenge and FEMTO bearing datasets yielded not only comparative result but improved extracted features when compared to other supervised traditional approaches, such as root mean square, kurtosis, and peak to peak. The proposed method is found to be effective in extracting features from huge amounts of data, and might be ineffective when applied to very small dataset. The proposed method is computationally effective and can be embedded in on-board sensors of highly valued assets in streaming data analysis of big data systems applications, such as in rail transportation systems, to compress the huge data generated, and ensure that only data representing the health index are stored for remaining useful life (RUL) estimation.

Keywords: *PHM, ELM, SAE, ARENA, Maintenance Policy*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-194

ON A ROBUST STATISTIC FOR SCALE AND ITS APPLICATION IN ENGINEERING

Hideki Nagatsuka[1]

[1] Department of Industrial and Systems Engineering, Chuo University and E-mail: hideki@indsys.chuo-u.ac.jp

Abstract

In this talk, I will firstly introduce some basic robust statistics and their properties.
After them, I will discuss the derivation of the exact finite-sample distribution of a robust statistic for scale under the very mild conditions. Then, I will apply the above results to provide an exact confidence intervals for the location-scale family.
Finally I will show the results of Monte Carlo studies to evaluate the methods developed and give an example in engineering area.

Keywords: *robust statistics, exact finite-sample distribution, confidence intervals*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-195

PRIVACY PRESERVING AND SECURE LOGISTIC REGRESSION FOR HORIZONTALLY DISTRIBUTED DATA

Junghye Lee[1]

[1] Ulsan National Institute of Science and Technology and E-mail: junghyelee@unist.ac.kr

Abstract

In the data market, which is growing at a tremendous pace, the subject of the most talked about lately is the sharing of data and the exchange of information. Data sharing and information exchange play an important role in data analysis because it can improve the quality of data analysis, accelerate meaningful information creation, and facilitate the meaningful secondary use of data. However, protecting privacy is a major concern. Advanced security methods have been developed to protect the storage and computation of sensitive data in distributed settings. However, they do not protect against information leakage due to data analysis results. To address this aspect, research on differential privacy (a state-of-the-art privacy framework) has been encouraged, but most of them do not apply to distributed scenarios. Combining security and privacy methodologies is a natural way to solve problems, but naïve solutions can degrade analytical performance. This study introduces a new strategy that combines different privacy methods with homomorphic encryption in a harmonious manner. Using logistic regression (a universal model in biomedicine), we use real world data sets (using AUC) to demonstrate the practicality of building secure and privacy-preserving models with high efficiency and good accuracy.

Keywords: *Privacy, Security, Logistic regression, Horizontally distributed data*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-196

MACHINE LEARNING ALGORITHM FOR SUBSTATION OF CHEMICAL REACTION MODEL

Jong-Ho Shin[1], Sun-Joong Kim[2]

[1] Dept. of Industrial Engineering, Chosun University and E-mail: jhshin@chosun.ac.kr
[2] Dept. of Materials Engineering, Chosun University and E-mail: ksjoong@chosun.ac.kr

Abstract

Due to the characteristics of Manganese which improves mechanical properties of steel, it is used as one of the key alloying elements in making advanced steel. As the demand of high quality steel is expanding recently, the consumption of Manganese is also growing concurrently. To meet the increasing demand of Manganese, more Manganese ores should be excavated from mines. However, the location of Manganese mines are restricted to few countries such as South Africa and Australia. Most of countries should rely on the import of Manganese in the production of high quality steel. Another possible source of Manganese is to extract Manganese from steelmaking slag. To recycle Manganese in steelmaking slag, some authors proposed an innovated process to extract manganese from steelmaking slag. The developed process is formulated as a reaction model and used to simulate the extraction procedure. In many ways, the advantage of reaction model in the field is no doubt. However, the development of the reaction model requires highly expertized knowledge and it is not always possible to build it. Due to these difficulties, most of companies rely on operation experts and he/she controls reaction process based on his/her experience. To cover these limitations and help field operators, it is necessary to provide an empirical method to substitute a reaction model. To do this, this study proposes a utilization of machine learning algorithm so as to substitute the reaction model with field data. The main objective of this study is to show the possibility to substitute the reaction model with ANN, which means the operators can control smelting process by machine learning based on the field data without special knowledge on reaction model.

Keywords: *Artificial neural network, smelting process, machine learning, chemical reaction model, artificial intelligence*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-197

BAYESIAN UPDATING METHODS FOR LIFE PREDICTION USING FIELD FAILURES DATA

Joo Ho Choi[1], Dawn An[2], Nam Ho Kim[3]

[1] Korea Aerospace University and E-mail: jhchoi@kau.ac.kr
[2] Korea Institute of Industrial Technology and E-mail: dawnan@kitech.re.kr
[3] University of Florida and E-mail: nkim@ufl.edu

Abstract

In the fatigue life design of mechanical components, uncertainties arising from materials and manufacturing processes are usually considered by applying a safety factor to the resulting design, which most likely relies on the designer's experience. Due to the ignorance of the uncertainties, however, the value tends to be high, leading to unnecessarily conservative design. In this paper, Bayesian technique, which incorporates the field failure data into the prior knowledge, is used to obtain a more dependable prediction of fatigue life. The highlight is to estimate the parameters of life distribution conditional on the field failures data, which are by themselves given by the distributions, hence, often given by confidence interval. As more data are provided, the values are updated to reduce the confidence interval. The results can be used in various needs such as a risk analysis, reliability based design optimization, maintenance scheduling, or validation of reliability analysis codes. Markov Chain Monte Carlo technique is employed to find out the posterior distribution of the interested parameters, which effectively draws the samples of the given distribution. Failure data gathered for the turbine blades made by the regular inspection and suspension springs of automotive made by durability tests are exploited to illustrate our approach.

Keywords: *Bayesian statistics, Field data, Life prediction*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-198

SOH-FLUCTUATION ANALYSIS FOR EARLY DETECTION OF UNHEALTHY LI-ION BATTERIES: A RANDOM FOREST APPROACH

Changyong Lee[1]

[1] Ulsan National Institute of Science and Technology and E-mail: changyong@unist.ac.kr

Abstract

This study proposes a state-of-health (SOH)-fluctuation analysis for early detection of unhealthy Li-ion batteries. For this, first, three types of quantitative indicators are defined to characterise the fluctuation patterns of SOH of Li-ion batteries during qualification tests. Second, random forests are employed to assess the remaining useful life (RUL) of batteries. Finally, the performance of our approach was assessed by using quantitative metrics. In addition, this study also investigates the importance of quantitative indicators for early detection of unhealthy Li-ion batteries. Experimental results on Li-ion batteries for consumer electronics confirm that the proposed approach can reduce the required number of cycles for qualification tests to 50 cycles, less than a month in practice.

Keywords: *qualification test, unhealthy Li-ion battery, early detection, SOH-fluctuation*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-200

FATIGUE LIFE SIMULATION OF PTH SOLDER JOINTS CONSIDERING ALEATORY UNCERTAINTY

Dongmin Lee[1], Hyunseok Oh[2]

[1] Gwangju Institute of Science and Technology and E-mail: dmlee4045@gist.ac.kr
[2] Gwangju Institute of Science and Technology and E-mail: hsoh@gist.ac.kr

Abstract

This study presents a pin-through-hole (PTH) simulation model and probabilistic life prediction of solder joints subjected to thermal cycling loads. The simulation model includes solder joints, the film capacitor and printed circuit boards. The Anand model is used to simulate the viscoplastic behavior of SAC 305 solders. Thermal cycling between -40°C and 125°C is used. Then, the strain energy density is calculated through the finite element analysis. The PTH solder joint fatigue lifetime is calculated using the Darveaux model. Finally, the statistical distribution of the solder joint fatigue life is estimated with the PTH simulation model and advanced uncertainty propagation methods considering aleatory uncertainty of the solder joints. The probabilistic approach to the PTH solder joints is anticipated to accurately estimate the reliability of the PTH solder joints during product design.

Keywords: *Inherent randomness, solder joints, Anand model, fatigue life, uncertainty propagation, SAC305*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-201

ESTIMATION OF HEALTHCARE ACCESSIBILITY USING AN ONLINE EXPERIMENT

Youngdeok Hwang[1]

[1] Sungkyunkwan University and E-mail: yhwang@g.skku.edu

Abstract

Automated data collection using various sources becomes an important part of statistical practice. In this talk, we present a statistical approach to study the accessibility and coverage of the urgent health care that determines the patient outcomes. I will discuss an online experiment to estimate the nation-wide statistics for first medical contact to door time using a large scale automated map queries, combined with national patient statistics. The proposed approach utilizes the design of experiment to plan the data collection process, while taking into account the survey sampling perspective.

Keywords: *healthcare accessibility*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-202

RESILIENCE ASSESSMENT OF POWER SYSTEMS USING TREND TESTS

Shen Lijuan[1], Tang Loon Ching[2]

[1] Department of Industrial Systems Engineering and Management, National University of Singapore and E-mail: isesl@nus.edu.sg
[2] Department of Industrial Systems Engineering and Management, National University of Singapore and E-mail: isetlc@nus.edu.sg

Abstract

Resilience is the ability of a system to reduce the chances of shocks, absorb the shock if it occurs and recover quickly if a shock causes disruption. Power systems become increasingly complex due to the growth in both the scale and interdependencies. With the increased population density and demand for electricity, major disruptions become more damaging. This study conducts an empirical evaluation on the resilience of the U.S. power grid based on the database of the Electric Disturbance Events. To assess trends in systems resilience, we look into three key components associated with each black-out and recovery of power systems, i.e., the time between disruptions, the performance loss of each disruption and the time needed for recovery. We present a combined measure that takes into account all the three components. A modified Lewis-Robinson test is then developed for trend detection in this combined measure. To support the trend analysis of this combined measure, we further perform trend test for the performance loss and the recovery time. It is found that, among various North American Electric Reliability Corporation (NERC) regions of the U.S. power grid, the resilience in the Northeast Power Coordinating Council (NPCC) region has become better. Empirical evidence from the government financial support is used to substantiate these statistical findings.

Keywords: *Trend test, Frequency of occurrence, Performance loss, Recovery time*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-203

A CONSTRUCTION OF COST-EFFICIENT DESIGNS WITH GUARANTEED REPEATED MEASUREMENTS ON INTERACTION EFFECTS

Yasmeen Akhtar[1], Frederick Kin Hing Phoa[2]

[1] Arizona State University and E-mail: ysa.yasmeen@gmail.com
[2] Academia Sinica and E-mail: fredphoa@stat.sinica.edu.tw

Abstract

This work introduced a useful class of cost-efficient designs for two-level multi-factor experiments. It provided guaranteed repeated measurements on all 2-tuples from any two factors and the number of repetitions was adjusted by the experimenters. Given the number of factors of interest, it utilized less resources than an orthogonal array while its repeated measurement provided a resistance towards outliers that a covering array failed to achieve. To bridge the wide spectrum between two extreme settings (orthogonal arrays and covering arrays) in terms of the number of repeated measures of tuples, we developed a systematic method to construct families of these designs, namely (supersaturated) repeated coverage design, with small run sizes under different number of factors and number of repetitions.

Keywords: *Orthogonal Arrays, Covering Arrays, Repeated Coverage*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-205

MATRIX COMPLETION FOR NETWORK ANALYSIS

Ji Zhu[1]

[1] Department of Statistics, University of Michigan and E-mail: jizhu@umich.edu

Abstract

Matrix completion is an active area of research in itself, and a natural tool to apply to network data, since many real networks are observed incompletely and/or with noise. However, developing matrix completion algorithms for networks requires taking into account the network structure. This talk will discuss two examples of matrix completion used for network tasks. First, we discuss the use of matrix completion for cross-validation or non-parametric bootstrap on network data, a long-standing problem in network analysis. The second example focuses on reconstructing incompletely observed networks, with structured missingness resulting from the egocentric sampling mechanism, where a set of nodes is selected first and then their connections to the entire network are observed. We show that matrix completion can generally be very helpful in solving network problems, as long as the network structure is taken into account. This talk is based on joint work with Elizaveta Levina, Tianxi Li and Yun-Jhong Wu.

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-206

AN ITERATIVE APPROACH FOR FIXTURE ALLOCATION AND SCHEDULING FOR A FLEXIBLE MACHINING SYSTEM WITH MULTI-FIXTURING PALLETS AND CONTROLLABLE PROCESSING TIMES

Jeong-Hoon Shin[1], Dong-Ho Lee[2]

[1] Department of Industrial Engineering, Hanyang University and E-mail: boost7@hanyang.ac.kr
[2] Department of Industrial Engineering, Hanyang University and E-mail: leman@hanyang.ac.kr

Abstract

This study addresses an integrated fixture allocation and scheduling problem for a flexible manufacturing system with multi-fixturing pallets that can load multiple parts of different types and controllable processing times that can change processing times with different processing costs. The fixture allocation sub-problem is to allocate fixtures to each multi-fixturing pallet and the resulting scheduling sub-problem is to determine the set of parts loaded on each pallet according to the fixture allocation, the sequence of pallets to be released initially through the loading/unloading station, the allocation of pallets to machines, the part processing times and the sequence of pallets allocated to each machine for their loaded parts to be processed. Practical constraints such as central buffer capacity, limited number of pallets, maintenance times and setup worker availability are also considered. To investigate the trade-off between scheduling and cost performance, a bi-criterion objective is considered that minimize the total tardiness and the total processing cost. After the problem is explained in more details, an overview of the solution approach is proposed that consists of two iterative phases: search heuristic for fixture allocation and priority rule based heuristic for scheduling.

Keywords: *Flexible manufacturing system, multi-fixturing pallets, controllable processing times, fixture allocation, scheduling, iterative solution approach*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-207

POWER LAW PROCESS WITH LEFT-TRUNCATED AND RIGHT-TRUNCATED DATA FOR MULTIPLE REPAIRABLE SYSTEMS

Byeong Min Mun[1] , Seong-Joon Kim [2], Suk Joo Bae[1,*]

[1] Department of Industrial Engineering, Hanyang University
[2] Department of Industrial Engineering, Chosun University
* Corresponding author: sjbae@hanyang.ac.kr

Abstract

A nonhomogeneous Poisson process (NHPP) has been widely used in modelling recurrent failure data of multiple repairable systems. Sometimes, the failure data of multiple repairable systems in the initial recording period may be missed, thus there is no information about how many failures have occurred before the recording time. Dealing with this kind of left truncated data is challenging because there exists various types of incompleteness in the aspect of missing positions in the early time epoch and the number of missing failures. Furthermore, we may encounter some circumstances that we should cease data collection at a predetermined time (right time-truncated) or at a predetermined number of failures (right failure-truncated).
In this paper, we propose the power law process (PLP) for recurrent failure data from left-truncated and right-truncated data for multiple repairable systems. We also propose frailty models to describe the system-to-system variability due to changes in operating conditions and working conditions for multiple repairable systems. A maximum likelihood estimation (MLE) is used to estimate PLP parameters and construct confidence intervals for the cumulative number of failures. The proposed models are applied to the failure data of multiple repairable systems.

Keywords: *frailty model, left truncated data, minimal repair, nonhomogeneous Poisson process, power law process*

The Fifth International Conference on the Interface between Statistics and Engineering 2019 (ICISE2019)

ICISE-208

SEQUENTIAL DESIGN OF EXPERIMENTS FOR MULTIRESPONSE SURFACE OPTIMIZATION

Donghee Lee[1] , Sohee Kim[2]

[1] Hanyang University and E-mail: dh@hanyang.ac.kr
[2] Hanyang University and E-mail: sohee.oreo@gmail.com

Abstract

Multiresponse problems are common in process optimization. Conventional approach for optimizing multiresponse is using response surface methodology and this approach is called as multiresponse surface optimization (MRSO). RSM is a sequential procedure in that it first searches region of optimum. Once the region of optimum has been found, elaborated models are employed and an analysis is performed to find the optimum. The first part in RSM (i.e., searching the region of optimum) has been rarely studied for MRSO while most of MRSO focuses on the later part (i.e., finding the optimum). This paper suggests a sequential design of experiments for searching the region of optimum in MRSO. In the proposed method, several candidates of directions for sequential experiments are generated by considering tradeoffs between the multiple responses. The process engineering selects one of direction for the sequential experiments based on his/her preference on the multiple responses. In this paper, a systematic procedure for generating and selecting the direction of the sequential experiments is suggested and a hypothetical example is employed for illustrating the proposed procedure.

Keywords: *Design of experiments, multiresponse optimization, sequential design*