UNIVERSITÀ DEGLI STUDI FIRENZE

Inter-University Research Center
StEering: Design Quality and Reliability

UNIVERSITÀ DEGLI STUDI FIRENZE
DISIA
DIPARTIMENTO DI STATISTICA, INFORMATICA, APPLICAZIONI "G. PARENTI"

WORKSHOP 2020

STATISTICS AND INNOVATION FOR INDUSTRY 4.0

February, 20-21, 2020

enbis European Network for Business and Industrial Statistics

RENMIN UNIVERSITY OF CHINA 1937 中国人民大学

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

Welcome to the StEering Workshop

Statistics and Innovations for Industry 4.0

Dear Participant,

On behalf of the Scientific Committee, it is a great pleasure for me to welcome you to the StEering Workshop at the Department of Statistics Computer Science Applications "G.Parenti"- DISIA- University of Florence (Italy).

We are proud to offer you a Workshop programme which confirms the effective relationship between academic research and practical applications in business and industrial statistics, also expounding the usual topics towards data science, in a 4.0 perspective. With a wide range of Invited Talks related to the main topics of Statistics and Engineering for Industry 4.0, we hope you find the programme stimulating and interesting.

The Workshop is organized by the Inter-University Research Center StEering (http://www.steering.academy/WS2020/index.html) and DISIA, with the Scientific Support of ENBIS (www.enbis.org), Bejing University of Technology, and Renmin University of China.

International Speakers will share their knowledge and innovation studies in order to improve researches related to the field of statistics, engineering, management and technology *in and for* a global and sustainable world. Seven Invited Sessions are organized considering: *Statistical for Engineering, Reliability, Big Data and Classification, Bayesian Methods, Computer Experiments, Business and Marketing Studies.*

Talks from industry (on successful applications) related to the illustration of particularly challenging technological issues will give the opportunity to discuss and suggest real solution for innovation in an Industry 4.0 perspective.

Besides the Invited, Contributed researches will be presented during the Poster Sessions scheduled during the two lunches.

The official journal of ISBIS, the International Society for Business and Industrial Statistics, *Applied Stochastic Models in Business and Industry* (ASMBI), http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1526-4025 is inviting contributions for a special issue "*Statistics, Statistical Engineering, and Innovation for Industry 4.0*" (Guest Editors: Rossella Berni, Jesus Fernando Lopez-Fidalgo, and Geoff G. Vining) devoted to advances in Statistics and Statistical Engineering towards Industry 4.0, following the Workshop, deadline June 30th, 2020.

Unfortunately, all the Invited Speakers, and Contributed, coming from China and

Hong Kong Region cannot be present (onsite) at the Workshop, nevertheless the Organizing Committee made every attempts for solving this sudden and undesirable issue through the videos and streaming connections.

I would like to thank the Scientific and Organizing Committee members, especially the Academic Chinese Collegues: Prof. Yanyun Zhao, Prof. Yubin Tian, and Prof. Quian Guo, that spent all the efforts for collaborating with me and Prof. Guido Ferrari for a successful event, the Collegues of the StEring Research Center, particularly Prof. Nicola Bonora, and all the attendants (onsite and by streaming), as well as all of you who will contribute to its success by presenting talks and posters, and taking an active part in discussions.

Enjoy the Workshop and have a good time in Florence,

Rossella Berni
Chair of the Scientific Committee

Inter-University Research Center
StEering: Design Quality and Reliability

The Inter-University Reseach Center StEering has been founded in 2017 http://www.steering.academy/WS2020/index.html.
Currently it has own administrative headquarters at the Department of Statistics Computer Science Applications 'G.Parenti' - DiSIA, University of Florence. https://www.disia.unifi.it/p186.html.

SteEring, (acronym of "Interuniversity Center for Statistics for Engineering: Design, Quality and Reliability") is an Interuniversity Research Center which promotes the scientific collaboration among Universities and Research centers or inter-university service centers.
Its mission is to promote and spread the knowledge and use of statistical tools and methods in the engineering field with particular reference to the business world and to support companies.

SteEring is composed by Italian Universities, through its own Departments, including Professors, Researchers and Research Laboratories.

StEering University (and Department) Members:

Università degli Studi di Firenze (https://www.unifi.it/vp-9240-dipartimenti.html)
Dipartimento di Statistica, Informatica, Applicazioni - "G. Parenti" (DiSIA)
Dipartimento di Ingegneria dell'Informazione (DINFO)
Dipartimento di Ingegneria Industriale (DIEF)

Università degli Studi "G. Marconi" (https://www.unimarconi.it/en/presentazione-6)
Dipartimento di Ingegneria dell'Innovazione e dell'Informazione (DIIE)

Università degli Studi di Cassino e del Lazio Meridionale
Dipartimento di Ingegneria Civile e Meccanica (DICeM)
(https://www.unicas.it/siti/dipartimenti/dicem.aspx)

SCIENTIFIC COMMITTEE

- **Rossella Berni – Chair**
  (Prof. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Gabriele Arcidiacono
  (Prof. -Università G. Marconi, Roma, Italy)
- Stefano Bacherini
  (Ing. – Alstom Group SpA)
- Nicola Bonora
  (Prof. - Università di Cassino e del Lazio Meridionale, Italy)
- Luciano Cantone
  (Prof. - Università degli Studi di Roma "Tor Vergata")
- Marcantonio Catelani
  (Prof. - Dept. of Information Engineering, University of Florence, Italy)
- Paolo Citti
  (Prof. -Università G. Marconi, Roma, Italy)
- Jesus-Fernando Lopez-Fidalgo
  (Prof. -University of Navarra, Spain)
- Marco Pierini
  (Prof. - Dept. of Industrial Engineering, University of Florence, Italy)
- Antonio Pietavolo
  (Dr. – CNR IMATI Milano, Italy)
- Silvia Ranfagni
  (Prof. - Department of Economics and Management, University of Florence, Italy)
- Andrew Ruggiero
  (Prof. - Università di Cassino e del Lazio Meridionale, Italy)
- Federico M. Stefanini
  (Prof. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Grazia Vicario
  (Prof. - Dipartimento di Scienze Matematiche "G. L. Lagrange"- Politecnico di Torino)
- G. Geoffrey Vining
  (Prof. – Virginia Tech, Virginia, USA)
- Yanyun Zhao
  (Prof. - School of Statistics, Renmin University of China; Secretary - General of the Education and Instruction Commission of Applied Statistics Master Degree of Ministry of Education of China)

ORGANIZING COMMITTEE

- **Rossella Berni – Chair**
  (Prof. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Nicola Bonora
  (Prof. -Università di Cassino e del Lazio Meridionale, Italy)
- Guido Ferrari
  (Prof. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Alessandro Magrini
  (Dr. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Nedka D. Nikiforova
  (Dr. - Dept. of Statistics Computer Science Applications 'G.Parenti'- Univ. Firenze)
- Silvia Ranfagni
  (Prof. - Department of Economics and Management, University of Florence, Italy)
- Andrew Ruggiero
  (Prof. - Università di Cassino e del Lazio Meridionale, Italy)
- Yubin Tian
  (Prof. - Dean of Science, Beijing University of Technology, China)
- Yanyun Zhao
  (Prof. - School of Statistics, Renmin University of China; Secretary-General of the Education and Instruction Commission of Applied Statistics Master Degree of Ministry of Education of China)

Detailed Programme Overview

Thursday, February 20th, 2020

9.00-9.30 Opening Ceremony

9.30-11.00 Statistical for Engineering

Chair and Discussant: Rossella Berni

Geoff G. Vining (*Virginia Tech, Virginia, USA*)- *Towards a Foundational Theory of Statistical Engineering*

Francesco Bertocci (*Esaote SpA*)- *Statistics for Continuous Improvement of Manufacturing Process of Ultrasound Probes for Medical Imaging*

Umberto Galietti (*Politecnico di Bari, Italy*)- *On Some Issues of DoE and Robust Design in Experimental Mechanics Supported by Numerical Simulations*

11.00-11.30 COFFEE BREAK

11.30-13.00 Reliability- Thursday session

Chair and Discussant: Marcantonio Catelani

Loredana Cristaldi (*Politecnico di Milano, Italy*)**-** *Copula Method In Power System*

Qingpei Hu *(Chinese Academy of Sciences)- Assessment of Reliability in Accelerated Degradation Testing with Initial Status Incorporated*

Nikolaus Haselgruber *(CIS- Consulting in Industrial Statistics GmbH)- Reliability Verification as Integrated part of a Product Development Process*

13.00-14.30 Lunch and 1st Poster Session

14.30-16.00 Big Data and Classification

Chair and Discussant: Grazia Vicario

Bart De Ketelaere *(KU Leuven- Belgium)- Big Data Classification in Presence of Label and Measurement noise*

Jesus Fernando López Fidalgo (*University of Navarra – Spain*)- *Robust Learning for classification*

Fugee Tsung *(Hong Kong University of Science & Technology -HKUST)- Quality Big Data*

16.00-16.30 COFFEE BREAK

16.30-18.30 Computer experiments

Chair and Discussant: Geoff G. Vining

Grazia Vicario (*Dipartimento di Scienze Matematiche "Lagrange"- Politecnico di Torino, Italy*) - *Inference On Errors in Industrial Parts: Kriging and Variograms Versus Geometrical Product Specifications (GPS) Standard*

Matthias HY Tan *(School of Data Science, City University of Hong Kong)-Gaussian Process Modeling with Boundary Information*

Alessandro Magrini, Nedka D. Nikiforova (*Dept. of Statistics Computer Science Applications 'G.Parenti'- Università di Firenze, Italy*) - *Optimization Methods And Simulations For Improving The Payload Distribution of Freight Trains*

Shifeng Xiong, *(Academy of Mathematics and Systems Science, Chinese Academy of Sciences)- Reconstruction Parameterization: From Interpolation to Regression, Classification and Numerical Computing*

20.30 Social dinner

Friday February 21st, 2020

9.00-10.30 Business and Marketing studies

Chair and Discussant: Silvia Ranfagni

Yanyun Zhao, **Kang Wang** (*School of Statistics, Renmin University of China*)- *Beijing Enterprise Innovation and Development of Sophisticated Industry*

Wilson Ozuem (*University of Cumbria, UK*)- *Differential Effects Of Online Brand Communities And Customer Loyalty In The Fashion Industry: An Empirical Investigation of Millennials*

Nicola Antonelli (*LUISAVIAROMA, Italy*) *Il Digital Luxury Retailer in International Market: the Case of Luisaviaroma*

10.30-11.00 COFFEE BREAK

11.00-12.00 Bayesian Session

Chair and Discussant: Jesus Fernando López Fidalgo

Yubin Tian, Dianpeng Wang *(School of Mathematics and Statistics, Beijing Institute of Technology, China)- Adaptive Bayesian Prediction of Reliability Based on Degradation Process*

Federico Mattia Stefanini (*Dept. of Statistics Computer Science Applications 'G.Parenti'- Università di Firenze , Italy*)- *Bayesian Analysis Of Mediation In Cell Transformation Assays For Testing The Carcinogenicity Of Chemicals*

12.00-13.30 Reliability – Friday session

Chair and Discussant: Gabriele Arcidiacono

Giovanni Belingardi (*Dipartimento di Ingegneria Meccanica e Aerospaziale, Politecnico di Torino, Italy*) - *Methodologies for The Health State Monitoring of Automotive Composite Parts*

Iacopo Trotta (*Alstom Group*)- *Dependability Data Uncertainty Influence on Safety Analyses for Railway Signaling*

Antonio Pievatolo (*CNR IMATI Milano, Italy*) – *Case Studies in Failure Data Analysis*

13.30-15.00 Lunch and 2nd Poster Session

15.00-15.30 Greetings and Closing Ceremony

Poster programme

13.00-14.30 Lunch and 1st Poster Session (Chair: Rossella Berni)

Chiara Galimberti
Identifying and Representing Clusters of Spatial Defects in Microelectronics Planar Artefacts

I. García-Camacha Gutiérrez
Development of Model-Robust Designs for Accelerated Failure Time Models with Right-Censored Data

Jose Moler
Heuristic Algorithm to Reduce the Number of Measurements in a Quality Control

Gianni Pistone
Empirical Variograms

Yonggwan Shin
An Iterative Method for Tuning Complex Simulation Code

13.30-15.00 Lunch and 2nd Poster Session (Chair: Federico M. Stefanini)

À. Sebastià Bargues
"Computing exact andapproximate D-optimal designs for electrical impedance measurements"

Rosamarie Frieri
"Design of experiment for multi-stage processes"

M. Blanca Palacios
"Restricted experimental designs to optimize quality standards in an electro-color process"

Luca Pegoraro
"A Combined Approach to Detect Key Variables in Thick Data Analytics"

Invited talks

StEering
Inter-University Research Center:
Design, Quality and Reliability

Copula method in power system

Michy Alice[1]
Loredana Cristaldi[1]
Enrico Ragaini[2]

[1)] Politecnico di Milano.
[2)] ABB.

e-mail address: loredana.cristaldi@polimi.it

Keywords: Industrial circuit breakers; Remaining useful life; Statistical test

ABSTRACT –Two data-driven prognostic algorithms based on copula application to power systems are proposed. The first is related to the estimation of the Remaining Useful Life (RUL) of a product, and the second aims at evaluating the performance and predict the behavior of an energy-consuming load. Obtained results are encouraging and candidate this approach as a useful method in Prognostics and Health Management (PHM) and energy monitoring applications.

1. INTRODUCTION

Prognostics and Health Management (PHM) and energy efficiency programs are assuming a central role in the transition of power systems towards the Industry 4.0 paradigm. PHM targets are reduction of maintenance and life-cycle management costs, increase of the systems availability and adoption of Predictive Maintenance (PdM) strategies. Energy efficiency programs take into account all the different phases in which energy is involved. Cost reduction and widespread availability of condition monitoring (CM) apparatus allow to develop and apply data-driven to predict both the health condition (HC) and the Remaining Useful Life (RUL) of a product and quantify the use of the energies by means of specific indexes. Predicting the future behavior of a device (in terms of HC and RUL) by learning from its history and from the past behavior of similar products is an essential objective of a PHM program [1],[2][3].

2. COPULAS IN PHM AND ENERGY EFFICIENCY

Copula is a function which joins (or couples) multivariate distribution functions to their one-dimensional marginal distribution functions. It allows extracting information about the correlation structure between random variables and, in general, can capture nonlinear relationships as well. These properties have been investigated in order to acquire new instruments for PHM and energy efficiency programs.

2.1 PHM case study

The Remaining Useful Life (RUL) of a system is the remaining time interval in which it will be able to meet its operating requirements. Health condition is defined as "the extent of degradation or deviation from an expected normal behavior" [2]. In order to set prognostic algorithms, three different approaches can be followed: model based, data driven and hybrid. In particular, data driven methods are mainly based on the exploitation of the collected run-to-failure data and do not usually require specific knowledge about the inherent failure mechanisms (strictly related to physical models describing the behavior of the systems under study)..

In the case of circuit breakers (CB), health condition refers to the degradation of the monitored components, such as wear of the switching contacts, leakages in the interrupting chamber, etc. Considering the complexity of a CB and the availability of inexpensive monitoring systems that allow the collection of condition monitoring data, a data driven approach allows reaching reasonable results [2], [3].

Considering the different contribution reported in literature, the key point is evaluating whether there are dependencies between random variables related to the functioning of the item. For a CB, two random variables are generally considered for the analysis of the degradation: the sampling time and the HC. If we consider the wearing of switching contacts, two mechanisms may play a role: a gentle degradation due to the aging and a strong degradation due to external events related to operations (e.g., when a large current is interrupted). Of course, both mechanisms affect the apparatus, so this consideration suggests exploiting statistical methods like copulas to take into account possible correlation between the sampling time and the health condition [2, 5].

To this purpose, a set of time series (extracted using the two-sample Kolmogorov-Smirnov Test - KST) of health condition and sampling time of different CB installed in different locations was considered. For each breaker, sampling time and the health condition measured at that time are provided until the breaker needs to be replaced. The condition monitoring data on the variation of the sampling time for the contact wear, that corresponds to each single breaking (or switching) operation, provides the knowledge on the usage profile of the breaker. In this scenario, the variation of the health condition represents the degradation profile of the product (i.e. the wearing of the contacts associated to each switching operation). It is possible to introduce a degradation rate (d_i) given by (Figure 1):

$$d_i = \frac{\Delta HC_i}{\Delta t_i}$$

where:

$\Delta t_i = t_i - t_{i-1}$ and $\Delta HC_i = HC_i - HC_{i-1}$

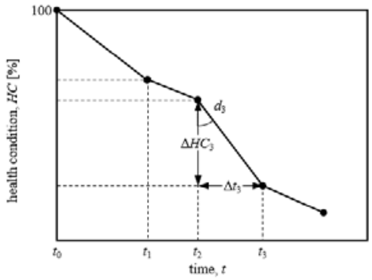

Figure 1 Degradation rate of a product

Independence test based on Kendall's τ has shown some possible dependence at 1% significance level.

The proposed algorithm gives results consistent with previous works [1, 4-6] and fairly accurate also for an observed degradation of 40%. Varying the selectivity α of the KST test for the subfleet seems to have little effect on the accuracy when using the copula (Figure 2).

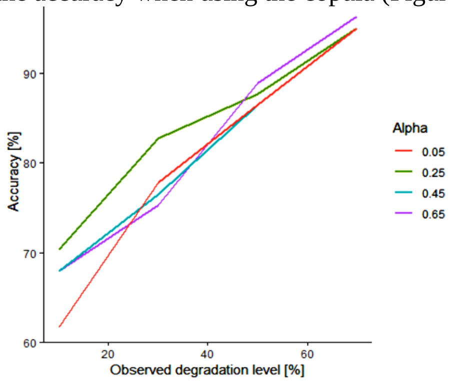

Figure 2 Algorithm performance (evaluated with 90 products)

2.2 Energy efficiency case study

Copula can be used in the modelling of multivariate time series as well [7]. For this case study, a dataset from a building management system, collecting the time series (with time interval of 30 minutes) related to the active power drawn by a set of aggregated loads, has been used. In Figure 3 it can be noted that there is a high positive correlation between one observation and the subsequent one, so conditional copula [6] can be used to simulate random load curves characterized by the fact that successive values are related by the dependence structure of the fitted copula.

3. SUMMARY

Two copula-based applications to power systems have been presented. The novelty of the first proposal is the attempt to exploit all the information enclosed in the product HC profile. The proposed method therefore represents a potential tool for an effective predictive maintenance strategy. In the second application, a prediction algorithm has been developed to predict the behavior of an energy-consuming load; based on the results, implementation of the algorithm in a monitoring apparatus able to model and to simulate the aggregate loads over time seems feasible.

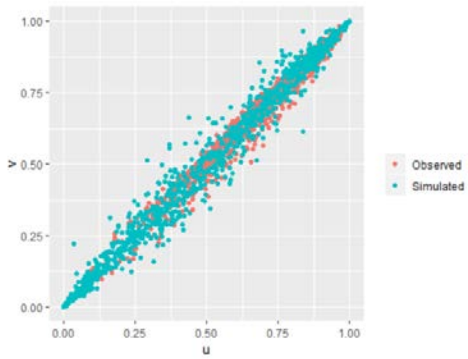

Figure 3 Correlation between subsequent observations in time series of active power drawn by a set of aggregated loads

4. REFERENCES

[1] G. Leone, L. Cristaldi, S. Turrin, "A data-driven prognostic approach based on statistical similarity: An application to industrial circuit breakers", *Measurement*, No. 108, pp. 163-170, 2017.

[2] M. G, Pecht, *Prognostics and Health Management*, Electronics, John Wiley & Sons, 2008

[3] M. Schwabacher, A survey of data-driven prognostics, in Proc. *AIAA Infotech@Aerosp. Conf*, Reston, VA, 2005.

[4] S. Turrin, S. Subbiah, G. Leone, L. Cristaldi, An algorithm for data-driven prognostics based on statistical analysis of condition monitoring data on a fleet level, in *2015 IEEE International Instr and Meas Tech. Conf. (I2MTC)*, 2015, pp. 629–634.

[5] M. Alice, D. Pejovski, L. Cristaldi, Remaining Useful Life Estimation of Industrial Circuit Breakers by Data-Driven Prognostic Algorithms Based on Statistical Similarity and Copula Correlation, in *16th IMEKO TC10 Conference "Testing, Diagnostics & Inspection as a comprehensive value chain for Quality & Safety"* Berlin, Germany, on September 3-4, 2019

[6] R. B. Nelsen, *An Introduction to Copulas*, Springer, 2006.

[7] B. Rémillard and C. Simard, "Forecasting time series with multivariate copulas," *Dependence Modeling*, vol. 3, no. 1, pp. 59 - 82, 2015.

Digitalization paths in a family-based luxury retailer: the case of Luisaviaroma

Nicola Antonelli[1)*]

[1)] LUISAVIAROMA SpA, Italy

[*]Corresponding author - e-mail address: n.antonelli@luisaviaroma.com

ABSTRACT – This chapter investigates the case of Luisaviaroma, an emblematiccompany that pioneered online business in the luxury fashion. Founded as a physical retailer in Florence, Italy, it exploited its points of sale to create an e-commerce platform to turn subsequently digital. In contrast to major online luxury retailers, it has remained a family business and developed business strategies based on a strong integration between the online and offline channels. What binds them strategically is its distinctive and unique ability to intercept iconic luxury brand products. Operationally, their combination relies on in-store bridge-technologies and services (free return, controlled logistics); to these are added digital marketing skills that blend with the culture and the strategic orientations of the company. Our intent is to investigate the factors that have made the digitalization of Luisaviaroma integrated with offline stores, retracing its evolutionary stages. Along all these stages, the family, together with the management, has always played a strategic role.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Dependability Data uncertainty influence on Safety Analyses for Railway Signaling

Stefano Bacherini*[1], Iacopo Trotta[1], Lorenzo Ciani[2], Marcantonio Catelani[2]

[1)] Alstom Ferroviaria S.p.A.
[2)] Università di Firenze, Scuola di Ingegneria.

*Corresponding author - e-mail address: stefano.bacherini@alstomgroup.com

Keywords: railway signaling; RAMS.

ABSTRACT – Railway Signaling is a highly standardized sector with strict requirements. Most of the reliability data on which safety target demonstration is based are affected by uncertainty, so alternatives to the classic deterministic methods (fuzzy methods and human reliability models) are explored.

1. INTRODUCTION

Railway signaling is a sector which is deeply driven by standards and Dependability, that is the combination of the RAMS disciplines (Reliability, Availability, Maintainability and Safety), is one of the most critical aspect of such systems. These standards are the CENELEC EN 50126, 50128 and 50129 series in Europe (and elsewhere) or the AREMA standards in North America. Such standards (CENELEC) specify targets for risk, defining a set of Safety Integrity Levels (SIL) associated to specific tolerable hazard rate (THR) ranges. For example, SIL4, the most restrictive level, requires the tolerable hazard rate to be lower than $1 \cdot 10^{-8}$ h^{-1}.

Demonstration that such targets are achieved mostly relays on the Fault Tree Analysis (FTA) [1] and the Failure Mode, Criticality and Effect (FMECA) [2] methods, which depend on parameters such as failure rates λ that, for their inherent nature, are affected by uncertainty. However, both the classical methods of FTA and FMECA do not consider such uncertainty in their results: alternative methods are thus considered to demonstrate that, even with a degree of uncertainty, Safety targets are achieved: fuzzy methods are introduced on top of the classic methods.

As human error is the one of the root cause of railway accidents today [3], human factor definition is also one of the elements considered**.**

2. METHODOLOGY

In signaling equipment failures are rare events (only a few per million hours) and collecting data on which to build a statistical "probability of failure" is costly and difficult. By allowing imprecision and approximate analysis, fuzzy logic can be included in reliability analyses, thus allowing uncertainty to be modeled (as in [4]). This is then used for FTA and FMECA.

Moreover, human reliability evaluation methods are applied, too.

2.1 Fuzzy FTA (FFTA)

Failure probabilities of the basic events of a fault tree can be represented as fuzzy numbers and these values can be used to determine the probability of the top event of fault tree in fuzzy form. FFTA is widely used, *e.g.* [5] and [6]. The process is as per Figure 1:

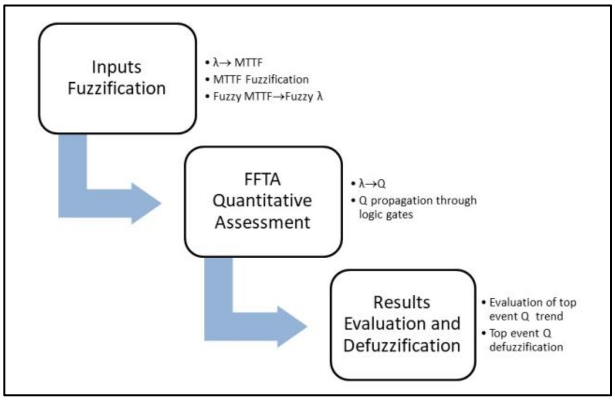

Figure 1 FFTA process

where MTTF is the Mean Time To Failure and Q is the Unavailability (*i.e.* the probability the system is faulty at time *t*).

A triangular fuzzification function is used, considering a *p* interval as the uncertainty for the MTTF value (see figure 2):

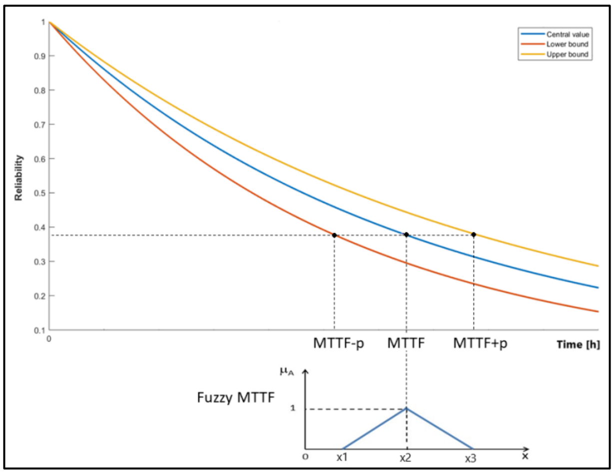

Figure 2 MTTF fuzzification

At the end of the process, it has been possible to compare the fuzzy, the de-fuzzy and classical failure rate curves with the target THR allowing to see that, with the chosen *p* interval, the target THR is still achieved even in case of uncertainty.

An alternative to this method is the use of Monte Carlo simulation for λ definition, which implies asymmetric MTTF triangles, but which provides similar results.

2.2 *Fuzzy FMECA (FFMECA)*

With respect to the classical FMECA, the approach has been to use triangular/trapezoidal fuzzification functions for the Occurrence (O), Severity (S) and Detection (D) parameters, which are defined by the analyst in the range from 1 (best) to 10 (worst) and whose product is the Risk Priority Number (RPN). This fuzzification copes with a possible wrong evaluation by the analyst. Considering an AND-OR fuzzy logic and after that an IF-THEN fuzzy implication rule, RPN is distributed as in figure 3.

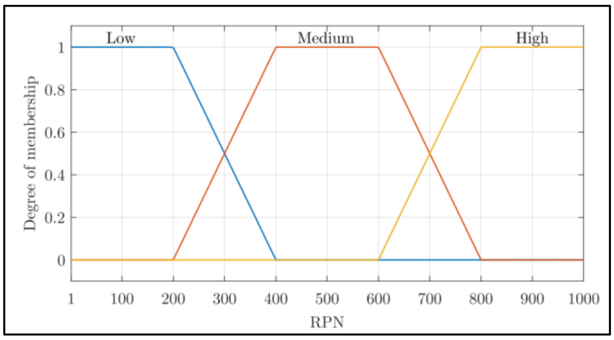

Figure 3 RPN fuzzification

RPN is then de-fuzzified and the resulting values are listed for each failure, thus allowing to understand what the most critical ones are. By comparison with the classical method, this gives some more granularity in the results.

2.3 *Human Factor*

The Railway Action Reliability Assessment (RARA, see [7]) method for the definition of Human Error Probability (HEP) has been used to compare different countermeasures in an otherwise classical Operating and Support Hazard Analysis (OSHA, see [8]). The introduction of this quantitative HEP assessment method has given the possibility to quantitatively evaluate the effects of the OSHA countermeasures instead of giving only a qualitative feedback, thus providing more support in the final decision of what cost-effective actions are more significant.

3. RESULTS AND DISCUSSION

The introduction of the proposed methods has allowed to cope with uncertainty, be this given by inherent physical properties of components (for failures), analyst misjudgments or impact in the implementation of a chosen countermeasure. This can be used both to confirm already existing results of classical analysis or to allow a finer judgment in borderline situations.

4. CONCLUSIONS

The concepts presented herein are an interesting innovation in the railway environment, as they are not contrasting with the prescriptions of the standards while they can indeed provide support in decision of the supplier (for design) and of the railway operators (for operation and maintenance), thus improving cost-effectively the Dependability of railway systems.

5. REFERENCES

[1] IEC 61025, *Fault Tree Analysis*, Technical report, International Electrotechnical Commission, 2007.

[2] IEC 60812, *Analysis techniques for system reliability - Procedure for failure mode and effects analysis (FMEA)*, Technical report, 2006.

[3] European Union Agency for Railways, *Railway Safety in the European Union – Safety overview 2017*, ERA, 2017.

[4] J. B. Bowles, C. E. Pelaez, "Application of fuzzy logic to reliability engineering", *Proceedings of the IEEE*, vol. 83, no. 3, pp. 435-449, 1995.

[5] H. Pan, W. Yun, "Fault tree analysis with fuzzy gates", *Comput. Ind. Eng.*, vol. 33, no. 3, pp. 569–572, 1997.

[6] L.P. Yang, "Analysis on dynamic fault tree based on fuzzy set", *Appl. Mech. Mater.*, vol. 110, pp. 2416–2420, 2011.

[7] RSSB, *Railway Action Reliability Assessment - A technique for the quantification of human error in the rail industry*, Rail Safety and Standards Board, 2012.

[8] Clifton A. Ericson, II, *Hazard Analysis Techniques for System Safety,* Wiley-Interscience, 2005.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Methodologies for the health state monitoring of automotive composite parts

Giovanni Belingardi[1,2], Alberto Ciampaglia[1], Nicola Amati[1,2], Maria Pia Cavatorta[1], Davide S. Paolino[1]

[1)] Politecnico di Torino – DIMEAS – Dept. of Mechanical and Aerospace Engineering.
[2)] Politecnico di Torino – CARS – Center for Automotive Research and Sustainable Mobility.

[*]Corresponding author - e-mail address: giovanni.belingardi@polito.it

Keywords: Health State Monitoring; Composite Materials

ABSTRACT – The paper proposes a methodology for the health state monitoring of composite parts, based on data gathering of the material mechanical behavior after fatigue or impact damage, and on a properly designed and trained neural network. Some preliminary results obtained on a simple plate are shown.

1. INTRODUCTION

The detection and evaluation of damage in components made of composite materials is one of the main concerns for automotive engineers. It is acknowledged that the damage of this type of components can originate during the manufacturing process but also as a result of impact events and/or fatigue loads [1,2].

Long fiber-reinforced plastic currently represents one of the most interesting composite materials for automotive components in the light weighting perspective that brings along reduction of fuel consumption and of Green House Gas emissions. Despite of a long-lasting experience in the aeronautical structures, automotive engineers are still looking for sound and cheap methodologies for the health assessment of composite material parts, not only during periodical inspections [3], but throughout vehicle operation.

Recently, the authors proposed an innovative simplified non-destructive methodology for the assessment of the damage sustained by composite materials. The methodology was specifically devoted to the detection and evaluation of the induced damage, in terms of residual elastic properties and, consequently, of residual strength.

The present paper proposes a step forward in the development of methodologies for the health assessment of composite material parts by means of continuous monitoring during vehicle operation.

2. HEALTH ASSESSMENT METHODOLOGY

It is well known that composite materials exhibit a fatigue behavior that is different from the typical behavior of metallic materials. In particular, in composites fatigue cycling induces a progressive degradation of the elastic modulus and strength. (see for example the diagram of figure 1, where the behavior of specimens submitted to constant amplitude fatigue cycling is shown). The degradation of the elastic modulus of the material, which is evident in the figure, can be used for detecting the position and the amount of damage that has occurred.

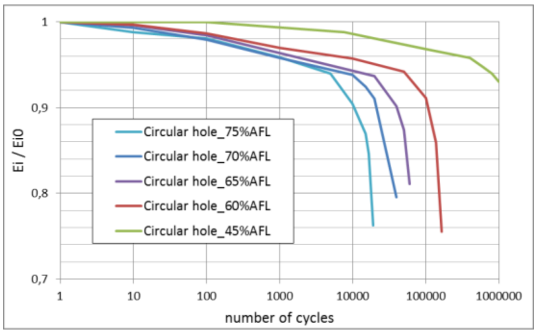

Figure 1 Progressive material degradation due to fatigue loading of notched specimens at different initial max load as a % of the Average static Failure Load (AFL).

There is a well-known relationship between the loss in elastic modulus and the loss in strength. If the strain field is measured in the considered part and the loading history can be reconstructed, an evaluation of the state of health of the part itself is then possible.

The measurement of the strain histories can be performed according to different procedures:

1 – by means of strain gages;
2 – by means of optical fibers;
3 – with the detecting damage index (DDI) technique;
4 – with digital image correlation (DIC) techniques.

Strain gages lead to a map of data based on a few, but properly conceived, locations. Through the use of a neural network, that has been instructed to analyze the relationship between the time history, it is possible to identify the locations on the part where the damage occurred and also the amount of damage. This technique can be effectively applied if one has the a-priori knowledge of the most convenient location for the application of the strain gages. The measurement chain is robust and relatively cheap, data processing can be made in real-time using an on-board computation unit.

The optical fibers lead to a more refined map of the strains, along the full length of the fiber itself. Recently developed techniques can provide the dynamic variation of thermal and mechanical strains. In this case, the measurement chain is robust although not cheap, the data processing can be made in real-time but it requires a dedicated computation unit.

The Detecting Damage Index procedure is basically a Lab test procedure. It leads to a refined map of the material damage. Quasi-non-destructive impact tests lead to the evaluation of the local value of the residual elastic modulus (see figure 2) and thus of the material damage.

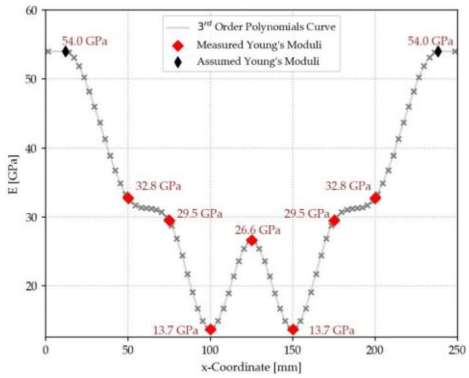

Figure 2 The residual elastic modulus calculated after impact test on the basis of the DDI technique. [4]

The Digital Image Correlation is basically a Lab test procedure, too. It leads to a full-size dynamic assessment of the strain field. Therefore, an accurate description of the state of health can be obtained with no need of an a-priori knowledge of the structure response. The measurement chain is robust although quite expensive and generally not suitable for in-field applications; the data processing can be made in real-time but it requires a dedicated large computation unit.

3. PRELIMINARY RESULTS

The proposed methodology has been assessed through some simulations. A simple composite laminate plate, fully constrained on one side and loaded at the opposite side, has been discretized and analyzed by FE, in two health conditions: without any damage (no initial damage, and no fatigue or impact damage) and with a localized damage.

The FE simulation provides a set of signals taken in selected locations that do not include the damage location. These signals are processed by means of a trained neural network, as it is schematically shown in figure 3. This process gives precise information on the location of the damage and on its entity. Obviously, the larger the number of signals and nodes of the neural network, the more precise the damage information will become.

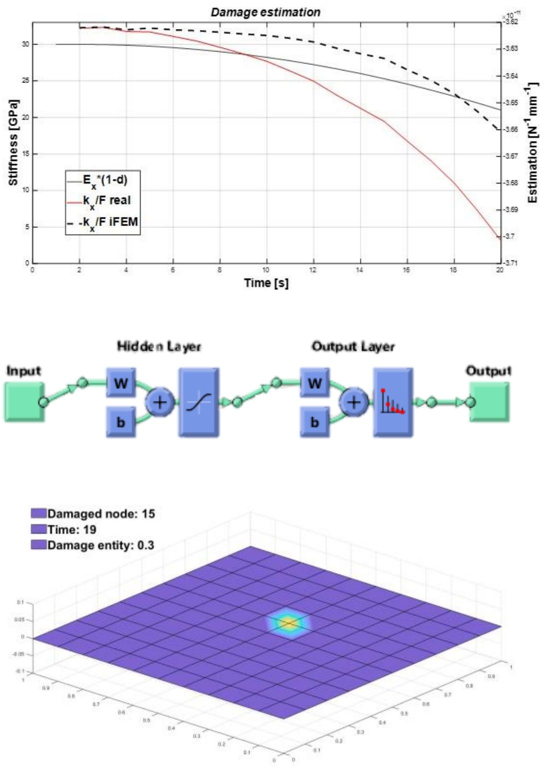

Figure 3 – a) Nodal data, b) neural network, c) identification of damage location and entity

4. CONCLUSIONS

A methodology for the health state monitoring of composite parts has been proposed, which is based on the knowledge of the material mechanical behavior , and on a properly designed and trained neural network,. Some encouraging preliminary results on a simple plate are shown.

5. REFERENCES

[1] A.T. Beyene, G. Belingardi, and E.G. Koricho, "Effect of notch on quasi-static and fatigue flexural performance of Twill E Glass/Epoxy composite", *Composite Structures*, vol. 153, pp. 825-842, 2016.

[2] D.S. Paolino, M.P. Cavatorta, and G. Belingardi, "Effect of thickness on the damage tolerance of glass/epoxy laminates subject to repeated impacts", *Proc. Inst.of Mech. Eng - Part C,* vol. 232, no. 8, pp. 1363-1373, 2018.

[3] G. Belingardi, M.P. Cavatorta, and D.S. Paolino – "Composite material components damaged by impact loading: a methodology for the assessment of their residual elastic properties", *J. of Achiev. in Mat. Manuf. Eng.*, vol. 87, no.1, pp 18-24, 2018.

[4] C. Boursier Niutta, A. Tridello, R. Ciardiello, G. Belingardi, and D.S. Paolino, "Assessment of residual elastic properties of a damaged composite plate with combined Damage Index and Finite Element methods", *Applied Sciences,* vol. 9, paper 2579, pp.1.14, 2019.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Statistics for Continuous Improvement of Manufacturing Process of Ultrasound Probes for Medical Imaging

Francesco, Bertocci[1], Andrea Grandoni[1], Monica Fidanza[1], Rossella Berni[2]

[1)] R&D Global Transducer Technology, Esaote spa, Via di Caciolle, 15, 50127, Firenze (Italy)
[2)] University of Florence, Department of Statistics, Computer Science, Applications "G. Parenti", Firenze (Italy)

[*]Francesco Bertocci - e-mail address: francesco.bertocci@esaote.com

Keywords: Plan-Do-Check-Act (P.D.C.A.); problem solving; statistical methods

ABSTRACT – The advanced robust process optimization techniques are widely adopted in the manufacturing field in order to enhance the productivity and to make the process reliable by limiting the variability and the scraps. Statistical methods for quality control and for higher productivity achievement, together with engineering techniques oriented to problem solving (including failure analysis, reverse engineering, electronic and acoustic microscopy and brainstorming of internal know-how), plays a key role for improving the manufacturing process in industry 4.0 scenario.

Starting from the high definition of the process, and thus providing a successful implementation of the PDCA (Plan-Do-Check-Act) methodology, the analysis is conducted via statistical models and through the detection of latent failures by means of NDT (Non-Destructive Testing), i.e. Scanning Acoustic Microscopy (SAM). The presented approach driven by the statistical analysis allows the engineering to distinguish the potential weak points of the manufacturing process of ultrasound (US) probes and to implement the corrective actions. The study leads an advantage for the manufacturers of US probes that are oriented toward continuous improvement devoted to the process capability, and product quality.

1. INTRODUCTION

Quality Control methodologies, such as Six Sigma, has been adopted by leading companies throughout the world for quality and process improvement aiming to enhance an organization's competitiveness [1]. The PDCA cycle (called also as Deming cycle) is used for making improvement in the existing processes in an industrial scenario. More recently, in a highly competitive and demanding electronics and biomedical market, the research studies can be implemented through advanced statistical methods, i.e. experimental designs and statistical modelling, jointly with the advanced instrumentation for defect detection, i.e. Scanning Acoustic Microscopy. The latter is a robust NDT technique that provides an efficient solution for quick identification and location of failure in the bulk of US probes [2].

The current study is related to a convex array probe for medical imaging (Figure 1) devoted to abdominal and women health diagnosis [3].

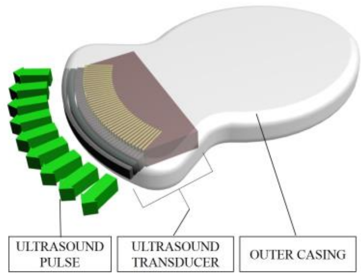

Figure 1 Ultrasound probe devoted to medical imaging: the structure in brief

2. METHODOLOGY

The PDCA cycle approach helped the company to identify the major root cause that limit the quality of the manufacturing process (Figure 2) and to improve the know-how on the manufacturing process. Following the Taguchi's philosophy, an inter-disciplinary team composed by engineers and physicists of R&D (Research & Development) of the US probe, mechanical designers together with the production and people in charge of data collection on M.E.S. (Manufacturing Execution System) platform has been defined. The beginning step consisted in the identification of the weak points, assessing a classification of the risk level-based critical phases with the aim to obtain benefits through corrective actions in a short time. For complex manufacturing such as US probes became obvious that the improvement of process efficiency needed medium-long term, due to critical phases of the production process in which different materials are involved. Furthermore, some stages are worker dependent.

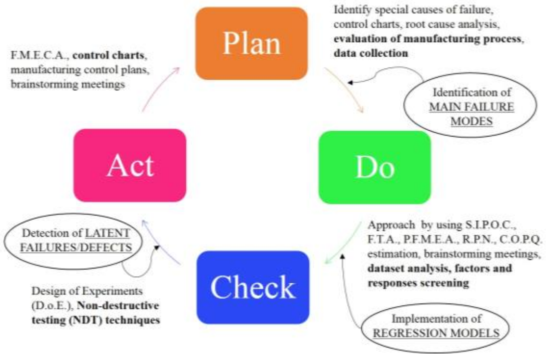

Figure 2 P.D.C.A. (Plan-Do-Check-Act) cycle for the continuous improvement in US probes manufacturing process

Therefore, the manufacturing process has been analyzed not only with process map, but also involving Failure Modes and Effects Analysis (FMEA) and Fault Tree Analysis (FTA), bottom-up and top-down techniques respectively. Moreover, the initial phases of the manufacturing process were certainly the most complex to analyze, because it is just during these initial steps that the main failure may occur, or some latent defects, risen during the construction of the US probe, may become evident. Therefore, in order to analyze these issues, it becomes strictly relevant the strenuous collection of historical (observational) data, based on processing cards and control charts. To this end, the use of M.E.S. system for internal data management allowed for implementing a dataset on an electronic spreadsheet.

The statistical modelling applied through a dedicated software (S.A.S. - Statistical Analysis System) has been implemented. In this context a method for the detection and the visualization of latent defects on US probes for medical imaging was carried out by means of SAM.

Once the corrective actions have been applied, graphical analyses such as control chart are utilized for monitoring and measuring the effectiveness of the improvement in the process.

3. RESULTS AND DISCUSSION

This study, in which the collaboration between the engineering and the statisticians is fundamental, provided clear and specific advantages, is articulated through the following step:

a. the re-examination of each ultrasound probe manufacturing phase under critical point of view;
b. the identification and the analysis of factors, never evaluated by the engineering, that can influence the variability of the production process, by also considering the distinction between systematic, noise and block effects for defining and planning the design of experiments.

Through this rigorous approach, the engineers, in agreement with the statistics, were able to manage and to organize the analysis by considering more than 36 different factors, and 38 response variables, both quantitative and qualitative.

Two different statistical models have been implemented, in which the continuous discussion between the engineering and the statisticians led to the identification, by successive refinements, of the final set of factors considered most significant within the process variability. As a contribution to reduce the failure probability, the two selected statistical models were used for implementing the actions necessary to improve of the production process.

The statistical results revealed:

1) significant interactions among factors;
2) the operators involved in the manufacturing play a key role in the variability of the production process;
3) the need to implement a D.o.E. (Design of Experiments) in the early stages of the production process for a better screening of the principal critical factors.

The statistical study led to implement SAM analysis campaign able to detect a latent failure as shown in Figure 3.

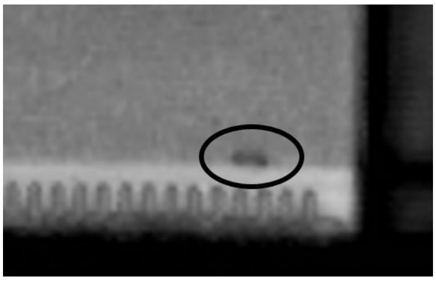

Figure 3 A latent defect in the black oval: crack in the US probe bulk. The detection is provided by SAM

4. CONCLUSIONS

The study presented the application of advanced methods for robust process optimization to identify the root causes in a US probes manufacturing process. In this context the multidisciplinary research was been decisive for the improvement of the production process. A new team of five dedicated operators has been formed, in order to reduce the variability of the production process and increase the quality of the product. The in-line inspection based on SAM analysis and the proposed statistical study can open a tangible contribution for the industry 4.0, in which the real time information is the key aspect for the continuous improvement.

5. REFERENCES

[1] E.V. Gijo, and J. Scaria, “Process improvement though Six Sigma with Beta correction: a case study on manufacturing company”, *The International Journal of Advanced Manufacturing Technology*, vol. 71, pp. 717–720, 2014.

[2] F. Bertocci, A. Grandoni, and T. Djuric-Rissner, “Scanning Acoustic Microscopy (SAM): a Robust Method for Defect Detection during the Manufacturing Process of Ultrasound Probes for Medical Imaging”, *Sensors*, vol. 19, pp. 4868–4887, 2019.

[3] R. Berni, F. Bertocci, N.D. Nikiforova, and G.G. Vining, “A tutorial on randomizing versus not randomizing Split-Plot experiments”, *Quality Engineering*, vol. 0, pp. 1–21, 2019.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Big data classification in the presence of label and measurement noise

Bart De Ketelaere[1*], Mia Hubert[2], Jakob Raymaekers[2], Peter Rousseeuw[2], Iwein Vranckx[1,2]

[1)] Mechatronics, Biostatistics and Sensors (MeBioS), Katholieke Universiteit Leuven, Belgium.
[2)] Statistics and Data Science Section, Katholieke Universiteit Leuven, Belgium

[*]Corresponding author - e-mail address: bart.deketelaere@kuleuven.be

Keywords: Classification; Robust statistics; Big data

ABSTRACT – We introduce an efficient approach for classifying noisy data in a big data context. The approach is based on recently published high-performant robust methods and takes into account noise both on the measurements as well as on the labels. We demonstrate the capabilities of the method on relevant data.

1. INTRODUCTION

Advanced sensor technologies are often used in modern industries to discriminate between good and bad product, or to separate many fractions that constitute a material stream (e.g. in food or plastics inspection). In order to get most out of the sensor systems, they have to be "trained" using a classification algorithm that can cope with prevalent practical challenges. Two of the most encountered challenges are label and measurement noise. Label noise (or mislabeling) occurs when a wrong label is attached to a given measurement, e.g. a label "good" is given to a "bad" product. Measurement noise is the presence of erratic sensor measurements, originating from a sensor fault or any disturbing factor that causes the sensor to output outlying values. Besides, the fact that sensor systems become faster and generate more data than ever is an additional challenge modern systems have to cope with.

Robustifying classification algorithms has been proposed by several authors. For example, Hubert and Van Driessen [1] proposed a robust version of quadratic discriminant analysis (QDA) by replacing the classical estimates of location and scatter by their robust counterparts. More specific, they proposed to use the (reweighted) Minimum Covariance Determinant (MCD) estimator which has very much desired properties in terms of robustness but is computationally expensive, making it unsuitable in a big data context where large amounts of data need to be processed in a fraction of a second. Recently, De Ketelaere et al. [2] proposed a novel algorithm that is based on the MCD principle but is several orders of magnitude faster so that it can be used in a big data context.

In this extended abstract, we briefly describe the proposed classification algorithm and show its performance under varying conditions.

2. METHODOLOGY

The proposed algorithm starts from the classical QDA algorithm and makes two main adaptations:

(1) The RT-DetMCD algorithm (De Ketelaere et al., 2020) is used to estimate location and scatter. It starts from the DetMCD algorithm introduced by Hubert et al. [3] but replaces the six initial estimators by two new ones proposed in Raymaekers et al. [4]. It further introduces a parallel topology during which the dataset is split into q partitions and applies refinement on these subsets, before merging them into one final estimator. These main changes make the algorithm several orders of magnitude faster than the DetMCD algorithm, especially when the dataset becomes large.

(2) After plugging in the robust estimators in the QDA context, we apply tolerance ellipses to allow measurements to be considered as outlying in their measurement values. Furthermore, this allows us to separate this measurement noise (i.e. measurements that have low probability to belong to any predefined class) from label noise (i.e. measurements that do have high probability of belonging to a given class but that were given erratic labels).

In order to demonstrate the ability of our proposed algorithm, we performed a simulation study, as well as an analysis of a real dataset in an Industry 4.0 context.

In our simulation study, we start from a three-class problem in three dimensions, each of the classes drawn from a Gaussian distribution with different location and scatter. We consider different outlier models (shift, cluster, point) to investigate the broad performance. Finally, label noise is introduced by randomly swapping a portion of the class labels. The sample sizes for the three classes were chosen to be 10,000, 6,000 and 4,000, respectively.

As an example of how the technology can be applied to an Industry 4.0 problem, we consider the classification of objects based on their multispectral reflection. In order to label each object, a mixture of material fractions is presented to a camera system, where a human operator manually labels the different fractions in the acquired images. This process is prone to human error thus leading to label noise. Besides, the measurements themselves can be contaminated due to bad sensor readings (measurement noise). A typical datasets consists of >1 million observations that need to be classified in a very short time (typically less than a second). We will present some speedup values in this extended abstract, and will give more flavor on the actual case in the presentation.

3. RESULTS

3.1 *Simulation study*

For the simulation study, we summarize results for a specific setting that includes label as well as measurement noise in the training set (Table 1). In this table, classical QDA (left part) is compared to the proposed methodology (right part) in the case 10% measurement noise is present as well as 10% label noise in the training set. Rows represent the real (true) class whereas columns denote classification results. Remark the presence of a column labeled 'π_0' which is used for measurements that fall outside all tolerance ellipses, thus indicative of measurement noise. All results shown relate to the performance on an independent test set with increasing fraction of measurement noise (ϵ).

Table 1 Classifiers trained with 10% label and 10% measurement noise, where ϵ outliers are inserted in the test set.

		Classic				RT-DetMCD			
ϵ		π_1	π_2	π_3	π_0	π_1	π_2	π_3	π_0
A: Test set measurement contamination									
0.0	π_1 :	0.76	0.20	0.03	0.00	0.80	0.09	0.10	0.01
	π_2 :	0.49	0.46	0.04	0.00	0.09	0.80	0.10	0.01
	π_3 :	0.47	0.36	0.17	0.00	0.09	0.11	0.79	0.01
0.05	π_1 :	0.79	0.16	0.06	0.00	0.78	0.09	0.07	0.06
	π_2 :	0.41	0.50	0.08	0.01	0.07	0.78	0.09	0.06
	π_3 :	0.39	0.28	0.33	0.00	0.09	0.10	0.75	0.06
0.1	π_1 :	0.82	0.13	0.04	0.00	0.75	0.08	0.06	0.11
	π_2 :	0.44	0.50	0.04	0.01	0.07	0.71	0.11	0.11
	π_3 :	0.62	0.18	0.20	0.00	0.07	0.10	0.72	0.11
0.15	π_1 :	0.86	0.09	0.06	0.00	0.68	0.08	0.08	0.16
	π_2 :	0.41	0.48	0.08	0.02	0.07	0.68	0.09	0.16
	π_3 :	0.42	0.17	0.40	0.00	0.07	0.09	0.68	0.16
0.2	π_1 :	0.89	0.08	0.03	0.00	0.65	0.08	0.06	0.21
	π_2 :	0.50	0.41	0.08	0.01	0.07	0.63	0.09	0.21
	π_3 :	0.52	0.11	0.37	0.00	0.07	0.08	0.64	0.21

The first observation is the fact that in case of the classical QDA tolerance ellipses are inflated during the model building phase so that almost no test set measurements fall outside of them. This is in strong contrast with the proposed methodology that clearly is able to label these outlying observations as measurement noise (π_0).

The second observation from Table 1 is the fact that classical QDA tends to break in case of measurement and label noise in the training set. This can easily be derived from inspecting the noise-free test set case ($\epsilon = 0$) where about half of the samples from class 2 are wrongly assigned to class 1. In case of our robust methodology, this is only 9%, much resembling the performance on the training set (9%, data not shown).

3.2 *Industry 4.0 case*

Table 2 shows the speedup that is attained when comparing the proposed methodology to the reference robust statistical approach that includes the standard DetMCD approach [3]. A typical multispectral setting with up to 16 dimensions is considered, where the number of samples is increased up to 2^{20}. It can easily be seen that especially in the case where datasets become large, the speedup is very strong, with a 10,000 times faster computation when 2^{20} measurements are considered in 8 dimensions. Remark that the speedup in case of 16 dimensions is not calculated due to the fact that the classical DetMCD algorithm took more than 24h on a typical CPU.

Table 2 Speedup in a Industry 4.0 context. ϵ denotes fraction of measurement noise as before, ω denotes the size of the partitions (De Ketelaere et al. [2]).

$\epsilon = 0.3$	$\omega = 2^{12} = 4096$		
n	$p = 4$	$p = 8$	$p = 16$
2^{10}	6.7	10.8	13.8
2^{11}	9.1	12.7	17.3
2^{12}	13.7	19.2	22.9
2^{13}	25.8	32.2	37.3
2^{14}	49.0	63.2	72.8
2^{15}	160.3	121.8	128.9
2^{16}	489.7	386.9	228.9
2^{17}	1188.2	1056.8	769.0
2^{18}	2489.7	2447.2	2152.3
2^{19}	5020.6	5251.2	5138.4
2^{20}	10 180.9	10 588.2	NA

4. CONCLUSIONS

We introduced a methodology for data classification in an Industry 4.0 context where label as well as measurement noise is present. We showed the ability of the method to handle a substantial amount of noise, as well as its speedup when compared to other robust methods.

5. REFERENCES

[1] M. Hubert and K. Van Driessen, "Fast and robust discriminant analysis", *Computational Statistics & Data Analysis*, Vol. 45(2) pp. 301-320, 2004.

[2] B. De Ketelaere, M. Hubert, J. Raymaekers, P.J. Rousseeuw, and I. Vranckx, "Real-time outlier detection for large datasets by RT-DetMCD," *Chemometrics and Intelligent Laboratory Systems*, under review, 2020.

[3] M. Hubert, P.J. Rousseeuw, and T. Verdonck, "A deterministic algorithm for robust location and scatter", *Journal of Computational and Graphical Statistics* Vol. 21, pp. 618–637, 2012.

[4] J. Raymaekers, and P.J. Rousseeuw, "Fast robust correlation for high dimensional data", Technometrics, DOI: 10.1080/00401706. 2019.1677270.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Robust learning for classification

Jesús López-Fidalgo[1*], Douglas P. Wiens[2]

[1)] Institute for Culture and Society, University of Navarre; Biblioteca de Humanidades; Pamplona, Spain 31009.
[2)] Department of Mathematical and Statistical Sciences; University of Alberta, Edmonton, Alberta; Canada T6G 2G1. e-mail: doug.wiens@ualberta.ca

[*]Corresponding author - e-mail address: fidalgo@unav.es

Keywords: Generalized linear models; optimal subsampling; sequential sampling

ABSTRACT – We introduce a method of Robust Learning for classification, and propose its use in situations where Active Learning is appropriate, and where sampling the predictors is easy and cheap, but learning the responses is hard and expensive. As well, we seek robustness against both modelling and misclassification errors. We seek to sample effectively from the population of predictors, and learn the responses only for an influential sub-population. It is somewhat remarkable that our sampling weights eliminate the bias due to both types of error. The estimation weights are then chosen to minimize the variance. Simulation and real-data confirmed the theory.

1. INTRODUCTION

With the advent of Big Data huge, indeed astronomical data sets are becoming increasingly prevalent. With them comes an increased need for data reduction, prior to the analysis, in order to reduce not only the computational burden but, as explained below, the costs of sampling. For this, data subsampling has been proposed [8, 5, 3, 6, 4, 7]. This in turn leads to revisiting methods of Optimal Experimental Design (OED) as [3] and [5] point out, the goals and techniques are not entirely dissimilar.

The major difference is perhaps that in OED we have complete control over the predictors, whereas in Sampling Design (SD) these and perhaps the responses as well, are sampled. But the sampling need not be done randomly. Simple Random Sampling (SRS) is termed Passive Learning (PASSL) in the machine learning literature; statisticians tend to be drawn instead to Active Learning, whereby the sampling is done in a more controlled, supervised manner, akin to OED.

[8] proposed a method of subsampling in active learning, with the goal being classification. In [7] this was done for linear regression instead. There are two major ways in which that work departs from what is presented here, in a classification context,

(i) The framework of [8, 7] was that one was to sub-sample from a population containing both predictors and responses. Here we instead envision a scenario in which predictors can be sampled quickly and inexpensively (e.g. through electronic medical databases), whereas the responses are much more difficult to obtain (e.g. medical outcomes). One hopes to have to query the responses only in a relatively small number of cases, although the number of predictors sampled might be much larger.

(ii) In [8, 7] the model fitted to the data was assumed to be of a correct form. Here we place only slight and tentative faith in the fitted model, and instead seek sampling designs which are robust against both model misspecifications, and possible misclassification errors. The former goal again has been well-developed in the field of Robustness of Experimental Design, see [9] for a comprehensive review.

We assume that, once gathered, the data will be analyzed by fitting a Generalized Linear Model (GLM), through solving a weighted likelihood equation. As a result of the robustness issues raised above, the asymptotically normal estimates and resulting classifiers are biased. We propose a sampling scheme whereby, for given weights, this asymptotic bias is eliminated. It seems somewhat remarkable that both sources of bias can be simultaneously eliminated. We then propose weights which minimize the variance. Comparisons with PASSL illustrate the gains to be made, in terms of decreased numbers of responses required to be queried, from our Robust Learning (ROBL) approach,

2. METHODOLOGY

We aim to develop a theory of robust classification, and a methodology for applying this to large data sets such as arise in machine learning. The general idea is that there is a (large) population $Q \in R^k$

of explanatory variables, which can be easily sampled. We denote by X the set of distinct values of $x \in Q$. With a certain probability, possibly depending on unknown parameters θ, an item with covariates x belongs to group A (indicated by $Y = 1$), otherwise it belongs to group B ($Y = 0$). We suppose that the determination of the appropriate group, given x, is difficult and expensive, so that the investigator wishes to sample from Q in a manner which is more efficient than random sampling.

More precisely, we let $q(x)$ be the probability mass function (pmf) of x. This is the relative frequency with which x appears in Q, and need not be known - SRS from Q yields an empirical pmf on X, converging in probability to q. We wish to find an optimal pmf $p(x)$ on X, and sample from the sub-population with this distribution. This will involve probability weighted sampling from X -.choose x with probability $q(x)$, and then accept this into the sample according to the weights $\pi(x) / p(x) = q(x)$. For this $p(x)$ need not be known.

We consider a generalized linear model framework, in which the investigator assumes, and fits, the possibly incorrect model,

$$P(1|x) = \Pr(Y = 1|x) = G(\beta(x, \theta)),$$

for a strictly increasing, absolutely continuous distribution function G with a density g possessing two bounded derivatives. We concentrate on logistic and probit regression. We take $\beta(x, \theta) = f'(x)\theta$ for regressors $f(x)$, and an unknown, d-dimensional parameter θ ranging over a compact, convex space.

Our focus is on efficient and robust estimation and prediction of $P(1|x)$. To quantify the robustness against both modelling and misclassification errors, first define

$$\beta^*(x, \theta_\gamma) = f'(x)\theta_\gamma + \varphi(x),$$

where θ_γ is the .true parameter under the misclassification model described in the next paragraph, and $\varphi(x)$ represents unknown model error constrained by

$$E_q[\varphi(x)^2] \leq \frac{\tau_1^2}{n},$$

for a constant τ_1 controlling the magnitude of the error in the incorrectly fitted model.

We suppose that

$$P(1|x) = (1 - \gamma)G(\beta(x, \theta)) + \gamma(1 - G(\beta(x, \theta))),$$

where $\gamma \in [0; 1]$ is a misclassification probability (see [2, 1]). We assume that

$$\gamma \leq \min\{\frac{\tau_2}{\sqrt{n}}, 0.5\},$$

for a user-specified τ_2.

3. RESULTS AND DISCUSSION

The main theoretical result establishes the asymptotic property that the unconditional mean squared error matrix of $\sqrt{n}\left(\hat{\theta} - \theta_0\right)$, is

$$MSE_{\varphi,\gamma}(\theta_0) =$$
$$M_1^{-1}(\theta_0)\{M_2(\theta_0) - nb_{\varphi,\gamma}(\theta_0)b'_{\varphi,\gamma}(\theta_0)\}M_1^{-1}(\theta_0) + O_p(n^{-1}),$$

as n tends to infinity, where matrices and vectors are computed through the model and weights based on influence functions. This result allows us to look for optimal sampling.

4. CONCLUSIONS

Table 1 shows a remarkable descent in the misclassification error in a real example classifying tweets as human or bot precedence. Columns 4th and 6th show means and standard deviations of the misclassification errors and columns 5th and 7th the sample sizes required.

Table 1. Tweet classification: Mean error rates (std. errors in parentheses) and mean sample sizes in 25 simulations with an initial random sample of size 30

Mean error rates and sample sizes						
α	**link**	**Class'n method**	**ROBL**	**#queries**	**PASSL**	**#queries**
.05	logistic	standard	.386 (.017)	176	.415 (.010)	419
	logistic	Neural net	.323 (.016)		.295 (.009)	
.05	probit	standard	.368 (.019)	171	.400 (.016)	458
	probit	Neural net	.333 (.014)		.286 (.004)	

5. ACKNOWLEDGEMENT

López–Fidalgo is sponsored by Ministerio de Economía y Competitividad MTM2016-80539-C2-1-R and by Program Salvador de Madariaga PRX18/00339. Wiens is sponsored by the Natural Sciences and Engineering Council of Canada. The authors thank L. Kong and J.A. Moler–Cuiral for helpful comments.

6. REFERENCES

[1] R. Carroll and S. Pederson. On robustness in the logistic regression model. Journal of the Royal Statistical Society, Series B, 55:693.706, 1993.

[2] J.P. Copas. Binary regression models for contaminated data. Journal of the Royal Statistical Society, Series B, 50:225.265, 1988.

[3] Ch.C. Drovandi, Ch.C. Holmes, J.M. McGree, K. Mengersen, S. Richardson, and E.G. Ryan. Principles of experimental design for big data analysis. Statistical Science, 32:385.404, 2017.

[4] A.C. Nachtsheim and J. Stufken. Comments on: Data science, big data and statistics. TEST, 28:345.348, 2019.

[5] R. Nie, D.P. Wiens, and Z. Zhai. Minimax robust active learning for approximately specified regression models. Canadian Journal of Statistics, 46:104.122, 2018.

[6] Ch. Wang, M.-H. Chen, J. Wu, J. Yan, Y. Zhang, and E. Schifano. Online updating method with new variables for big data streams. Canadian Journal of Statistics, 46:123.146, 2018.

[7] H.-Y. Wang, M. Yang, and J. Stufken. Information-based optimal subdata selection for big data linear regression. Journal of the American Statistical Association, 114:393.405, 2019.

[8] H.-Y. Wang, R. Zhu, and P. Ma. Optimal subsampling for large sample logistic regression. Journal of the American Statistical Association, 113:829.844, 2017.

[9] D.P. Wiens. Robustness of Design, in .Handbook of Design and Analysis of Experiments. Chapman & Hall/CRC, 2015.

On some issues of DoE and Robust Design in experimental mechanics supported by numerical simulations

Umberto Galietti*, Davide Palumbo, Rosa De Finis, Francesca Di Carolo

[1)] DMMM - Politecnico di Bari.

[*] e-mail address: umberto.galietti@poliba.it

Keywords: Test Design, Experimental Mechanics

ABSTRACT – The use of complex analytical models can be a starting point to assess the capability of an experimental technique to obtain the required measurement with the available sensors.
A general model in Thermoelastic Stress Analysis was used jointly with the simulation of possible sources of noise for an error assessment of the experimental technique providing information also on how to perform a proper test when the residual stress is the wanted quantity and not and unknown source of error.
The proposed procedure showed the capability to help in properly design the test setup starting from simulations.

1. INTRODUCTION

One of the main aims of experimental mechanics is the measurement of mechanical quantities called stresses and strain on bodies that have to carry static or dynamic loading. Despite numerical modelling (Finite Element Analysis) is nowadays the main design tool, the need to validate those models require a good capability to measure stresses and strains.

The Thermoelastic Stress Analysis is a very powerful technique. In particular conditions (adiabatic conditions, homogenous and isotropic materials) it is based on a very simple equation that provide a very simple relationship between the change of Temperature and the changes of the applied load [1].

Measuring the change of temperature with a high level Thermocamera the change of the stresses can be assessed [2]. In this equation the change of stress and then change of temperature are in terms of amplitude of the quantity due to dynamic load variation, usually set as sinusoidal[1,2].

Some materials, such as titanium and aluminum behave in a more complex way and the general equation showed in eq. 1 [3,4] describes an influence also of the mean stress and the residual stress. Of course neither the residual stress nor the mean stress vary with the time.

In the showed equation the relation among general variables are in terms of change amplitude.

$$\Delta T = (\rho C_\varepsilon)^{-1} T_0 \left[\left(\bar{\bar{C}}^{-1} \left(\bar{\sigma}_m + \bar{\bar{R}} \bar{\sigma}_r \right) \right)^T \frac{\partial \bar{\bar{C}}^T}{\partial T} - \bar{\alpha}^T \bar{\bar{C}}^T \right] \bar{\bar{C}}^{-1} \Delta\bar{\sigma} \quad (3)$$

Where: ΔT are the temperature changes, $\bar{\alpha}^T$ is the vector of linear expansion coefficients, ρ is the density, C_ε is the specific heat at constant strain, $\Delta\bar{\sigma}$ stress amplitude vector, $\bar{\sigma}_m$ mean stress vector, $\bar{\bar{R}}\bar{\sigma}_r$ is the residual stresses vector in the absolute reference system, T_0 is the reference temperature, $\bar{\bar{C}}$ is the stiffness matrix.

On one side this means that not considering these effects will lead to measurement errors [4, 5] and on the other side this can be a good opportunity to assess residual stresses, very difficult to be measured otherwise on the field [6].

Aim of this work is to present the results obtained using statistic applied on an analytical model simulating also the various types of errors that can be made on the process. The same model is used: to assess the Interval of Confidence of the measured Stresses under different error conditions; to assess the β error so to evaluate the minimum residual stress measurable and to assess residual stresses with different set-up conditions to obtain the highest Signal to Noise ratio [7].

2. METHODOLOGY

Despite the presence of an analytical model, deal with this second order effect is not easy at all. There are numerous variables with different kind of error that can be made.

The same variable involved in this problem can be Input factors, Noise factors or even output factors according to the aim of the analysis. For example residual stresses can be noise factor when Δσ is assessed but the residual stress filed is unknown of neglected, can be an output if it is the investigated variable.

For all analyses carried out the fixed variables are: $\bar{\sigma}_m$ mean stress, fixed as proportional to Δσ; Mat – material and its characteristics;

This study forecasted different steps detailed in table 1:

Table 1. Analysis performed with the definition of the investigated variables

Analysis	**Input**	**Noise**	**Output**
Error Analysis	$\Delta\bar{\sigma}$	$TCNoise$ MatNoise; $\bar{\bar{R}}\bar{\sigma}_r$	ΔT; $IC_{\Delta T}$ $\Delta\sigma_{meas}$; $IC_{\Delta\sigma meas}$
Sensitivity to residual stress analysis	$\Delta\bar{\sigma}$; $\bar{\bar{R}}\bar{\sigma}_r$	$TCNoise$;	β error
Residual stress measurement	**set up** $\Delta\bar{\sigma}$ and ϕ 3 levels of $\Delta\bar{\sigma}$; ϕ ΔT_{sim}	$TCNoise$;	$\bar{\bar{R}}\bar{\sigma}_r$; $IC_{R\sigma r}$ SN ratio($\bar{\bar{R}}\bar{\sigma}_r$)

Where input factors are: $\Delta\bar{\sigma}$= Stress change vector; $\bar{\bar{R}}\bar{\sigma}_r$ is residual stress vector expressed in the applied stress orientation system; Set up $\Delta\bar{\sigma}$ and ϕ (i.e different load configurations in terms of stress amplitude and applied load angle); level of $\Delta\bar{\sigma}$ and ϕ (the needed 3 values to define the solution for the output variable that is a vector): ΔT_{sim} (Temperature amplitude obtained by the model using all correct values of the variable as input).

The main Noise parameters are: TCNoise (Experimental Noise due to the measurement chain - experimentally assessed); MatNoise (Material Characteristics that are not the nominal ones – literature assessed); $\bar{\bar{R}}\bar{\sigma}_r$ considered source of noise in the error analysis when it is unknown and not neglectable.

The output for the analyses are quite different. At first the information needed is about the possible interval of confidence of the Δσ obtained from TSA, then a β error analysis is performed to assess the minimum level of residual stress that provides significant change of the measured ΔT and at last, but not at least, a comparison is obtained amongst different way to load the component, in order to enhance the results, when the residual stress assessment is performed.

3. RESULTS AND DISCUSSION

The error analysis provided information on the variables for which eventual error has more influence. Moreover the error in the measurement increase from a quite usual 3% in case of presence of only TCNoise to more than 70% in case of unknown residual stresses. An example of the results obtained is reported in Fig. 1 where the influence of the residual stress is showed when an experimental error is introduced.

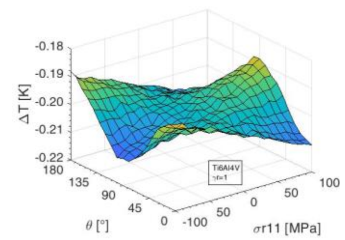

Figure 1 Influence of the residual stress considered as noise for Δσ measurement

The β error analysis provided a model that is a function of the applied stress. As shown in fig. 2 it is possible to assess the minimum significant residual stress level.

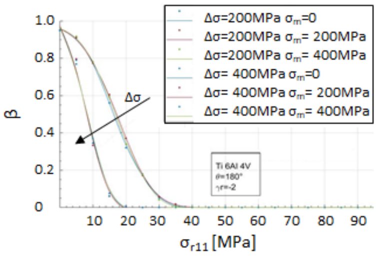

Figure 2: example β error as function of $\bar{\bar{R}}\bar{\sigma}_r$

For the setup analysis the model provided indication of the importance to forecast applied loads with different angle.

4. CONCLUSIONS

Test simulation analysis can be a powerful tool to design the best setup when a model is available. The possibility to use statistical tools allow an easy use also of very complex models providing information on both possible errors induced by neglected variables or underestimated variables (materials characteristics).

In particular a very important information to TSA operator concern the importance of the residual stress orientation on the measured stress variation with a Thermoelastic Stress Analysis.

In this work an articulated simulation campaign allowed to define also the better setup (considering all needed compromises) in term of statistically significant residual stress measurement reducing the variability of the measurement.

5. REFERENCES

[1] N. Harwood and W.M Cummings, "*Thermoelastic Stress Analysis*, Philadelphia: Adam Hilger; 1991.
[2] G. Pitarresi, E.A. Patterson, "A review of the general theory of thermoelastic stress analysis", *J Strain Anal Eng*, vol. 38, pp. 405-417, 2003.
[3] A.K. Wong, S.A. Dunn, J.G. Sparrow, "Residual stress measurement by means of the thermoelastic effect", *Nature vol. 332*, pp. 613-615, 1988.
[4] F. Di Carolo, R. De Finis, D. Palumbo, U. Galietti. "A Thermoelastic Stress Analysis General Model: Study of the Influence of Biaxial Residual Stress on Aluminium and Titanium", *Metals*, vol. 9, pp. 671, 2019;
[5] U. Galietti, D. Palumbo. "Thermoelastic stress analysis of titanium components and simultaneous assessment of residual stress". In *Advanced Developments Applied to Structures, Materials and Environmental Issue*: *Thermomechanics, Biomechanics, Fluids-solids Interaction,* Poitiers, France, 4-9 July.
[6] J. Lu, *Handbook of Measurement of Residual Stresses, 1st* ed., Fairmont Pr, 1996
[7] D. Montgomery, Design and Analysis of Experiments, 8th ed., John Wiley & Sons Inc, 2012.Materials Technology Conference and Exhibition, 2014, pp. 1–4.

Reliability verification as integrated part of a product development process

Nikolaus Haselgruber

CIS consulting in industrial statistics, A-4170 Haslach/Austria.

nh@cis-on.com

Keywords: reliability, verification

1. ABSTRACT

This talk discusses the integration of reliability verification methods in a product development process with the objective that the fulfillment of the reliability requirements can be confirmed prior to product release.

2. INTRODUCTION

Reliability is a key characteristic for high-tech products, in particular if those are in commerical use. Then, unreliability leads directly to downtime, an interruption of its function, and finally costs for both repair as well as compensation of losses. Corrective actions are required to restore the original functional state. This needs understanding of the failure mechanisms, i.e., the root causes before countermeasures can be developed and implemented. They have to be verified to confirm their effectivity.

For long time, reliability activities were synonym to lifetime or failure rate analysis following the classical scheme *data – model – inference*. I.e., reliability was mainly a reactive task but it has changed to a proactive discipline. The contents of important reliability textbooks around the turn of the 21st century, such as Nelson (1990), Meeker and Escobar (1998), Birolini (1999), Rigdon and Basu (2000), Singpurwalla (2006), Hamada et al. (2008) confirm this by increasing the focus on reliability verification.

3. METHODOLOGY

The integration of reliability verification activities in a product development process means to design, develop, train, execute, track and improve reliability methods starting from the concept phase throughout and beyond market release to avoid unreliability as soon as possible or at least detect it early than the customer.

The spectrum of methods contains the derivation of adequate reliability targets, a risk assessment to decide for which parts or components of the product to carry out the reliability verification as well as actual verification methods, such as design of accelerated tests, lifetime tests under consideration of competing risks, reliability growth monitoring and testing as well as early fault tree analyses in case of failures where the root cause has not been understood sufficiently.

An important aspect is the selection of appropriate models and parameter estimation procedures. Both depends on the availability and completeness of data as well as on the knowledge about the product or service. In particular, Bayesian approaches appear very effective as long as relevant prior information is available.

4. RESULTS AND DISCUSSION

The result of this work is a roadmap throughout a development process, indicating adequate methods for the different process phases. Reliability methods are proposed for every phase and investigated regarding requirements and suitability.

An example from the automotive industry is persented to discuss the methodology described above where specific reliability verification activities come effectively into play along the development process.

5. CONCLUSIONS

The reliability verification strategy needs to be tailored to the requirements of the product as well as to the availability of reliability data in different stages of the development process.

6. REFERENCES

[1] Nelson, W. (1990): Accelerated Testing. Wiley, New York.
[2] Meeker, W.Q and L.A. Escobar (1998): Statistical Methods for Reliability Data. Wiley, New York.

[3] Rigdon, S.E., and A.P. Basu (2000): Statistical Methods for the Reliability of Repariable Systems, Wiley, New York.
[4] Singpurwally, N. D. (2006): Reliability and Risk, Wiley, New York.
[5] Birolini, E (2007): Reliability Engineering. 5th ed., Springer, Berlin.
[6] Hamada, M.S. et al. (2008). Bayesian Reliability. Springer, New York.

StEering
Inter-University Research Center:
Design, Quality and Reliability

An innovative methodology through optimal designs for freight trains optimization

Alessandro Magrini[1], Nedka Dechkova Nikiforova[1*], Rossella Berni[1], Luciano Cantone[2,] Gabriele Arcidiacono[3]

[1)] Department of Statistics Computer Science Applications "G. Parenti",
University of Florence, Viale Morgagni 59, 50134 Florence, Italy
[2)] Department of Engineering for Enterprise "Mario Lucertini", University of Rome "Tor Vergata", Rome, Italy
[3)] Department of Innovation and Information Engineering, Marconi University, Rome, Italy

[*]Corresponding author - e-mail address: n.nikiforova@unifi.it
Keywords: D-optimal designs; freight trains; optimization

ABSTRACT – This study aims to improve previous studies relating to the optimization of payload distributions of freight trains through the reduction of in-train forces.
Moreover, the innovative procedure, here briefly described, evaluates the successful Kriging researches and exploits the empirical evidences given by sets of "basic" freight trains. Optimal designs are then defined. Following, simulated trains are modelled and optimized in order to find the best empirical train, given specific and required constraints on hauled mass and train length.

1. INTRODUCTION

Previous researches [1, 2] are related to the innovative contribution consisting in applying Kriging models to the railway field, in order to optimize the payload distribution of freight trains. More precisely, starting from the first study [1], the last developments [3] are related to improve the selection of experimental runs involving Latin Hypercube Designs (LHD) generated through new orthogonal arrays, called strong orthogonal arrays, and recently developed in [4]. The Strong Orthogonal Array based LHD (SOA-based-LHD) is a further development of the Orthogonal Array based LHD (OA-LHD), illustrated in [5].
The new obtained designs are then applied, through the software *TrainDy* [6], to simulate and evaluate the braking performance (considered in terms of in-train forces exchanged by consecutive vehicles of train) of a freight train for container traffic. The SOA-based-LH design proposed in [3] allowed to describe accurately the behavior of the various payload distributions with a relatively small computational effort for the train simulations. Moreover, through the estimated Universal Kriging models it was possible to find the best payload distribution able to reduce the in-train forces for the specific train set arrangement considered.
Despite the benefits of the approach proposed in [3], further developments are essential in order to further improve the optimization of the in-train forces not only in terms of several train sections as in [3] but also in terms of the position of the various vehicles of the train. Moreover the entire procedure should be implemented in an automatic software program able to provide to the Railway Undertakings-RU, in a real time, the best train configuration. To this end, in this paper we suggest an innovative procedure for solving the gap existing between the achieved theoretical improvements in [3], and the empirical requirements when dealing with the composition of the vehicles of actual freight trains, given hauled mass and train length. The starting point is the use of sets of real freight trains generated through the consolidated and used algorithms [7, 8]. This key-point allows us to involve in the optimization procedure empirical information useful for better defining the Information Matrix, and then the optimal designs, subsequently applied to find the "best" freight trains configuration. Lastly, an optimization procedure will be implemented in order to find for each freight train the best composition/configuration of the train vehicles able to reduce the in-train forces.
The final aim, when the complete new procedure will be implemented, should allow to compose actual freight trains, respecting target requirements

(in-train forces), but with higher hauled mass, i.e. higher transportation efficiency. Moreover, one of the key-points of the proposal is that the entire procedure should be implemented in an automatic software program to be used for composing real freight trains by interested RU.

2. METHODOLOGY

The starting point of our proposed methodology is the generation of a set of freight trains through the consolidated algorithm in the railway field [7, 8]. Following, the "best" configuration of the generated freight trains is obtained by considering optimal design methodology [9]. More precisely, D-optimal experimental designs are built by considering the Universal Kriging model expressed as follows:

$$Y(\boldsymbol{x}) = \mu(x) + Z(\boldsymbol{x})$$
$$\mu(\boldsymbol{x}) = \boldsymbol{f}'(\boldsymbol{x})\boldsymbol{\beta} \qquad (1)$$

In formula (1), $\mu(\boldsymbol{x})$ is the deterministic part of the model in which $\boldsymbol{f}'(\boldsymbol{x})$ is a set of unknown trend functions, while $\boldsymbol{\beta}$ is the vector of unknown coefficients. Furthermore, $Z(\boldsymbol{x})$ is a Gaussian random process with zero mean and stationary covariance function expressed as follows:

$$Cov[Z(\boldsymbol{x}), Z(\boldsymbol{x}+\boldsymbol{h})] = \sigma_y^2 R(\boldsymbol{h};\ \boldsymbol{\omega}) \qquad (2)$$

where σ_y^2 is the process variance, and $R(\boldsymbol{h};\ \boldsymbol{\omega})$ is a stationary correlation function depending on the displacement vector $\boldsymbol{h}$ and on the set of parameters $\boldsymbol{\omega}$. In general, the mostly applied covariance functions are the Matérn, the Exponential, the Power Exponential and the Gaussian ones. For the case-study here illustrated, we consider the anisotropic Matérn covariance function with $\boldsymbol{\omega} = (\boldsymbol{\varphi}, v)$ and for $v = 3/2$ [10].
For obtaining the D-optimal experimental design, the Fisher Information matrix is derived for model (1), and it is expressed as follows [11, 12]:

$$I(\boldsymbol{\beta}, \boldsymbol{\varphi}) = \begin{bmatrix} X'C^{-1}X & 0 \\ 0 & \frac{1}{2} tr[C^{-1}\left(\frac{\partial}{\partial\varphi}C\right)C^{-1}\left(\frac{\partial}{\partial\varphi}C\right)] \end{bmatrix}$$

where C is the covariance function, formula (2). Once the D-optimal design is built, the modelling step and the optimization procedure are carried out.

3. PRELIMINARY RESULTS

Firstly, a set of N=50 basic freight trains are generated through the consolidated algorithm [7, 8] according to: i) a train mass defined in the range 1200-1600 tons, and ii) a train length defined in the range 600-700 meters. Following, the selected D-optimal experimental design obtained for this generated freight train set, is used to simulate the responses, e.g. in-train forces, through the software *TrainDy* [6] by considering three different shapes for the payload distribution (uniform, triangular and trapezoidal). Finally, all the procedure is completed by including the optimization step, in which the "best" train is defined conditioned to: mass, length, and shape for improving the braking system through in-train forces against derailments.

4. CONCLUSIONS

We have proposed an innovative methodology for optimizing the payload distribution of freight trains able to reduce the in-train forces. The preliminary results are promising; further developments are currently underway in order to obtain a complete and valid procedure to be successfully applied by interested Railway Undertakings.

REFERENCES

[1] G. Arcidiacono, R. Berni, L. Cantone, and P. Placidoli, "Kriging models for payload distribution optimization of freight trains", *International Journal of Production Research*, vol. 55, pp. 4878–4890, 2017
[2] G. Arcidiacono, R. Berni, L. Cantone, N.D. Nikiforova, and P. Placidoli, "Fast method to evaluate payload effect on In-Train forces of freight trains", *The Open Transportation Journal*, vol. 12, pp. 77–87, 2018.
[3] N.D. Nikiforova, R. Berni, G. Arcidiacono, L. Cantone, and P. Placidoli (2020). "Latin Hypercube Designs based on Strong Orthogonal Arrays and Kriging Modelling to Improve the Payload Distribution of Trains", (under review).
[4] Y. He and B. Tang, "Strong orthogonal arrays and associated Latin Hypercubes for computer experiments", *Biometrika*, vol. 100, pp. 254–260, 2013.
[5] B. Tang, "Orthogonal array-based Latin Hypercubes", *Journal of the American Statistical Association*, vol. 88, pp. 1392–1397, 1993.
[6] L. Cantone, "TrainDy: the new Union Internationale des Chemins de Fer software for freight train interoperability", *Journal of Rail and Rapid Transit*, vol. 225, pp. 57–70, 2011.

[7] UIC Leaflet 421, Rules of the consist and braking of international freight trains, 9th edition, January 2012.
[8] L. Cantone, "New Random Trains Generation According to UIC Code 421", *International Journal of Mechanical Engineering and Technology*, Vol. 9, no. 10, pp. 890-899, 2018.
[9] A.C. Atkinson, A.N. Donev, and R.D. Tobias (2007), *Optimum Experimental Designs, with SAS*, New York: Oxford University Press, 2007.
[10] C.E. Rasmussen, and C.K.I. Williams, *Gaussian Processes for Machine Learning*, Boston: The MIT Press, 2006.
[11] A. Baldi Antognini, and M. Zagoraiou, *Space filling and locally optimal designs for Gaussian Universal Kriging*. In: P. Mantovan, P. Secchi (Eds.), Complex data modeling and computationally intensive statistical methods. Springer, pp. 27-39, 2010.
[12] W. Müller and M. Stehlík, "Issues in the optimal design of computer simulation experiments" *Applied Stochastic Models in Business and Industry*, vol. 25, no. 2, pp. 163–177, 2009.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Differential effects of Millennial online brand community and customer loyalty in the fashion industry

Wilson Ozuem[1*], Michelle Willis[2], Guida Helal[3], Kerry Howell[4], Silvia Ranfagni[5] and Geoff Lancaster[6]

[1)] University of Cumbria, UK
[2)] University of Hertfordshire, UK
[3)] American University of Beirut, Lebanon
[4)] University of Plymouth, UK
[5)] Department of Economics and Management, University of Florence, Florence, Italy
[6)] London School of Commerce, UK
[*]Corresponding author - e-mail address: will.oz@talk21.com

Keywords: social media, online brand communities, electronic word of mouth, customer loyalty, qualitative research, millennials, fashion

1. ABSTRACT – In today's digital world, consumers share their lives on social media, interacting with peers about dizzyingly arrays of consumption behaviours and patterns in emerging online brand communities. Grounded in understanding that loyalty is seen and understood differently by people who inhabit and participate in online brand communities, this study is based on constructivist perspective combined with hermeneutic methodology and embedded case study research strategy to examine how online brand communities activate multi-dimensional customer loyalty commitment. Empirical data were generated through 37 in-depth interviews of Millennials. The analysis proposes a framework that categories customer loyalty into: Ambassador loyalists; Public-voting loyalists; Loveless loyalists; Mercenary-loyalists. Each stream contains one additional sub-category mediated by consumer levels of participation in online brand communities. This paper contributes to existing literature. Unlike extant studies, it specifically argues that customer loyalty commitment in online brand communities depend on individual people and context involved, and categorise loyalty into different levels. Practical steps by which companies may utilise these categories, and further theoretical implications for wider consideration, are proposed.

1. INTRODUCTION AND RATIONALE

Social media has revolutionised communication. It hosts worldwide user networks through real-time interactions to become a successful global communication medium (Alves, Fernandes & Raposo, 2016; Felix, Rauschnabel & Hinsch, 2017; Mas-Tur, Tur-Porcar & Llorca, 2016). Brands recognise opportunities to engage with consumers through social networking (Malthouse et al., 2013, Nisar & Whitehead, 2016). Social media channels are populated with brand-related activities connecting customers and brands through free-flow exchange of content (Ibrahim, Wang & Bourne, 2017; Ou, et al 2010). Brand-consumer dynamics have evolved, allowing users to access common platforms and interact with one another and brands alike (Chen, et al., 2013; Cheung, Chiu & Lee, 2011). A collective environment is born through virtual outlets, sparking cognitive and emotional exchanges among participants (McAlexander, Schouten & Koenig, 2002). Online social media brand communities provide consumers with co-creative shared platforms that enrich brand/customer relationships (Hajli, et al., 2017; Payne & Frow, 2005; Zwass, 2010). Millennials have innately adopted the shift to social media in engaging with brands (Helal, Ozuem & Lancaster, 2018). Degrees of engagement vary depending on motivation; information seeking is located on the passive spectrum of engagement as opposed to social interaction involving a more active approach (Khan, 2017).

Millennials represent the first tech-savvy age group to incorporate digital media in dayto-day communication. Millennial birth years fall between the early 1980s and 2000s (Ng, Schweitzer & Lyons, 2010; Rainer & Rainer, 2011). Howe & Strauss (2009) describe a technologically advanced generation born between 1982 and 2002 who value social networking convenience and high levels of digital involvement in socialisation, information sharing and shopping. Deloitte (2019), showed that millennials pursue or halt brand relationships based on insights into business operations and influences on society. Retailers have observed digital unfolding that has empowered active consumers with brand information (Hur, Lee & Choo, 2017). Social media manifestations of virtual communities and electronic word-of-mouth prompts peer influence over millennials' purchase decisions (Eastman &, Liu, 2012; Liu, Wu & Li, 2019; Akman & Mishra, 2017). Kong, et al., (2019) address the impact of social brand referrals on millennial trust within shared commerce settings. Consumers experience value through online brand community engagement contributing to outcomes of identification and trust.

Cheng, Wu & Chen (2018) highlight that information completeness and need for bridging social capital are predictors of customer satisfaction and relationship commitment for brand communities. Building on these ideas, Cheng and colleagues argued that the needs to acquire connection with new people is why people build relationships with online brand communities. Little is known about how consumer commitment in online brand community influences levels of loyalty commitment of millennials. Given its importance in the fashion clothing industry, we develop a theoretical framework that draws on social influence theory to consider how online brand communities link to the emergence of different customer loyalty commitments. Building on this premise and focusing on online brand communities, we refine and extend existing work by providing a framework that explains how consumer participation in online brand communities provide different streams of loyalty. The theoretical insights emerging from our study illuminate different customer commitments with important implications for our understanding of millennials' categorisation in online brand communities to inform novel action.

Case studies in failure data analysis

Antonio Pievatolo

CNR IMATI

e-mail address: antonio.pievatolo@cnr.it

Keywords: Field failures; forecast; stochastic process

ABSTRACT – Field failure data of manufactured items provide information on future failures and to the quality department as a means of identifying problems in product design, materials or the production line. Two case studies regarding home appliances and one on train doors are illustrated. Forecasts are obtained by stochastic process models, after estimating the unknown parameters with the available data. The method has the potential of being integrated with a failure monitoring system for a timely update of models and forecasts.

1. INTRODUCTION

Lawless [1] provides an early review on the analysis of warranty claim data of manufactured items, which are an example of field failure data. This type of data provides information on future failures and to the quality department as a means of identifying problems in product design, materials or the production line.

In this work we focus in particular on failure forecasting with the specific purpose of obtaining an early warning on incoming problems, when a retrospective analysis does not signal any malfunction, or of programming maintenance activities.

In this context, stochastic process models are very useful because they provide a probabilistic law for future events. We have considered a nonlinear state-space model for failures of home appliances (Pievatolo [2]) and a nonhomogeneous Poisson process model for failures of train doors (Pievatolo and Ruggeri) [3]. In Section 2 we illustrate the modelling and computational methodology and in Section 3 we present and discuss the results. In Section 4 we highlight the potential of the method for an integration with a real time failure monitoring system.

2. METHODOLOGY

An ensemble analysis of the failure of the door opening system of a fleet of 40 trains can be adequately modelled via a nonhomogeneous Poisson process (NHPP), because of the number of train involved and the system-level approach. In this case, observed events are the result of the superposition of several component-level failure processes and the system level process is well approximated by an NHPP (see Theorem 1 in Pievatolo and Ruggeri [3]).

With failure (calendar) times, also the distance run by each train at every failure is available. Although they are very much related to each other, they also contain different information (for example, the failure time includes the season), therefore we consider an NHPP on time and space, with a bivariate intensity function $\lambda_i(t, s)$ for train i, which has a form of a baseline modified to take into account the following: failures are concentrated along a space against time linear function and there is a yearly periodicity in the occurrence of failures, as shown in Fig. 1:

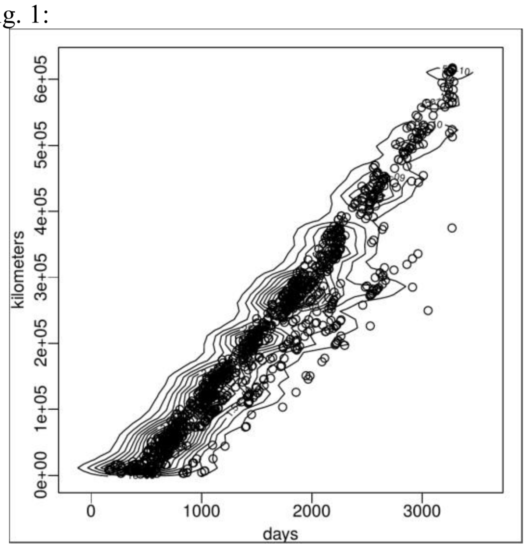

Figure 1: failure occurrences with contours from kernel smoother

After integrating over space is possible to derive the marginal intensity function $\lambda_i(t)$ in closed form and perform all the inference and forecasting on the time scale:

$$\lambda_i(t) = \mu \sqrt{\frac{\pi}{\gamma w(t-t_{0i})}}\, \Phi\left\{\left(a_i + c_i(t - t_{0i})\right) \times \sqrt{2\gamma w(t - t_{0i})}\right\} \exp\left\{A \cos\left(\omega(t - d)\right)\right\} \lambda_0(t - t_{0i})$$

For details on the parameters and their meaning see Pievatolo and Ruggeri [3]. Parameter estimates and forecasting have been obtained in a Bayesian framework by an MCMC Metropolis-Hastings algorithm, after assigning prior distributions to the unknown parameters.

For the failures of the home appliances we have segmented the items in the field (a very large number) by production plant. Failure data are collected as monthly number of failures for every monthly production batch (or cohort). Failure data show both batch-to-batch variation and within batch variation. The separate modelling of these two sources of variation, including stochastic dependence between batches and between failure counts within batches has led to the following state-space formulation:

$$\begin{aligned} y_{tj} &\sim Poisson(n_i e^{\alpha_{tj}}), \quad j-1,\dots,d \\ \alpha_t &= \alpha_{t-1} + \varepsilon_t \end{aligned}$$

where y_{tj} is the number of failures for batch t at the j-th month after sale, n_i is the number of items in the batch, $e^{-\alpha_{tj}}$ is the single item failure rate, α_t is the vector of all log failure rates and ε_t is a Gaussian random vector with zero mean and a covariance matrix Σ incorporating the within-batch correlation. Conditional on α_{t-1}, this model corresponds with the multivariate Poisson-lognormal model of Aitchison and Ho [4], which has an overdispersion property with respect to the Poisson distribution, because $Var(Y_{tj}) > E(Y_{tj})$. A Wishart prior distribution on Σ makes it possible to incorporate information on correlation in the procedure and start a particle learning algorithm (see Carvalho et al. [5] and Pitt and Shephard [6]) to obtain filtered failure rates and failure forecasts.

3. RESULTS

In Fig. 2 the plot of the cumulative number of failures over 8 years and 9 months for the door opening system of the train fleet is shown, along with an estimate from the NHPP model, with a baseline cumulative intensity function $\Lambda_0(t) = (1 - e^{-bt})/b$. One-year-ahead failure forecasts always include the observed number of failures when historical data amount to at least three years.

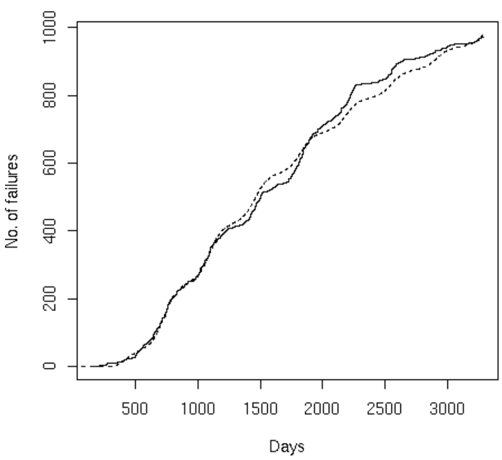

Figure 2: cumulative no. of failures of a train fleet vs time

In Fig. 3 we have reproduced the forecast of the failure rate (expected number of failures in 24 after-sales months divided by batch size) versus batch production times. As time gets closer to the time of the analysis we have less and less information on the number of failures; the prediction works satisfactorily until there are very little observed failure data.

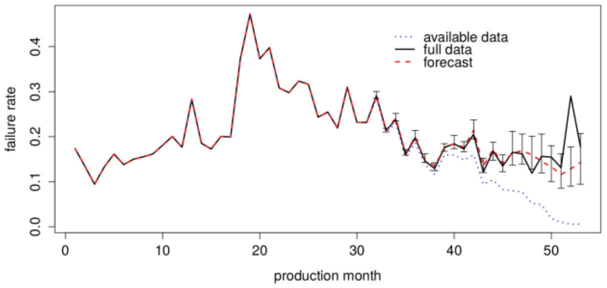

Figure 3: post-sales failure rate forecast of home appliances versus batch up to 24 months ahead with diminishing information.

4. CONCLUDING REMARKS

The failure forecasting methods that we have discussed involve stochastic process models with a sufficient degree of generality and flexibility. They have been thought and developed for application in a context where real-time reporting of failures was not implemented, but, once set up, they can be continuously updated with incoming data and integrated within a predictive analytics system.

5. REFERENCES

[1] J.F. Lawless, "Statistical analysis of product warranty data", *Int. Stat. Rev.*, vol. 66, pp. 41-60, 1998.

[2] A. Pievatolo, "A nonlinear state-space model for the forecasting of field failures", in *Book of short papers SIS 2018*, 2018, pp. 624-631.

[3] A. Pievatolo, F. Ruggeri, "Bayesian modelling of train door reliability", in *The Oxford Handbook of Applied Bayesian Statistics,* Oxford University Press, 2010.

[4] J. Aitchison, C.H. Ho, "The multivariate Poisson-log normal distribution", *Biometrika*, vol. 76, pp. 643-653, 1989.

[5] C.M. Carvalho, M.S. Johannes, H.F. Lopes, N.G. Polson, "Particle learning and smoothing", *Statistical Sciences*, vol. 25, pp. 88-106, 2010.

[6] M.K. Pitt, N. Shephard, "Filtering via simulation: Auxiliary particle filters", *Journal of the American Statistical Association*, vol. 94, pp 590-599, 1999.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Bayesian analysis of mediation in Cell Transformation Assays for testing the carcinogenicity of chemicals

Federico M. Stefanini

Department of Statistics, Computer Science, Applications, University of Florence, Italy.

*Corresponding author - email address: federico.stefanini@unifi.it

Keywords: Structural Causal Model; CTA

ABSTRACT – Cell Transformation Assays (CTAs) are in-vitro tests of carcinogenicity where a minimum of 10 Petri dishes are treated with the same dose of a substance and 5 or more doses are considered in the same experiment. The number of fully transformed *foci* grown on each dish is the outcome collected after 5 weeks from treatment [1]. A Bayesian structural causal model for CTAs is proposed and exploited to estimate mediated effects of a known carcinogen on the outcome. Null estimates of direct effects at several doses suggest that some features of CTAs might be reconsidered.

1. INTRODUCTION

Fast and cheap assays for probing the carcinogenic potential of substances to be used by cosmetic, food and pharma companies are of wide interest to perform preliminary screenings without resorting to rodent-based long and expensive studies. A typical CTA corresponds to a one-way randomized experiment, nevertheless recommended statistical models from the literature [2] do not seem to cover the whole range of distributions that might be required by very different chemicals, e.g. those with a variance smaller than the mean.

A structural causal analysis [3,4] based on the Directed Acyclic Graph (DAG) shown in Figure 1 is here proposed to describe several causal relationships existing in this class of experiments. Node D refers to the dose level applied to a Petri dish in which a random number of cells will be alive after treatment (node A), and that will develop a total number of *foci* (node T) after 5 weeks, some of them being fully transformed (node F). When the same experiment is replicated using different batches of cells (node C) and of substratum-reagents (node S) then confounding of A→T and T→ F relationships may occur. In order to illustrate the approach, a case study on a known carcinogenic chemical with 9 increasing dose levels is considered. The whole experiment was not replicated, thus estimates are conditional to a specific batch.

2. METHODOLOGY

Let $M_{F:g,i,j}$ be the random variable "manifest number of fully transformed *foci*" at dose level i in Petri dish j with batch of reagents g. The random variable $M_{A:g,i,j}$ is the "manifest number of alive cells after treatment", while $M_{T:g,i,j}$ is the "manifest total number of *foci*". Non-parametric equations relating endogenous variables are derived according to the DAG (Figure 1):

$$m_{F:g,i,j} = h_F(m_{D:i}, m_{T:g,i,j}, m_{S:g}, m_{C:g}, u_F) \quad (1)$$

Dose levels are assigned at random, therefore two more similar equations for variables $M_{A:g,i,j}$ and $M_{T:g,i,j}$ are enough to complete the structural part of the specification. Furthermore, if batch does not change within the same experiment then the specialized functions required in our case study are:

$$m_{F:i,j} = h'_F(m_{D:i}, m_{T:i,j}, u_F) \quad (2)$$

$$m_{T:i,j} = h'_T(m_{D:i}, m_{A:i,j}, u_T) \quad (3)$$

$$m_{A:i,j} = h'_A(m_{D:i}, u_A) \quad (4)$$

We assume that exogenous variables are independent, thus the induced manifest random variables $M_{A:g,i,j}$, $M_{T:g,i,j}$ and $M_{F:g,i,j}$ may be considered without making functions in equations (1-4) explicit; modularity of relationships coded by the DAG is also assumed.

In a typical CTA, the sample size n is small, for example $n = 89$ in our case study. A parametric Bayesian model was elicited by exploiting constraints coded in the CTA protocol, e.g. that the lowest dose is close to the No-Observed-Adverse-Effect-Level (NOAEL), while the highest dose is close to the median lethal dose (DL50):

$$M_{A:i,j} \sim c_{ini} \cdot \text{Beta}(\phi_i \mid \alpha_{A:i}, \beta_{A:i}) \quad (5)$$

with c_{ini} the initial total number of seeded cells, ϕ the survival rate, and i=1,2,…,9. Some of the alive cells will later originate transformed foci according to rate η:

$$M_{T:i,j} \sim \text{Poisson}(\eta_i \cdot m_{A:i,j}) \quad (6)$$

$$\eta_i \sim \text{Gamma}(\alpha_{T:i}, \beta_{T:i}) \quad (7)$$

Finally, the number of fully transformed *foci* are present as a proportion π of the number of total *foci*, that is:

$$M_{F:i,j} \sim \text{Binomial}(\pi_i, m_{T:i,j}) \quad (8)$$

$$\pi_i \sim \text{Beta}(\alpha_{F:i}, \beta_{F:i}) \quad (9)$$

After eliciting the initial distributions of model parameters (results not shown), total effects (TE), natural direct and natural indirect effects (NDE, NIE) [3,4] were estimated by Markov Chain Monte Carlo imputation of counterfactuals using Bayesian predictive distributions. Sufficient conditions for identifying all effects [3] are satisfied within stratum g (one batch-experiment).

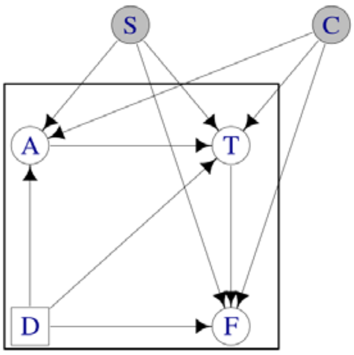

Figure 1 DAG with dish-specific variables inside plate

3. RESULTS AND DISCUSSION

Estimates of TE, NDE, NIE were obtained taking water as common reference value (Figure 2) and then taking dose level i as reference for dose level $i+1$, thus in this last case the effect captures changes due to an increase of dose and is useful to detect turning points (results not shown).

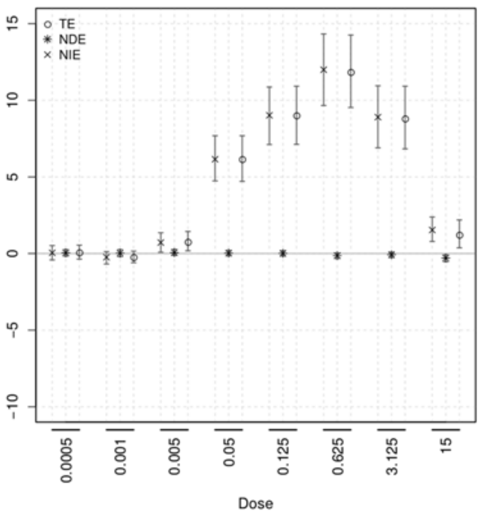

Figure 2 TE, NDE and NIE with water as reference

In Figure 2, Bayesian point estimates and 95% credibility intervals are shown: TE is almost entirely due to NIE at all dose levels but the highest. The number of fully transformed *foci* seems mainly determined by a change in the total number of *foci* instead of a change in the probability of transition to the fully transformed state. We conjecture that the role of M_T should be reconsidered, for example by modifying the CTA protocol to have high values of M_T even at low doses but with a small M_F.

At the highest dose, NDE is negative, a result that could be due to the increase of toxicity of the tested chemical. The relationship between log-dose and the expected value of parameter π in the posterior distribution was found almost constant for all but the 3 highest doses where a linear decrease was observed.

Model adequacy was checked by calculating values of discrepancy variables "number of distinct counts" and "range of the number of *foci*", both for M_T and M_F (not shown): 2 of the 9 dose levels might benefit from model improvement.

A large number of other different carcinogenic chemicals should be tested to assess whether the no-NDE pattern is a general feature in CTAs or it depends on the tested chemical. In all of these experiments the number of alive cells should be also measured, a practice not yet fully implemented in routine work.

4. CONCLUSIONS

Estimates obtained in our case study are conditional to a fixed batch g, but the natural extension to several batches is not trivial in practice. Current CTAs are replicated at most in 3 times, furthermore heterogeneity among cell batches might be due to dozens different genes and involve complex intercellular signaling. This is an important area to address in future research, for example by including molecular information and/or foci image descriptors [5] into the model.

5. REFERENCES

[1] OECD, "Detailed Review Paper on Cell Transformation Assays for detection of chemical carcinogens", *Series on testing and assessment*, Number 31, 2009.

[2] S. Hoffmann, L. A. Hothorn, L. Edler, A. Kleensang, M. Suzuki, P. Phrakonkham, D. Gerhard, "Two new approaches to improve the analysis of BALB/c 3T3 cell transformation assay data", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, vol. 744, no. 1, pp. 36-41, 2012.

[3] J. Pearl, "The deductive approach to causal inference", *Journal Causal Inference*, vol. 2, no. 2, pp. 115-129, 2014.

[4] J. Pearl, *Causality: models, reasoning, inference*, 2th ed. Cambridge: Cambridge University Press; 2009.

[5] G. Callegaro, K. Malkoc, R. Corvi, C. Urani, F.M. Stefanini, "A comprehensive statistical classifier of foci in the cell transformation assay for carcinogenicity testing.", *Toxicology in Vitro*, vol. 45, no. 3, pp.351-358, 2017.

6. ACKNOWLEDGEMENTS

The author acknowledges the financial support provided by the "Dipartimenti Eccellenti 2018-2022" ministerial funds. Thanks are due to Chiara Urani for sharing her view on CTAs besides the case study, and to Raffaella Corvi for criticism that helped to refine this work.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Gaussian Process Modeling with Boundary Information

Matthias Hwai Yong Tan[1*]

[1)] School of Data Science, City University of Hong Kong
[*]Corresponding author - e-mail address: matthtan@cityu.edu.hk

Keywords: Uncertainty quantification, Constrained emulator, Boundary conditions

ABSTRACT – In this talk, I will introduce the problem of constructing a Gaussian process (GP) emulator that reproduces known input-output relationships of a simulator at the input domain boundary. The proposed boundary modified GP emulator is shown to outperform the standard GP emulator with a real example.

1. INTRODUCTION

GP emulators are widely used to approximate time consuming deterministic simulators, which are often models of physical systems based on partial differential equations (PDEs). The PDE solution at the boundary of the input parameter space can be known based on physics or mathematical analysis. In particular, suppose that $y: [0,1]^2 \to \mathbb{R}$ denotes the *continuous* functional relationship represented by the simulator, where we assume that $[0,1]^2$ is the input domain for simplicity. In practice, there is often information given by one of more of the following expressions

$$y(0, x_2) = b_{01}(x_2) \forall x_2 \in [0,1], \quad (1)$$
$$y(1, x_2) = b_{11}(x_2) \forall x_2 \in [0,1], \quad (2)$$
$$y(x_1, 0) = b_{02}(x_1) \forall x_1 \in [0,1], \quad (3)$$
$$y(x_1, 1) = b_{12}(x_1) \forall x_1 \in [0,1]. \quad (4)$$

In other words, the simulator output at one or more of the edges of the input domain is known. When (1)-(4) is available, the goal of an emulator is to approximate the simulator well in the interior $(0,1)^2$ of the input domain.

Unfortunately, commonly employed stationary GP priors do not take the boundary information given by (1)-(4) into account. It is expected that if the GP prior is constrained by the known boundary behavior, it will give better prediction performance. In this talk, I will discuss how a GP emulator that reproduces known boundary behavior of the simulator can be built, and show that the proposed model gives improved prediction performance. This talk is based on the work of Tan (2018). Since the work of Tan (2018), there has been increasing interest in using boundary information in emulator construction. Ding et al. (2019) provide a rigorous theoretical analysis on the emulator convergence rate when the emulator satisfies known boundary information while Vernon et al. (2019) propose an alternative method.

2. METHODOLOGY

This section first provides a motivating example where boundary information is obtained through physical considerations. Then, the standard stationary GP emulator and the proposed emulator that exploits boundary information are described.

2.1 Motivating example

In this section, we consider a simulator implemented in Matlab PDE Toolbox (Mathworks, 2020) that predicts the temperature in a cylindrical rod of length $l = 3$m and diameter $d = 0.4$m (Figure 1). The cylindrical surface and right end of the rod are insulated. The initial (time 0) temperature of the entire rod is 27 Celsius. Immediately after time 0, the left end of the rod is uniformly raised to a temperature of x_2 Celsius. The rod has uniform thermal reffusivity x_1, where thermal reffusivity is the inverse of the thermal diffusivity, i.e., the volumetric specific heat divided by the thermal conductivity. The range of x_1 of interest is 0 to 30000 $\mathrm{s/m^2}$, while the range of x_2 of interest is 27 Celsius to 100 Celsius. The response of interest y is the temperature at the right end of the rod at time 86400 s = 24 hours.

Suppose we rescale x_1 (by dividing by 30000) so that its range is $[0,1]$ and we standardize x_2 so that its range is $[0,1]$. It is known that (1) holds with $b_{01}(x_2) = x_2$ since a rod with zero reffusivity will have infinite thermal conduction speed, making the temperature at the right end instantaneously equal to the left end. It is also known that (3) holds with $b_{02}(x_1) = 27$ since if the temperature at the left end of the rod becomes equal to the initial temperature of 27, the temperature at the right end will remain constant at the initial temperature also.

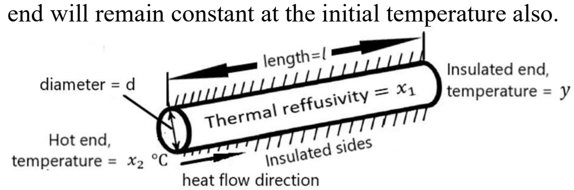

Figure 1: Schematic of rod temperature problem

2.2 Standard stationary GP emulator

The commonly employed GP emulator assumes a stationary real valued GP prior $\{Y(x): x \in [0,1]^2\}$ for y.

The prior mean and prior variance of Y are constants denoted by μ and σ^2 respectively. The prior correlation function, denoted by $\rho(x-x'|\theta)$, where $x,x' \in [0,1]^2$ and θ is a vector of correlation parameters, is chosen so that the GP is several times or infinitely mean square differentiable. A common choice of ρ is

$$\rho(x-x'|\theta)=\prod_{i=1}^{2}\exp\left(-\frac{|x_i-x_i'|}{\theta_i}\right)\left(\frac{|x_i-x_i'|}{\theta_i}+1\right), \quad (5)$$

which is a member of the Matérn correlation function family that yields a once mean square differentiable GP.

The prior GP is updated with the output data $\boldsymbol{y}=\left(y(x^1),\dots,y(x^n)\right)^T$ obtained by evaluating the simulator at the design points $\mathcal{D}=\{x^1,\dots,x^n\}$. This gives the posterior GP with posterior mean and covariance functions that are easily obtained. The hyperparameters μ,σ^2, and θ are estimated by maximizing the likelihood, and these estimates are plugged into the formulas for the posterior mean and covariance functions.

2.3 Boundary modified GP emulator

Our proposed boundary modified GP (BMGP) emulator employs a nonstationary GP prior $\{\mathcal{Y}(x): x \in [0,1]^2\}$ that satisfies the boundary constraints in mean square sense. Suppose that (1) and (3) are known to hold. Then, we choose the prior mean function $m(x)$ to be a *continuously differentiable* function such that

$$m(0,x_2)=b_{01}(x_2)\forall x_2 \in [0,1], \quad (6)$$
$$m(x_1,0)=b_{02}(x_1)\forall x_1 \in [0,1]. \quad (7)$$

We choose the variance function $v(x)$ to be a *continuously differentiable* function such that

$$v(0,x_2)=0\forall x_2 \in [0,1], \quad (8)$$
$$v(x_1,0)=0\forall x_1 \in [0,1], \quad (9)$$

and $v(x)>0\forall x \in (0,1]^2$. The correlation function for our BMGP model is (5). Given these choices, we see that

$$\lim_{x_1\to 0} E\{[\mathcal{Y}(x_1,x_2)-b_{01}(x_2)]^2\}=0\forall x_2 \in [0,1], \quad (10)$$
$$\lim_{x_2\to 0} E\{[\mathcal{Y}(x_1,x_2)-b_{02}(x_1)]^2\}=0\forall x_1 \in [0,1], \quad (11)$$

which implies that the BMGP converges in mean square to the boundary functions $b_{01}(x_2)$ and $b_{02}(x_1)$. Moreover, the BMGP model is mean square differentiable as the correlation function (5) and variance function $v(x)$ together give a covariance function $\rho(x-x'|\theta)\sqrt{v(x)}\sqrt{v(x')}$ with continuous second order mixed partial derivatives with respect to x_i and x_i'.

Our proposed form for the mean function is

$$m(x)=\lambda_0(x)\mu+\lambda_1(x)b_{01}(x_2)+\lambda_2(x)b_{02}(x_1), \quad (12)$$
$$\lambda_0(x)=\left(\sum_{i=1}^{2}d^2(x_i)\right)\Big/\left[\sum_{i=1}^{2}d^2(x_i)+\sum_{i=1}^{2}\frac{\alpha}{d^2(x_i)}\right],$$
$$\lambda_i(x)=\frac{\alpha}{d^2(x_i)}\Big/\left[\sum_{i=1}^{2}d^2(x_i)+\sum_{i=1}^{2}\frac{\alpha}{d^2(x_i)}\right], i=1,2,$$
$$d^2(x_i)=|0.5/(0.5+x_i)-1|^2,$$

where $d(\cdot)$ is a distance metric, and μ and α are parameters. Our proposed variance function is

$$v(x)=s^2\prod_{i=1}^{2}d^{2\eta}(x_i), \quad (13)$$

where s^2 and η are parameters. The mean and variance functions (12)-(13) given here are special cases of more general forms given in Tan (2018). We estimate the parameters μ,α,s^2,η, and θ with the maximum likelihood method.

3. RESULTS AND DISCUSSION

We generate a sliced Latin hypercube design with two slices for two inputs, where each slice is a Latin hypercube design with 20 points, using the R package SLHD (Ba et al., 2015). The first slice is used as the design to build the standard GP emulator and BMGP emulator for the rod temperature simulator described in Section 2.1. The second slice, which has an empty intersection with the first slice, is the test input set. The simulator is run at all points in the design and test input set, and the predictions at the test input sites given by the standard GP and BMGP emulators are compared. Table 1 gives the mean absolute error (MAE), root mean square error (RMSE), average length of 98% prediction intervals (ALPI), and coverage (Cov) of 98% prediction intervals for the two GP emulators. We see that the MAE, RMSE, and ALPI for the BMGP emulator are much smaller than that for the standard GP emulator, and the coverage of the BMGP emulator is slightly closer to nominal than the standard GP emulator.

Table 1 Mean absolute error (MAE), root mean square error (RMSE), and average length (ALPI) and coverage (Cov) of 98% prediction intervals for standard GP and BMGP emulators

Emulator	MAE	RMSE	ALPI	Cov
Standard GP	0.529	0.978	2.687	0.95
BMGP	0.255	0.324	1.772	1

4. CONCLUSIONS

In this talk, we consider the problem of using known simulator input-output relationships at the boundary of the input domain to improve GP emulator construction. These relationships can often be discovered by physical considerations or mathematical analysis, as the boundaries of the input domain frequently correspond to simplified physical processes. We propose to use judiciously chosen nonstationary mean and variance functions to build a GP emulator so that the emulator converges to the known boundary functions in mean square sense. It is demonstrated numerically via a real example that the proposed GP emulator, called the BMGP emulator, has substantially better prediction performance than the standard stationary GP emulator.

Acknowledgement This talk and research are supported by General Research Fund Project 11201117 sponsored by the Research Grants Council of Hong Kong.

5. REFERENCES

[1] Ba, S., Myers, W.R., & Brenneman, W.A. (2015). “Optimal sliced Latin hypercube designs,” *Technometrics*, 57(4), 479-487.

[2] Ding, L., Mak, S., & Wu, C. F. (2019). “BdryGP: a new Gaussian process model for incorporating boundary information,” *arXiv preprint arXiv:1908.08868*.

[3] MathWorks, (2020). Partial Differential Equation Toolbox: User's Guide (R2019b). Accessed at https://www.mathworks.com/help/pdf_doc/pde/pde.pdf on January 26, 2020.

[4] Tan, M.H.Y. (2018). “Gaussian process modeling with boundary information,” *Statistica Sinica*, 28(2), 621-648.

[5] Vernon, I., Jackson, S. E., & Cumming, J. A. (2019). “Known boundary emulation of complex computer models,” *SIAM/ASA Journal on Uncertainty Quantification*, 7(3), 838-876.

Adaptive Bayesian Prediction of Reliability Based on Degradation Process

Yubin Tian
Dianpeng Wang*

School of Mathematics and Statistics
Beijing Institute of Technology

*Corresponding author - e-mail address: wdp@bit.edu.cn

Keywords: accelerated life tests; Arrhenius model; change-point analysis

ABSTRACT – For long-time running electric devices used in satellites, the accurate reliability prediction is crucial in engineering. The reliability of these devices is often directly related to the degradation of a performance characteristic. In this paper, an adaptive Bayesian method with a new multivariate Normal-Gamma (abbreviated as MNG) prior is proposed based on an additive degradation process to predict the reliability efficiently. The prediction process mainly consists of two phases: 1) a Bayesian conditional c-optimal criterion is suggested to select observations from real-time data flow effectively; 2) a criterion is given based on the Bayesian conditional c-optimal criterion and the MNG conjugate prior to choose the subset from the observations which will make the prediction more robust.

1. INTRODUCTION

In electric industry, especially in satellite applications, many electric devices, such as Metal-Oxide-Semiconductor Field-Effect Transistor (MOS), have high reliability and need to run a long time such as 15 years. Accurate reliability prediction of these devices has attracted more and more attention in field applications. For electric devices with high reliability in satellites, engineers often use degradation data on the ground to establish a degradation model and then predict the reliability of the device for 16~20 years. In order to integrate the information of the ground data and the real-time data in orbit, Wang and Tian [1] proposed a local c-optimal method to choose effective data from real-time data flow and used the recursive maximum likelihood method to predict the reliability, which outperformed the traditional method. But the local c-optimal method doesn't take prior information about the parameters into account.

In this paper, a Bayesian strategy which takes advantage of the prior information about the parameters is proposed to improve the prediction. Assume $s(t_1), s(t_2), \cdots, s(t_m)$ are observed data and $s(t)$ is the observation at the future time t, where $s(t) = \big(n(t), z(t)\big)$, $n(t)$ denotes the shocks received before time t and $z(t)$ denotes the degradation of characteristic of devices (more details will be given in Section 2). The originality and novelty of the proposed Bayesian conditional c-optimal criterion is that it need to maximize the variance of the posterior mean of the reliability with respect to the marginal distribution of $s(t)$ based on $s(t_1), \cdots, s(t_m)$, which doesn't depend on the unknown parameters.

2. METHODOLOGY

Let T denote the lifetime of the device, then the reliability function is defined as:

$$R(t) = P\{T > t\}.$$

In this paper, MOS devices are considered as an illustration example. Wang and Tian [1] suggested an additive degradation process, which combined Poisson and Gaussian processes. The lifetime of MOS devices is defined as

$$T = inf\{t: r(t)/r(0) > l, t > 0\},$$

where $r(t)$ is the degradation value of the Rdson, which is a characteristic of MOS devices, $r(0)$ is the initial value and l is the failure threshold for the ratio. The logarithm of Rdson can be simply modeled using an additive process as,

$$x(t) = lnr(t) = \alpha t + \beta n(t) + x(0) + \epsilon,$$

where $n(t)$ denotes the shocks received before time t which follows the Poisson process with parameter τt, the error ϵ follows a normal distribution with mean 0 and variance σ^2 and it is assumed to be independent of $n(t)$. Similar to Wang [2], under the monotonic assumption of the degradation paths, we can obtain that

$$P\{T > t\} = P\left\{\frac{r(t)}{r(0)} < l\right\}.$$

If we know that the device is in operation at time t, the reliability at time y can be derived as

$$R(y;t|\theta) = \frac{\sum_{k=0}^{\infty}\Phi(\frac{\ln l-\alpha y-\beta k}{\sigma})\frac{(\tau y)^k e^{-\tau y}}{k!}}{\sum_{k=0}^{\infty}\Phi(\frac{\ln l-\alpha t-\beta k}{\sigma})\frac{(\tau t)^k e^{-\tau t}}{k!}}.$$

Let $z(t) = x(t) - x(0)$ and $s(t) = (n(t), z(t))$ be the observation at time t. Assume $\{s(t_1), \cdots, s(t_m)\}$ are the observations obtained so far. Then, the posterior of θ based on the current observations can be given as

$$\pi(\theta|s(t_1), \cdots, s(t_m)) \propto$$

$$\pi(\theta)\prod_{i=1}^{m}\frac{1}{\sqrt{2\pi\gamma^{-1}}}\exp\{-\frac{(z(t_i)-\alpha t_i-\beta n(t_i))^2}{2\gamma^{-1}}\}\frac{(\tau t_i)^{n(t_i)}e^{-\tau t_i}}{[n(t_i)]!},$$

where $\pi(\theta)$ is the prior of θ. The reliability at time y based on the observations, which is called the conditional reliability, can be estimated as

$$\hat{R}(y;t_m) = E_\theta R(y;t_m|\theta)$$
$$= \int R(y;t_m|\theta)\pi(\theta|s(t_1), \cdots, s(t_m))d\theta.$$

In this paper, a two-phase procedure is proposed to obtain the prediction of the reliability. In practice, only limited observations in orbit can be used to predict the reliability due to the high cost of computation, data storage and transmission. Thus, a Bayesian conditional c-optimal criterion is used in phase 1 to select observations from a real-time data flow. Then, the next observing time t_{m+1} is selected by minimizing the variation of reliability prediction based on the collected observations and a new observation at time t, i.e.,

$$t_{m+1} = argmin_t\, Var\{\hat{R}(y;t|s(t_1), \cdots, s(t_m), s(t))\},$$

where the variation is taken with respect to the conditional distribution of the observation at the new time t. As the observing time gets longer, the earlier observations may not fit the current degradation model well. In order to remove this fluctuation, a robust criterion is used in phase 2 to choose the subset from the observations, which is used to get the robust prediction. Let $\{s(t_1), \cdots, s(t_N)\}$ be the current observations selected from real-time data flow and let $\hat{R}_j(y;t_N|s(t_j), \cdots, s(t_N))$ be the reliability estimation based on the selected observations $\{s(t_j), \cdots, s(t_N)\}$. Note that here $N = m + 1$. Choosing j^* which makes

$$|\hat{R}_j\left(y;t_N\middle|s(t_j), \cdots, s(t_N)\right)$$
$$- \hat{R}_{j-1}(y;t_N|s(t_{j-1}), \cdots, s(t_N))|$$

largest from $j = 2$ to $N - 3$. For the selection of $\pi(\theta)$, it is common sense to choose the conjugate prior since it is an algebraic convenience to give a closed form expression for the posterior and can update the prior transparently using likelihood function. Thus in this paper, a multivariate Normal-Gamma prior, dubbed as $MNG(\mu_0, \Lambda_0, a_0, b_0)$, is suggested for θ.

3. SUMMARY

The reliability prediction for long-time running electric devices has always been a big problem in engineering due to the high cost of the experiments and the low efficiency of the use of real-time data. In this paper, we propose an adaptive Bayesian method for the prediction of reliability by using real-time data flow. A new MNG distribution is given which is proved to be the conjugate prior for the parameters in our model. A Bayesian conditional c-optimal criterion and a robust criterion are given to select observations from real-time data flow and choose the robust subset.

MOS device is considered as an illustration example in this paper, which is the key component in solid-state power controller. The simulation studies based on three different models are also done to compare the performance of our proposed adaptive Bayesian method with that of local c-optimal method and the adaptive Bayesian method with Jeffreys's non-informative prior. The simulation results show that our method is better than the other two methods. Besides, in practice, engineers can get the means of α, β and σ^2 and the variances of α and β based on historical experience. In this case the prediction based on informative prior usually performs better.

In this paper, we use a discretization strategy to perform optimization which suffers from high computational cost. A key and remaining question is an effective and fast algorithm of minimization for searching the next observing time t_{m+1}. Thus, we think there is room for the adaptive Bayesian method to be further improved by using much better optimization algorithms or criteria.

4. ACKNOWLEDGEMENT

This research is supported by NSFC grant 11371054.

5. REFERENCES

[1] J. Wang and Y.B. Tian. "An adaptive reliability prediction method for the intelligent satellite power distribution system". *IEEE Access*, vol. 6, pp. 58719–58727, 2018.

[2] X. Wang. "Wiener processes with random effects for degradation data". *Journal of Multivariate Analysis*, vol. 101, pp. 340 – 351, 2010.

Quality Big Data

Fugee Tsung[1, *]

[1)] Department of Industrial Engineering and Decision Analytic,
Hong Kong University of Science & Technology

[*]Corresponding author - e-mail address: season@ust.hk

Keywords: accelerated life tests; Arrhenius model; change-point analysis

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.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Inference On Errors In Industrial Parts: Kriging And Variograms Versus Geometrical Product Specifications (GPS) Standard

Grazia Vicario

Department of Mathematical Sciences, Politecnico di Torino
C.so Duca degli Abruzzi 24, 10129 Torino (Italy)

*Corresponding author - e-mail address: grazia.vicario@polito.it

Keywords: Geometric Errors, Variogram, Anisotropy

ABSTRACT – This paper focuses on the inference on the errors in manufactured parts controlled by using a coordinate measuring machines (CMM), optical profilometers, coherence scanning interferometers. The parts usually show typical deterministic geometric deviation pattern due to the manufacturing processes used for their production under the same condition. These patterns are referred as manufacturing signatures. In a number of situation, the measurements may be also affected by systematic errors of the CMMs measurement process, e.g. caused by a bad part alignment during the measurement process.
The International Standards, developed and approved by the International Organization for Standardization (ISO) with the Committees, define the parameters characterizing a feature and supply methods and formula apt to computation dealing with this issue. They suggest the analysis with the autocorrelation function (ACF), and linked parameters to it, because ACF is vital for this purpose.
In the present paper, we consider a spatial dependence between the measured points and we suggest the use of the empirical variogram to identify patterns in the parts. A comparison between the two methodologies is based on real set of measured points.

1. INTRODUCTION

Manufactured parts are affected by form and size errors, assessed against geometrical and dimensional tolerances. These errors, that usually show typical patterns related to the manufacturing process, are limited by means of dimensional and geometrical tolerances (such as straightness, roundness, flatness, profile) that are used to control more precisely the shape and the form of the parts (see Fig. 1). As examples, straightness is defined as the minimum distance between two parallel lines enclosing the actual feature (minimum zone criterion); similarly, the flat tolerance is the zone between two parallel planes within which a surface must lie. These definition are common to main international standards: they are part of the scope of the International Organization for Standardization (ISO) Technical Committee 213 and of the American National Standards Institute (ANSI/ASME B46.1, 2010).
The ISO Technical Committee 213 works on standards concerning Geometrical Product Specification and Verification (GPS) in modern industry with the aim to enable an effective inspection regardless of the verification tool. In fact, the compliance assessment is built on measurement and unavoidably affected by several uncertainty contributions; these are clearly identified and managed by GPS standards through the definition of specification and verification operators. In this paper, we refer to the ISO 25178-2:2012, due to the Technical Committee ISO/TC213. In particular we focus on its part where terms, definitions and parameters for determining the surface texture by areal methods are defined.
The comparison of the above mentioned methods is carried on with the analysis of the empirical variograms, whose use for identifying the correlation structure is recommended by the geostatisticians [2, 3]. In the last decades, the use of Kriging models (even if originated for geostatistics investigations) has acquired increasing popularity in dimensional metrology, because these models provide accurate predictions [4, 5, 6]. Since the fundamental of Kriging, a stochastic linear interpolation technique, is the rate at which the variance between points changes over space, this can be expressed as a variogram which shows how the average difference between values at points changes. The variogram is a function of the distance and of the corresponding direction of any pair of points, therefore it depicts their correlation extent.
The comparison is aimed to highlight not only their

differences but also their peculiarities in order to provide elements suitable for a rationale choice.

2. COMPARISON OF DIFFERENT METHODS IN WAVINESS RECONSTRUCTION

Most manufactured parts undergo some procedures of control of their surface characteristics, because their surfaces can be regarded as non-ideal surface morphologies caused by manufacturing errors.
For a long time, the profile method was used for measuring and characterizing a surface (for a comprehensive discussion see [7]). Recently, it was considered more appropriate to analyze the areal surface topography; and a further reason is that the areal measurements have more statistically significant information (large sample size and a closer representation of the real surface).
We considered the field parameters of surface texture, since they allow the characterization of the features involved in the comparison. The autocorrelation function (ACF) is the most popular field parameter and it is:

$$f_{ACF}(t_x, t_y) = \frac{\iint z(x,y)z(x-t_x, y-t_y)\,dxdy}{\iint z^2(x,y)\,dxdy}$$

where the integral is over the definition area. It expresses the correlation of a part of surface with respect to the whole surface. The range of ACF is between -1 and +1 for each point on the surface (the value |ACF| = 1 means a perfect correlation and ACF = 0 means no correlation). The ACF is used to study periodicities on a surface, i.e. when a texture pattern is reproduced several times on the surface, or is used to assess the isotropy of a surface.
There are other important field parameters, related to the ACF:

- *Autocorrelation length* is the horizontal distance of the ACF (t_x, t_y) that has the fastest decay to a specific value s ($0 \le s \le 1$):

  $$S_{al} = min\sqrt{t_x^2 + t_y^2}$$

  and the *min* is for any $(t_x, t_y) \in R$, where $R = \{(t_x, t_y): f_{ACF}((t_x, t_y) \le s)\}$. The value of s is suggested to be 0.2 for relatively smooth surface (ISO 25178, part 3, 2012), although other values can be used. If the surface is anisotropic, S_{al} is in the direction perpendicular to the surface lay.

- *Texture Aspect Ratio*, S_{tr}, is one of the most important parameter since it characterizes the isotropy of the surface. Given that, the standards offer a graphical representation for computing S_{tr}. It is unit-less and its values is between 0 and 1: if S_{tr} is close to unity, then the surface is isotropic (i.e. it has the same properties regardless of the direction); if it is close to 0, then the surface is anisotropic (i.e. it has a dominant texture direction). In this case, the parameter S_{td} (Texture direction) will give the direction angle of the texture. It is assessed from the Fourier spectrum of the surface; the Fourier spectrum gives the energy content of each spatial frequency on the surface and is usually represented as a plot where amplitudes are coded with a color or grey level (see Fig. 2).

Another approach for checking form tolerances is provided by the Kriging models. These models have been introduced in geological applications to predict spatial noisy data; at the end of 80's, they acquired popularity in computer experiment as a tool to approximate the output of deterministic simulation. In recent times, a novel application of Kriging in metrology brought back these models to their origin. The literature concerning the Kriging models in industry, in industrial processes and industrial metrology is quite rich nowadays (see [8] as pioneer paper in industrial metrology and [6] as one of the most recent ones).
The general approach is to fit data to some underlying model, in order to obtain an artificial surface. Kriging models have been exploited in Metrology because of their acknowledged ability in providing high-quality predictions in space-like situations. The Krige's inspiring idea is that response values at spatially close sites are more alike than responses at distant ones. Consequently, the identification of the correlation structure is a crucial issue since the prediction improves refining the specification of the correlation. The correlation structure can be found either by estimating the parameters of the spatial correlation function (SCF) or by variogram. For a careful choice of the correlation function, we suggest to evaluate it by means of the variogram, which does not rely on a-priori beliefs:

$$2C(\mathbf{h}) = \frac{1}{\#N(\mathbf{h})}\sum_{N(\mathbf{h})}\left(Z(\mathbf{x}_i) - Z(\mathbf{x}_j)\right)^2$$

where $N(\mathbf{h}) = \{(\mathbf{x}_i, \mathbf{x}_j): \mathbf{x}_i - \mathbf{x}_j = \mathbf{h}; i, j = 1,2,...n\}$
and #N(**h**) is the number of pairs N(h) that are distinct.
The empirical variograms for the input variables (Cartesian coordinates) of real case studies brought to light unusual behaviours. This finding suggested to focus on the analysis of changes in the estimated variogram due to stimuli of a noticeable trend in the model. In [6] we simulated a few parts, affected by the most recurring errors. The empirical variograms have been evaluated: the omni-directional variogram for pinpointing possible correlation structure and the two one-dimensional ones for identifying origins and causes.

3. CONCLUSIONS AND FINAL REMARKS

The issue addressed in this paper is the form tolerances checking. The standards provide a number of indices in order to detect possible technological errors and signatures in the parts. We adopted the ordinary Kriging model, because they have been used in an industrial context proving their efficiency in predicting, and the variograms for modelling a possible covariance, as geostatisticians do for dealing with very noisy data. The variograms brought to light behaviors, that can be

considered as anomalous for classical covariance models but very indicative for the form errors.
The anisotropy in the empirical variograms along different directions (usually 0°, 45° and 90°) in presence of a noticeable trend in the model, is suggested for detecting a spatial correlation confounded with specifying trends.
To summarize, the Kriging models and the variograms in detecting the covariance structure can be considered as an encouraging preliminary research step to be used as a guide for further developments in detecting anomalies.

4. REFERENCES

[1] ANSI/ASME B46.1, Surface texture, surface roughness, waviness and lay. American National Standards Institute, Washington, 2010.
[2] ISO 25178-2, Geometrical product specifications (GPS) – Surface texture: Areal - Part 2: Terms, definitions and surfacee texture parameters. International Organization for Standardization, 2012.
[3] N.A. Cressie, "Spatial Prediction and ordinary Kriging", *Mathematical Geology*, 20(4), 407-421, 1997.
[4] [3] J. Haslett, "On sample variogram and the sample autocovariance for non-stationary time series", *The Statistician*, 46, 475-485, 1997.
[5] [4] S. Ruffa, G. Vicario, and G. Pistone, "Analysis of the Covariance Structure in Manufactured Parts", *Communications in Statistics—Theory and Methods*, 44(21), 4540-4551, 2015.
[6] [5] G. Pistone and G. Vicario, "A note on Semivariogram". In Topics on Methodological and Applied Statistical Inference (T. Di Battista, E. Moreno, W. Racugno Eds), Springer, 2016, pp. 181-190.
[7] [6] G. Vicario and G. Pistone, "Simulated variogram-based error inspection of manufactured parts", *Statistical Papers*, 59, 1411-1423. doi:10.1007/s00362-018-1030-0, 2018.
[8] [7] K.J. Hume, *A history of engineering metrology*, Mechanical Engineering Publications Ltd, London, 1980.
[9] [8] P. Pedone, G. Vicario, and D. Romano, "Kriging-Based Sequential Inspection Plans for Coordinate Measuring Machines", *Applied Stochastic Models in Business and Industry*, 25, 133-149, 2009.

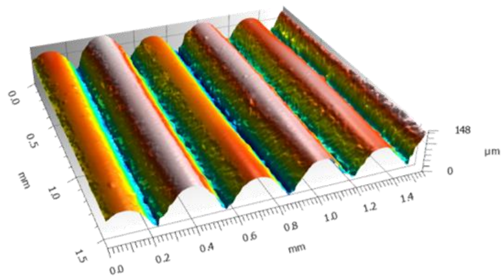

Figure 1 Example of a manufactured part affected by form error (waviness).

ISO 25178		
Height Parameters		
Sa	26.0	µm
Sq	31.0	µm
Sp	53.5	µm
Sv	94.0	µm
Sz	148	µm
Spatial Parameters		
Sal	0.0654	mm
Str	0.0838	
Std	0.315	°

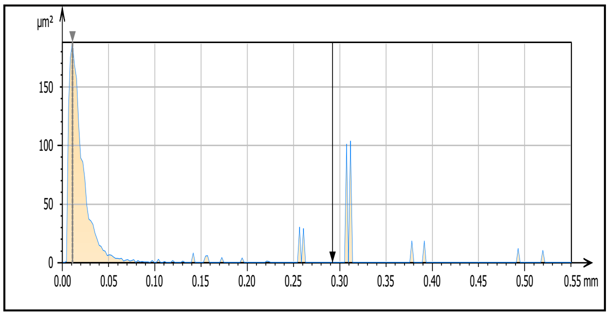

Figure 2 Output of the part in Fig. 1 characterization according to ISO 25178-2

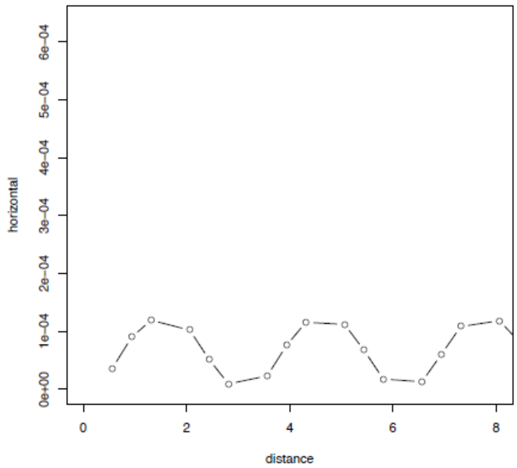

Figure 3 Empirical (horizontal) variogram related to the surface in Fig. 1. The waviness in the planar surface is reflected in the variogram.

Towards a Foundational Theory of Statistical Engineering

Geoff Vining*
Virginia Tech
USA

*Corresponding author - vining@vt.edu

Keywords: digital twin, first principles models, data science, analytics, and expert opinion

ABSTRACT – Statistical engineering is a new discipline dedicated to the solution of interesting problems/opportunities with data. The paradigm for this discipline is chemical engineering. Chemical engineering uses fundamental results from chemistry to solve important and often complex product/process problems. Chemical engineers can make fundamental contributions to the science of chemistry; however, chemical engineers primary focus is how to apply the fundamental science to these critical product/process opportunities. The proper application of the fundamental science requires far more tools than the field of chemistry provides. Ultimately, the theory of chemical engineering details the most efficient and effective approaches for applying the broad array of tools, including basic chemical theory, to solve these product/process problems. The theory of chemical engineering forms the basis for best practices, which guide the application of the appropriate tools to specific types of problems.

The foundational theory of statistical engineering distinguishes the new discipline from statistics, data science, predictive analytics, data mining, etc. These areas provide important tools that are critical for the solution of complex opportunities; however, ultimately, these areas are much more about the tools than the most efficient and effective approaches for their application. The foundational theory of statistical engineering defines the most efficient and effective approaches to combine the tools of statistics/data science with other critical toolsets to create effective solutions to important opportunities.

This talk introduces the foundational theory of statistical engineering, much of which is currently unfolding. As a result, the major purpose of this talk is to initiate a conversation on two major topics:

- The additional toolsets necessary to complement the tools and methods of statistics/data science
- Best practices for applying the statistical engineering toolsets to solve complex problems.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Assessment of Reliability in Accelerated Degradation Testing with Initial Status Incorporated

Chengjie Wang, Qingpei Hu*, and Dan Yu
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing

* qingpeihu@amss.ac.cn

Keywords: reliability, accelerated degradation testing, initial status

ABSTRACT – Accelerated Degradation Tests (ADT) provide effective information for reliability assessment of performance characteristic of long life and high-reliability products. Existing typical models and analysis usually assume that the products under test are of high consistency level during the manufacturing process, which implies that the individual differences of the initial performance of the products can be ignored. However, this may not be the case, and the initial performance of the test units may have great impact on the subsequent degradation rate. Both positively related and negatively related are possible. This phenomenon can be observed in many different examples, such as the performance of inkjet printer heads. It means that reliability-related information can be obtained before accelerated degradation test. The study considers the impact of initial performance on the reliability assessment. Based on the existing typical accelerated degradation test model and analysis process, this paper introduces the initial information of the products to carry out reliability assessment and test plan. The asymptotic variance of a lifetime quantile at normal use conditions is considered to obtain the optimum test plan. Results show that the initial performance of the test units can be made use of to improve the accuracy of estimators. The impact of fisher information has been taken into account.

1. INTRODUCTION

In many practical applications, the initial status of test units is not always consistent. There are differences among individuals, such as the examples of disk error units in [1]. The initial difference may further affect the subsequent degradation rate. Lu et al. assume that the initial state and degradation rate follow the multivariate normal distribution [2], which is also the model to be used in this paper. Ye et al. make a theoretical study of the initial degradation and carry out numerical simulation [3]. In this study, the factors mentioned above are taken into account in the modelling and analysis of accelerated degradation test, and the effects on reliability evaluation accuracy and test scheme design are studied.

2. MODELLING AND INFERENCE

For sample i, we record the stress level as s_i, the measurement time point $\boldsymbol{t}_i = \left(t_{i1}, \cdots, t_{im_i}\right)'$, where m_i is the number of measurements besides the initial state of sample i, $X_{(i)}$ is the initial state of sample i and the corresponding measurement data as $\boldsymbol{Y}_i = \left(Y_{i1}, \cdots, Y_{im_i}\right)'$. Thus, all the data collected are $D = \left\{\left(X_{(i)}, s_i, \boldsymbol{t}_i, \boldsymbol{Y}_i\right) \middle| i = 1,2, \cdots, n\right\}$. A random effect model is used in this paper to analyse the degradation data. This model was firstly introduced by J.-C. Lu *et al.* (1997). For test units with stress level s, the degradation level at time t is measured

$$Y(t,s) = D(t,s) + \varepsilon = b_0 + b_1 \exp(\gamma s)\, t + \varepsilon \qquad (1)$$

where b_0 is the initial performance, b_1 is the slope, indicating the degradation rate in normal use condition, γ is the model parameter which is used to fit the effect of stress level on the degradation rate, and $\varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ is the measurement error. In order to consider the initial quality difference among test units and its effect on degradation rate, it is assumed that (b_0, b_1) follows normal distribution

$$\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}\right)$$

the correlation coefficient $\rho = \sigma_{01}/\sigma_0\sigma_1$ may be positive or negative.

In order to simplify the calculation process, we can always conduct a monotone transformation on the stress level. Assume s_0 to be the normal use level of the test unit and s_H be the maximum stress level applicable to model (1). The following transformation is used to normalize the stress level

$$s = \frac{\psi(\tilde{s}) - \psi(\tilde{s}_0)}{\psi(\tilde{s}_H) - \psi(\tilde{s}_0)}$$

where $\psi(\cdot)$ is a monotone function. Particularly, $\psi(\tilde{s}) = 1/\tilde{s}$ while in the Arrhenius model, $\psi(\tilde{s}) = \ln \tilde{s}$ in the power law model. After the normalization, all the stress levels involved in the tests are compressed into the interval [0,1], and the normal use level becomes 0.

We use the maximum likelihood estimation method to estimate the parameters of the model. In the process, we derive the Fisher information matrix for subsequent estimation of variance. The degradation level of the initial measurement is $X = b_0 + \varepsilon \sim \mathrm{N}(\alpha_0, \sigma_0^2 + \sigma_\varepsilon^2)$. All the measurements $(X_i, \boldsymbol{Y}_i)'$ of sample i also follow the multivariate normal distribution.

$$\begin{pmatrix} X_i \\ \boldsymbol{Y}_i \end{pmatrix} \sim \mathrm{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$$

where $\boldsymbol{\mu}_i = \boldsymbol{Z}_i \boldsymbol{z}_i \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}$, $\boldsymbol{\Sigma}_i = \boldsymbol{Z}_i \boldsymbol{z}_i \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \boldsymbol{z}_i \boldsymbol{Z}_i' + \sigma_\varepsilon^2 I_{m_i}$, and

$$\boldsymbol{Z}_i = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & t_{i1} & \cdots & t_{im_i} \end{pmatrix}', \boldsymbol{z}_i = \begin{pmatrix} 1 & 0 \\ 0 & \exp(\gamma s(i)) \end{pmatrix} \quad (2)$$

Let $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \gamma, \sigma_0^2, \sigma_1^2, \sigma_\varepsilon^2, \sigma_{01})'$ be all the parameters involved in the model. We only need to solve the likelihood equation to get the estimator $\widehat{\boldsymbol{\theta}}$. The logarithm of likelihood function is

$$l(\boldsymbol{\theta}|\boldsymbol{D}) = -\frac{1}{2}\sum_{i=1}^{n} \ln \det(\boldsymbol{\Sigma}_i) - \frac{1}{2}\sum_{i=1}^{n}\left(\begin{pmatrix} X_i \\ \boldsymbol{Y}_i \end{pmatrix} - \boldsymbol{\mu}_i\right)' \boldsymbol{\Sigma}_i^{-1} \left(\begin{pmatrix} X_i \\ \boldsymbol{Y}_i \end{pmatrix} - \boldsymbol{\mu}_i\right) \quad (3)$$

This likelihood equation contains a non-linear exponential function. The analytic expression of parameter estimation $\widehat{\boldsymbol{\theta}}$ cannot be obtained by solving the equation. The likelihood equation can be solved by numerical method. The Fisher information matrix is $I(\boldsymbol{\theta}) = E\{-\partial^2 l(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\}$. The solving process is a bit complicated. Formula (16) in Klein *et al.* (2000) gives the result directly. $I(\boldsymbol{\theta}) = \tilde{I} + I$, $\tilde{I}$ is an additional positive semidefinite matrix which can be seen as the additional information from the initial degradation level of the test units

$$\tilde{I} = \begin{pmatrix} \frac{n}{\sigma_0^2+\sigma_\varepsilon^2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{n}{2(\sigma_0^2+\sigma_\varepsilon^2)^2} & 0 & \frac{n}{2(\sigma_0^2+\sigma_\varepsilon^2)^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{n}{2(\sigma_0^2+\sigma_\varepsilon^2)^2} & 0 & \frac{n}{2(\sigma_0^2+\sigma_\varepsilon^2)^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (4)$$

while I represents the information provided by the degradation process. The (k, l)-th entry of the matrix I is

$$I_{\boldsymbol{\theta}}(k,l) = \frac{1}{2}\sum_{i=1}^{n} tr\left(\boldsymbol{\Sigma}_i^{-1}\frac{\partial \boldsymbol{\Sigma}_i}{\partial \boldsymbol{\theta}_k}\boldsymbol{\Sigma}_i^{-1}\frac{\partial \boldsymbol{\Sigma}_i}{\partial \boldsymbol{\theta}_l}\right) + \sum_{i=1}^{n} E\left(\frac{\partial \boldsymbol{\mu}_i'}{\partial \boldsymbol{\theta}_k}\boldsymbol{\Sigma}_i^{-1}\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\theta}_l}\right) \quad (5)$$

where $\partial/\partial\, \boldsymbol{\theta}_l$ denotes the partial derivative of the l-th parameter. For any given function $g(\boldsymbol{\theta})$, the MLE of $g(\boldsymbol{\theta})$ is $g(\widehat{\boldsymbol{\theta}}_{ML})$ by the equivalence property of maximum likelihood estimation. Then, according to the asymptotic normality of MLE and Delta method, the asymptotic variance of $g(\widehat{\boldsymbol{\theta}}_{ML})$ is obtained

$$AVar\left(g(\widehat{\boldsymbol{\theta}}_{ML})\right) = \nabla g(\widehat{\boldsymbol{\theta}}_{ML})' AVar(\widehat{\boldsymbol{\theta}}) \nabla g(\widehat{\boldsymbol{\theta}}_{ML}) = \nabla g(\widehat{\boldsymbol{\theta}}_{ML})' [I(\boldsymbol{\theta})]^{-1} \nabla g(\widehat{\boldsymbol{\theta}}_{ML}) \quad (6)$$

where $\nabla g(\cdot)$ denotes the partial derivative of function $g(\cdot)$ with respect to parameter $\boldsymbol{\theta}$.

3. SUMMARY

The influence of initial state of product on assessment of reliability and test plan considered. The general results are obtained through theoretical derivation. In the assessment of reliability, the Fisher information matrix is larger when considering the initial state than when not considering it.

4. REFERENCES

[1] W. Q. Meeker, & L. A. Escobar, *Statistical methods for reliability data*: John Wiley & Sons, 1998.

[2] J.-C. Lu, J.Park, & Q. Yang, Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics*, 39(4), 391-400, 1997.

[3] Z.-S. Ye, Q.Hu, & D.Yu, Strategic allocation of test units in an accelerated degradation test plan. *Journal of Quality Technology*, 51(1), 64-80, 2019.

Reconstruction learning: From interpolation to regression and classification

Shifeng Xiong*

Chinese Academy of Sciences.

*Corresponding author - e-mail address: xiong@amss.ac.cn

Keywords: Gaussian process; kernel learning; SVM

ABSTRACT – This paper introduces an interpolation-based method, for nonparametric regression and classification. It is shown that it has broad applications with easy implementations.

1. INTRODUCTION

This paper introduces an interpolation-based method, called the reconstruction approach, for nonparametric regression and classification.

2. METHODOLOGY

Based on the fact that interpolation usually has negligible errors compared to statistical estimation, the reconstruction approach uses an interpolator to parameterize the unknown function with its values at finite knots, and then estimates these values by minimizing the regularized empirical risk. Some popular methods including kernel ridge regression and kernel support vector machine can be viewed 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. In particular, for some methods of complexity $O(n^3)$ where n is the sample size, this approach provides efficient surrogates with much less computational burden. This point makes it very suitable for large datasets. The details can be found in [1].

3. SUMMARY

In this paper we have proposed the reconstruction approach for nonparametric regression and classification. Theoretical basis behind the proposed approach is that interpolators of sufficiently smooth functions usually yield negligible errors compared to statistical errors. We have shown several features:

(a). This approach is easy to understand and to implement. The parameters in it have intuitive interpretation: they are function values at selected knots. This point is much different from other parameterization methods, and can bring some convenience in its implementation.

(b). It provides an interpolation angle to examine existing methods. Several popular methods can be viewed as its special cases. In particular, for some methods with n basis functions like GPR, it yields surrogates with much less basis functions. Therefore, the proposed approach is suitable for large scale problems.

(c). It allows flexible implementation with different sizes of knot sets. Our numerical experiments show that the proposed approach has satisfactory performance for both small/moderate and large scale problems.

(d). It systematically connects the two important areas: interpolation and regression. By the reconstruction idea, regression and classification problems can be handled based on an efficient interpolator. Other possible applications include density estimation, quantile regression, and semiparametric regression. The reconstruction approach broadens the applications of interpolation in statistics and machine learning.

4. REFERENCES

[1] S. Xiong, "The reconstruction approach: From interpolation to regression," *ArVix, 2019.*

StEering
Inter-University Research Center:
Design, Quality and Reliability

Beijing Enterprise Innovation and Development of Sophisticated Industry

Yanyun Zhao[1] Kang Wang[2]

[1] School of Statistics, Renmin University of China
[2] School of Economics, Beijing Technology and Business University

*Corresponding author - e-mail address: wk0536@126.com

Keywords: Enterprise Innovation; Development of Sophisticated Industry

ABSTRACT – Under the background of building the national science and technology innovation center, based on enterprise data, using statistical model, this paper analyzes the status, mechanism and countermeasures of Beijing enterprise innovation and sophisticated industry development. This paper conducts research according to the research idea of "innovation elements analysis - development path of sophisticated industry - sophisticated enterprise positioning - power source of sophisticated industry". The purpose of this paper is to provide empirical evidence for better improving the sophisticated industry pattern and entering the new stage of high-quality development.

1. INTRODUCTION

Under the goal for building the national science and technology innovation center, Beijing has continuously optimized the innovation ecology, improved the innovation level and forged a pattern of sophisticated industries in recent years. In this context, based on enterprise data, using statistical model, this paper analyzes the status, mechanism and countermeasures of Beijing enterprise innovation and sophisticated industry development, with the expectation to provide empirical evidence for better improving the sophisticated industry pattern and entering the new stage of high-quality development.

2. METHODOLOGY

This paper conducts research according to the research idea of "innovation elements analysis - development path of sophisticated industry - sophisticated enterprise positioning - power source of sophisticated industry". Firstly, this paper comprehensively depicts the innovation situation of Beijing enterprises from the two aspects, i.e., innovation input (divided into basic elements and environmental elements) and innovation output, and analyzes their advantages as well as weak links; secondly, this paper explores the joint action of basic elements and environmental elements to “the innovative breakthroughs” (potential variables) using the structural equation model so as to push forward the path of the forming of sophisticated industries; thirdly, this paper positions the sophisticated enterprises, that is, constructs the sophisticated standards system, identifies the sophisticated enterprises using the Weaver-Thomas Index, and compare them for level analysis; finally, this paper analyzes the power source of sophisticated industry development, explores the influencing factors of sophisticated industry development and analyzes the sophisticated endogenous dynamic action of the sophisticated talent elements so as to reveal the micro-mechanism for enterprises to promote the endogenous development path of sophisticated industries through innovation.

3. RESULTS AND DISCUSSION

The research findings of this paper are as follows: Firstly, in recent years, Beijing has maintained a strong growth trend in basic elements and environmental elements of innovation, thereof, the growth rate of basic elements is relatively smaller, which indicates that Beijing is entering a high-quality development stage of "rapid, stable and quality improvement"; secondly, under the joint action of basic elements and environmental elements, Beijing enterprises can achieve "innovative breakthroughs" and promote sophisticated industries while the action mechanism among certain elements inside the system needs to be further consummated; thirdly, senior talents and modern finance are important factors in the development of sophisticated industries, while the number of full-time R&D personnel and information technology can indirectly affect the forming of sophisticated industries with the help of matching effect.

4. CONCLUSIONS

This paper analyzes the status, mechanism and countermeasures of Beijing enterprise innovation and sophisticated industry development. The innovation vitality and innovation ability of enterprises in Beijing maintain a strong growth trend, and the sophisticated industrial pattern is gradually formed. However, there are some weak links in the development path and power source mechanism of sophisticated industries. Based on the research conclusions, this paper puts forward the basic ideas of further accelerating Beijing enterprise innovation and building a sophisticated economic structure.

5. REFERENCES

[1] Y.Y. Zhao, F. Zhou, and W.J. Fu, *Research on the theory, method and application of sophisticated industrial structure in Beijing.* Bei Jing: Capital Normal University Press; 2013.

[2] Y.L. Wu. *China's total innovation system theory and empirical analysis of innovation model.* Bei Jing: Renmin University of China; 2011.

[3] Z.Q. Li, “A Study of Breakthrough Innovation of Strategic Management: from the Perspective of Risk Investment” *Contemporary Finance & Economics,* pp. 61–71, 2012.

[4] D.T. Lei, “Industry evolution and competence development: the imperatives of technological convergence” *International Journal of Technology Management,* pp. 72–76, 2000.

[5] X.H. Huang and S.L. Zhang, “Industry evolution and competence development: the imperatives of technological convergence” *China Industrial Economics,* pp. 60–78, 2019.

Contributed posters

A Combined Approach To Detect Key Variables In Thick Data Analytics

Giovanni Antonelli[1], Rosa Arboretti Giancristofaro[2], Riccardo Ceccato[3], Paolo Centomo[1], Luca Pegoraro[3], Luigi Salmaso[3*] and Marco Zecca[1]
[1)] Department of Chemical Sciences, University of Padova, Padua, Italy.
[2)] Department of Civil, Environmental and Architectural Engineering, University of Padova, Padua, Italy.
[3)] Department of Management and Engineering, University of Padova, Vicenza, Italy.

[*]Corresponding author - e-mail address: luigi.salmaso@unipd.it
Keywords: Variable selection; permutation;

ABSTRACT – In machine learning one of the strategic tasks is the selection of only significant variables as predictors for the response(s). In this paper an approach is proposed which consists in the application of permutation tests on the candidate predictor variables in the aim of identifying only the most informative ones. Several industrial problems may benefit from such an approach, and an application in the field of chemical analysis is presented. A comparison is carried out between the approach proposed and Lasso, that is one of the most common alternatives for feature selection available in the literature.

1. INTRODUCTION

Feature selection is a critical step in data preparation before machine learning modelling. Several machine learning methods have already been developed which can perform feature selection such as Lasso or tree-based methods, namely decision trees and random forests. In this paper an alternative approach is proposed which is based on NonParametric Combination (NPC) procedures [1].

The relevance of feature selection is mainly dictated by the risk of overfitting and to achieve model simplification and diminish computational time. Furthermore, in industrial environments, feature selection may lead to additional savings such as cost savings in sensors for equipment monitoring.

Several fields are affected by the problem of high dimensionality of data, therefore methods for feature selection can find application in many diverse areas. Condition monitoring of industrial equipment is one of these areas: at the present, it is not unusual to measure tens or hundreds of parameters in a device or equipment, and the more complex is the system considered the higher is the number of sensors installed on it. The recent topics of Industry 4.0 and predictive maintenance are an additional boost for the increasing number of features measured [2]. Prediction of air pollution is another topic in which feature selection plays a relevant role. Datasets for air pollution forecasting tend to have a significant number of variables, not only carrying information on the atmospheric conditions but also including covariates such as seasonality, weekday and location [3]. The techniques employed in analytical chemistry for qualitative and quantitative assessments in several sectors including chemical, petrochemical, pharmaceutical, environmental, and agricultural readily produce datasets containing hundreds of data points, hence can be used for the prediction of a dependent response based on hundreds of variables [4]. In this context, an appropriate selection of the significant range for the collection of experimental data points leads to better accuracy and performance of the analytical method together with a simpler model to use and interpret.

The approach proposed of adopting NPC procedures for selection of informative variables is here validated by its application in a problem of analyte detection in a mixture.

2. METHODOLOGY

The problem on hand deals with the recognition of analyte concentration in a mixture. A Design of Experiments (DOE) study is planned to obtain 250 samples that are then split in 75% for training and 25% for model validation. Each of the samples is obtained as a mixture of 6 components that are dissolved in a solvent and each of the components is split into 3 levels of concentration. For each of the 250 combinations 5 repetitions of the analytical measure are carried out in such a way that 130 data points are collected per single repetition. This results in a multivariate regression problem in which 130 variables are available for the prediction of 6 response variables. The goal is to develop a model which can reliably detect the composition of the mixture. To this end, the range of the datapoint collection is selected by means of an NPC approach.

Variable selection is carried out on the training data. The average value for the 5 repeated measures is calculated for each one of the 130 variables of each sample, and a pairwise comparison is performed (using the difference in means as test statistic) for all the samples. A permutation approach is adopted in order to perform the testing procedure and p-values are calculated accordingly [1]. This results in 17391 pairwise comparisons, and the familywise error rate is controlled by means of the false discovery rate (FDR) correction [5]. The corrected p-values are evaluated at a significance level $\alpha=0.05$, and as such a certain number of significances are found for each data point. At this level, cutoffs must be found for the number of significances to determine how many variables to select. Some relevant cutoffs have been selected to the scope of this paper.

Lasso is used as a baseline for assessment of the NPC procedure proposed. Because of the multivariate nature of the problem on hand, a variable is included if it is selected by Lasso for at least one of the responses. In this setting, the multivariate Lasso selects 25 variables. After variable selection is executed, machine learning models are fit to the remaining variables of the training set and the predictive ability of the algorithms is assessed by application on the vaildation set. The validation set error is used as a means to evaluate the goodness of the subset of variables included by the different approaches.

3. RESULTS AND DISCUSSION

A comparison of the set of variables selected by the NPC approach and Lasso shows that the first tends to select bands of data points, while the latter tends to select data points that are rather sparse (Fig. 1). This behavior is imputable to the different rationale that drives the selection in the two cases. In presence of correlated predictors Lasso tends to pick one variable and discard the others. On the other hand, the proposed NPC approach picks all the most significant variables (according to the chosen cutoff) even if they are strongly correlated. Since in the present setting adjacent variables are more correlated than distant variables, it is no surprise that the two methods behave differently.

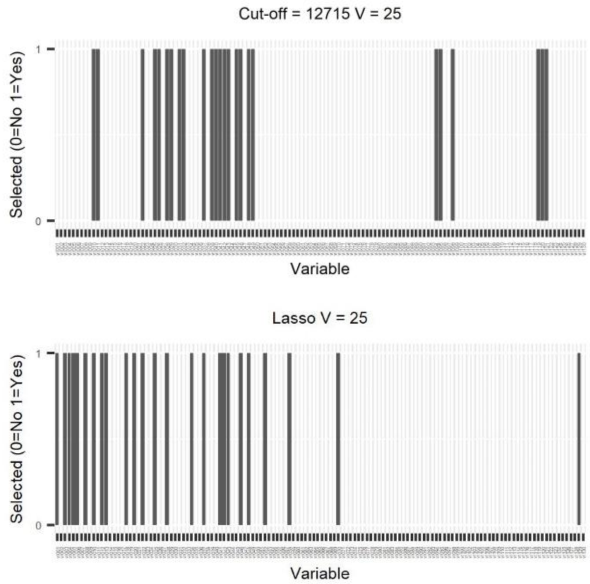

Figure 1 Variables selected by the NPC approach (top) and Lasso (bottom)

Ridge regression is selected as algorithm for machine learning modelling. The performance of the method is reported in Figure 2. The selection of only a subset of variables leads to a decrease of the predictive ability of the algorithms. The NPC approach outperforms Lasso when the number of variables selected is similar. The flexibility given by the choice of the cutoff in the NPC approach can further reduce the mean absolute error by including additional variables. Furthermore, the ability of the NPC approach to select bands of correlated variables may represent an advantage e.g. in the selection of a sensor for measuring such variables.

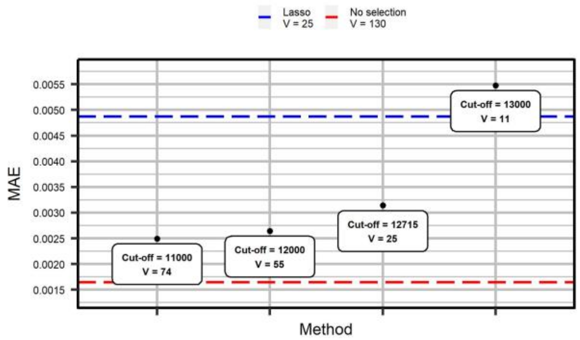

Figure 2 Validation set mean absolute error (MAE) for Ridge regression for different variable selection strategies

4. CONCLUSIONS

In this paper a novel strategy for feature selection based on permutation tests has been discussed and applied to a problem of chemical analysis. The performances of the approach proposed are compared to Lasso and used for model validation of a regression technique.

5. REFERENCES

[1] F. Pesarin, and L. Salmaso, *Permutation Tests for Complex Data: Theory, Applications and Software*, Chichester, UK: John Wiley & Sons; 2010.

[2] T. P. Carvalho, F. A.A.M.N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, S. G.S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Computers & Industrial Engineering*, vol. 137, 2019.

[3] K. Siwek, S. Osowski, "Data mining methods for prediction of air pollution," *International Journal of Applied Mathematics and Computer Science*, vol. 26, no. 2, pp. 467-478, 2016.

[4] R. M. Balabin, S. V. Smirnov, "Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data," *Analytica Chimica Acta*, vol. 692, no. 1–2, pp. 63-72, 2011.

[5] Y. Benjamini, and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," Journal of the Royal Statistical Society. Series B (Methodological), vol. 57, no. 1, pp. 289–300, 1995.

Identifying and representing clusters of spatial defects in microelectronics planar artefacts

Riccardo Borgoni[1], Chiara Galimberti[1], Diego Zappa[2]

[1)] Università di Milano-Bicocca, Milano, Italy.
[2)] Università Cattolica del Sacro Cuore, Milano, Italy.

Riccardo Borgoni - e-mail address: riccardo.borgoni@unimib.it
Chiara Galimberti - e-mail address: c.galimberti19@campus.unimib.it
Diego Zappa - e-mail address: diego.zappa@unicatt.it

Keywords: CSR, p-value control chart, Monte Carlo test

ABSTRACT – The aim of this work is to investigate the spatial structure of defectivity in integrated circuits fabrication. A prompt detection of an excess of defects and their spatial structure is critical for the entire fabrication process in order to limit yield loss. We propose first a control chart for spatial defectivity on the wafer area based on p-value. Then, a clustering procedure based on the minimum spanning tree algorithm is suggested to identify the areas of the wafer more prone to high defectivity and display their shape effectively.

1. INTRODUCTION

Integrated circuits are built through several different physical and chemical steps that are performed on thin silicon slices, called wafers. Thousands of chips can be obtained from each single wafer and the precision of the manufactured products is essential. The presence of defects is unavoidable, however structured defectivity, such as defect tendency to occur in clusters for instance, can be usually ascribed to specific causes of the production process. Therefore, a quick identification of defects leads to improvement of yield. We propose a flexible methodology which does not rely on strong parametric assumption and it is computationally convenient and easily implementable, also in-line to the process.

Firstly, as proposed by Lei et al. [1], a p-value control chart is introduced to detect potential structured patterns on the wafer area when a process distributional shift occurs, considering as benchmark a process with complete spatial randomness.

Secondly, an agglomerative clustering algorithm based on minimum spanning tree (MST) algorithm is used to identify the main clusters.

2. METHODOLOGY

It is assumed that a defect can be represented by a random point occurring in the wafer area W, and the spatial dynamics of defectivity is assumed to be governed by a spatial point process over W. The lack of interaction among points is named complete spatial randomness (CSR) property. Being N(W) the (random) number of events occurring in W and supposing that N(W) = n, the conditional property of CSR states that these n points are independent and uniformly distributed in W. If the CSR condition holds true, no structures are present in the defectivity process and defects are somehow a physiological result of the fabrication process. To test the CSR condition, the empirical cumulative distribution function (ecdf) of the nearest neighbor distance (NND) is compared to the cdf expected under the CSR condition. The p-value of the statistic test is obtained by approximating the distribution of under the null hypothesis of CSR via Monte Carlo simulations.

The p-value control chart is construct using a CUSUM approach [2]: the defects occurred up to time t are cumulated using the ecdf of the NND of the cumulated point patterns, the test statistics and the p-value at time t are computed.

In Figure 1 a control chart is showed: the time is showed on the x-axis and the p-values on the y-axis, the horizontal line represents the significance level of the test and the confidence interval is obtained via bootstrap method.

Once an out-of-control is detected it is relevant to assess the shape of the clusters in order to understand the defectivity process and plan remediation actions. Therefore, a clustering algorithm based on MST is implemented. The algorithm starts from the construction of MST considering all the defects. Then, sequential operations of separating are executed eliminating the longest edge of the tree. In this way, different clusters are defined as groups of defects remained connected to each other. An edge is eligible to being removed if its length is greater than the 90th percentile of the simulated distribution of the length of edges under the assumption of CSR, given the number of defects.

3. RESULT AND DISCUSSION

The methodology discussed above is applied retrospectively to a dataset of defects occurred in the microchip fabrication process. The coordinates of the defects occurred during the production process were detected by a laser scan of their surface. Wafers were scanned one after the other and the defects occurred at each time point were layered on the top of those previously scanned. At each time point the test statistics is computed on the set of cumulated defects and the p-value has been computed using simulated replicates drawn if the CSR property holds true. Instead, the p-value has been bootstrapped several times in order to obtain a 95% percentile confidence interval to account for the variability induced by the simulation. The result is shown in Figure 1. According to the chosen level of significance, the chart displays at what time the fabrication should be stopped. In general, the chart is a useful tool to monitor the behavior of the process. Then, the clustering algorithm is applied and the obtained result by separating the MST is shown in Figure 2. The distribution of the average length of the MST edges has been estimated via simulations. The skeleton of the sub-trees obtained by the procedure represents, however, a somewhat crude shape of the scratches occurred on the wafer area during the fabrication process.

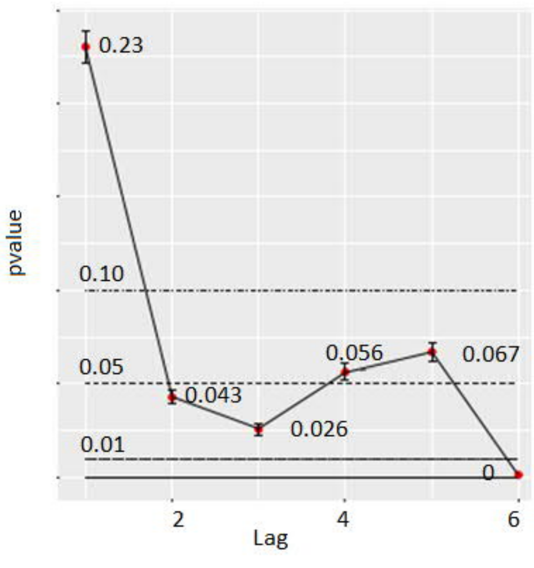

Figure 1: P-value control chart

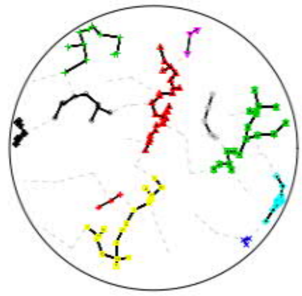

Figure 2: Identification of clusters with MST (the gray dashed lines correspond to the edges of the MST before separating step)

4. CONCLUSIONS

This work proposes a fully non-parametric procedure to detect defect patterns in wafer fabrication process control in the microelectronic industry. The procedure is based on two successive steps. Assuming that defectivity is governed by a spatial point process, a p-value control chart is constructed to detect any possible spatial structure of defect patterns using a test for the CSR property at each time point. The second phase aims at identifying the location of defects. Particular attention is paid to the identification of clusters having a complex, possibly non-convex, patterns typically difficult to detect and display. The advantage of the procedure discussed in this paper is twofold. Firstly, it is found reasonably easy to implement and computationally efficient. It also represents a powerful exploratory tool to identify effectively spatial structures in defect patterns that can be followed by more in-depth, possibly model based, off-line analysis of the clustering. The procedure in fact proves to be effective in detecting clusters of defects. Having used prospectively retrospective off-line data, we were in the position to cross-validate our methodology using defects data collected on successive wafers processed by the equipment considered and prove that most defectivity in new wafers of the same production line is concentrated in small areas on the surface. This fact identifies structured defectivity within the process. Secondly, the p-chart described proved to be extremely fast in detecting spatial structures requiring that only a few wafers are processed.

5. REFERENCES

[1] Li, Z., P., Qiu, S. Chatterjee, Wang Z.: Using p values to design statistical process control charts. Statistical Papers. 54, 523-539 (2013)

[2] Page, E.: Continuous inspection schemes. Biometrika. 42, 100-115. (1954)

StEering
Inter-University Research Center:
Design, Quality and Reliability

Design of experiment for multi-stage processes

Rosamarie, Frieri*[1,2], Marco, Mariti[2], Marilena, Paludi[2], Alessandro Pieri[2]
[1)] Department of Statistical Sciences, University of Bologna, Italy
[2)]GSK, Siena, Italy

*Corresponding author - e-mail address: rosamarie.frieri2@unibo.it

Keywords: Design of Experiment; Multi-stage process; Average D-optimal design

ABSTRACT
We propose a design strategy for the experimental planning of a process composed by multiple stages. Starting from a description of the process in which the response of each stage is characterized by a regression model whose covariates are the input factors of that stage and the output of the previous one, we implement D-optimality in average to plan the experiment in each stage; experiments are then performed sequentially. This strategy allows to estimate the model for the final response by taking into account the interactions among the subsequent stages with, in general, less experimental effort with respect to the classical DoE approach.

1. INTRODUCTION

Most industrial processes are composed by multiple stages, as summarized in Figure 1. We denote with $\mathbf{x}^{(i)}$ the vector of input factors of S_i (i-th stage) and with y_i the unidimensional outcome of stage i for $\mathrm{i} = 1, .., \mathrm{V}$.

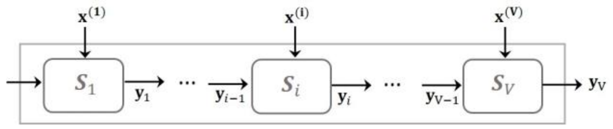

Figure 1: V-stage process

In these processes, the behavior of the final outcome, y_V, depends on all the input factors of the process and on the complex interactions among the subsequent stages. The experimenter is interested in finding the model for y_V to control, predict and optimize the process. The design of an experiment to fit such a complex model would require a high number of runs due to the high number of parameters to estimate. We propose an alternative description that models the output of each stage of the process. The idea is to include each intermediate output ($y_i, i = 1, .., V-1$) as input factor in the next stage so that the model for y_i is a function of the input factor specific of i-th stage and the output y_{i-1}. Hence, y_i is influenced by the input factors of S_i, the $\mathbf{x}^{(i)}$'s, which can be set directly, and by the additional input factor y_{i-1}. However, since y_{i-1} can be controlled only through the model for S_{i-1}, it cannot be set precisely to the desired level because of the prediction error of the model. Within this framework, V experiments have to be carried out to estimate the V models, but the error in setting the level of an input factor has to be taken into account in planning the experiment. We propose to use methodologies for designing experiments in presence of error in factor levels to handle with multi-stage experiments. To the best of our knowledge, the first author interested in the effect of error in setting factor levels was Box [1] and this topic has been more recently discussed by Pronzato [2] and Donev [3]. In the context of multi-stage processes, we apply these methodologies to plan the experiment of each single stage.

2. METHODOLOGY

Let us consider the i-th stage. Assume that $\mathbf{x}^{(i)}$ are set without error while y_{i-1} has the prediction error caused by the model estimated at S_{i-1}. Thus, the information matrix of the model for y_i has random components due to the values of y_{i-1} which fluctuates around the desired levels. If the experimenter is interested in the design that provides an estimate of model parameters as precise as possible, the well-known D-optimal criterion is employed in absence of error in setting the experimental conditions. As an extension, consistently with Pronzato [2], we consider the Average D-optimality that is aimed at minimizing the expected value of the determinant of the random information matrix. So the design is chosen as the one optimizing the average D-optimality criterion and we assume that when the experiment takes place the real values of the y_{i-1} are known, in other words the information matrix is random at the design phase but it is deterministic at the analysis phase. To implement the above mentioned criterion, we have used a Fedorov's type algorithm [4] that at each iteration computes a Monte Carlo estimate of the expected determinant. Notice that the generated design is the one that is optimal in average and in an experiment a single design is used that may be better or worse w.r.t. the expected criterion. To address this issue, we have extended the concept of Fraction of Design Space plot [5] for designs where some input factors are set with error to pre-experimentally evaluate D-optimal designs in average.

2.1 Evaluate the prediction capabilities of average D-optimal designs. Consider the Scaled Prediction Variance (SPV) of the predicted mean in a point of the design space (defined for example in [5]). The FDS plot shows the fraction of the design space at or below any SPV value. In cases of information matrix with random components, the SPV itself, at each point of the design space, is a random variable. In the FDS in Figure 2 we then consider the average scaled prediction variance (solid line) and the 95% quantile (dashed line) of the SPV for each point on the design space. For example, the cross in Figure 2 means that the chosen design has the 80% of the design space with average SPV less than or equal than 0.41 while the black dot tells us that the 95% of the possible realizations of the chosen design have the 80% of design space with SPV less than or equal than 0.57.

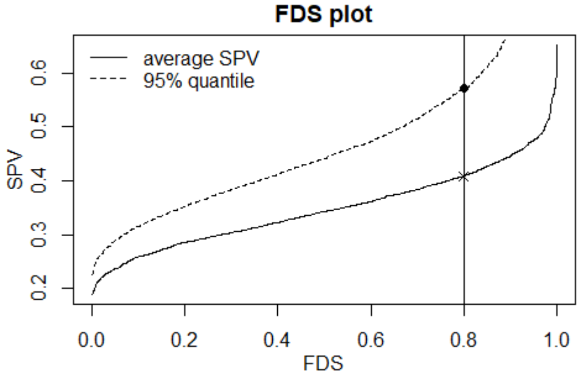

Figure 2: Fraction of design space plot

3. RESULTS AND DISCUSSION

We now summarize the procedure to design experiments to fit a model for the final outcome of the process.

(i) Stage 1: input factors can be controlled directly, then D-optimal criterion can be used to plan the experiment. Experiment for S_1 is carried out and the model for y_1 is fitted.

(ii) Stage 2: $\mathbf{x}^{(2)}$ can be controlled directly while y_1 can be only set by changing $\mathbf{x}^{(1)}$ and using the model estimated at S_1. The error in setting y_1 to each level is computed and average D-optimality is implemented to find the optimal design for Stage 2. Before performing the experiment for S_2, we apply Stage 1 to prepare the starting material to the levels required by the experimental design of S_2. When experiment for S_2 takes place, the real experimental conditions are measured and the model for y_2 is estimated via least squares.

(iii) *… the procedure is iterated up to the last Stage*

(iv) Stage V: the experimental conditions are determined by $\mathbf{x}^{(V)}$ and y_{V-1}. Similarly to (ii) the design is chosen according the average D-optimal criterion and after performing the stages up to S_{V-1} to prepare the starting material, the experiment to fit the model for y_V can be carried out.

Example: V=3 stage process.

Assuming that the tentative model for each step include intercept, main effects, quadratic effects and two factor interactions effects, we compare the number of parameters to estimate in the "full model approach" in which classical DoE is used (**Classic**) and in the proposed one (**New**). In Table 1 we show that, if each stage has two or more input factors, our strategy requires less experimental effort and the saving of resources increases as the number of input factors per stage increases.

Table 1: Number of parameters to estimate

Num. controlled input factors per stage	**Tot. parameters**	
	Classic	**New**
2,1,1	15	18
3,2,2	36	30
4,3,3	66	45
5,4,4	105	63
$n, n-1, n-1$	$\frac{3}{2}(3n^2 - n)$	$\frac{3}{2}(n^2 + 3n + 2)$

4. CONCLUSIONS

We propose a new design strategy to plan the experiment for a multi-stage process with the objective of finding a model for the final output of the process. Moreover, this approach allows the experimenter to take into account the interaction between the stages, that is also of main interest in multi-stage process.

5. REFERENCES

[1] G. E. P. Box, "The Effects of Errors in the Factor Levels and Experimental Design". *Technometrics*, 5:2, pp. 247-262, 1963.

[2] L. Pronzato, "Information matrices with random regressors. application to experimental design". *Journal of statistical planning and inference*, vol. 108, no. 1-2, pp. 189-200, 2002.

[3] A.N., Donev. "Design of experiments in the presence of errors in factor levels", *Journal of statistical planning and inference*, vol. 126, no. 2, pp. 569-585, 2004.

[4] Fedorov, V. V. (1972). Theory of Optimal Experiments. Translated and edited by W. J. Studden and E. M. Klimko. New York: Academic Press.

[5] A. Zahran, C. M. Anderson-Cook & R. H. Myers "Fraction of Design Space to Assess Prediction Capability of Response Surface Designs", *Journal of Quality Technology*, vol 35, no. 4, pp. 377-386, 2003

This work was sponsored by GlaxoSmithKline Biologicals SA. The authors have declared the following conflicts of interest: Rosamarie Frieri is a PhD student at GSK, supervised by Marilena Paludi. Marilena Paludi, Marco Mariti and Alessandro Pieri are employees of the GSK group of companies.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Research on the Current Situation of Intelligent Logistics Management and Internet Logistics Statistics in China

Qian Guo

Beijing Wuzi University

*Corresponding author - e-mail address:guoqian561@126.com

Keywords: Intelligent Logistics Management; Internet Logistics Statistics

ABSTRACT – During the "13th Five-Year" period, the implementation of the "Internet +" strategy in China, China's smart logistics is in the period of development opportunities, and the accelerated transformation and upgrading of intelligent logistics has become an inevitable trend. The rapid development of intelligent logistics has promoted the upgrading and transformation of the logistics industry, and the emergence of new forms of Internet logistics has brought challenges to the government's statistical work. This paper introduces the concept and characteristics of intelligent logistics, and puts forward the statistical classification system and statistical environment of Internet logistics.

1. Introduction

Intelligent logistics is a new ecology based on Logistics Internet and big data, through collaborative sharing innovation mode and advanced artificial intelligence technology, reshaping industrial division, rebuilding industrial structure and changing industrial development mode. "As the most advanced development form of logistics industry.

2. Characteristics of intelligent logistics

2.1 Auto sensing

Intelligent logistics uses the Internet of things technology to realize the digitization of all logistics elements and business data, and realize the automatic perception, transmission and storage of all kinds of data information, so as to realize the real-time interconnection between intelligent logistics network and physical network.

2.2 Self judgment

Intelligent logistics monitors logistics activities in real time, analyzes data timely according to preset logic and rules, finds weak links or loopholes in the logistics process, and predicts the potential impact of problems.

2.3 Intelligent decision making

On the basis of information analysis and judgment, the paper puts forward corresponding methods, measures and schemes according to the constraints, so as to realize the program control and automation development of logistics system.

2.4 Automatic execution

Within the scope of human authorization, data-driven decision-making implementation and execution will minimize the risk and reduce human intervention.

2.5 Deep collaboration

It can realize the overall optimization based on the global optimization intelligent algorithm of the logistics system, and dispatch the efficient division of labor and cooperation of all participants in the whole logistics system.

2.6 Intelligent learning

Through active learning of new knowledge, it is to create language and thinking ability to adapt to the environment.

3. Internet logistics and its influence

Internet logistics is the integration of physical world logistics system and online Internet world logistics information system. In this system, the Internet becomes the leading and control core of logistics entity operation, the "brain" and nervous system of logistics system, and extends and seamlessly connects to the offline logistics system through the logistics information internet, so as to realize the all-round interconnection of the physical world logistics system. The rapid development of logistics Internet has triggered a new revolution in the field of logistics, making modern logistics really enter the "intelligent logistics era".

With the development of Internet, logistics management using Internet has the characteristics of low cost, real-time dynamic and customer driven. In terms of employee management, the management requirements for employees are higher. Due to the lack of talents, it is also necessary to actively introduce advanced talents in human resource management. Internet and logistics industry complement each other. With the mutual promotion of the two, enterprises have established a variety of management platforms to assist business development, including decision support system, transportation information system, inventory information system, distribution information system and e-commerce platform.

4. Statistical classification system of Internet Logistics

There are five classification systems of Internet logistics enterprises, as table 1.

Table 1 Statistical Classification System

Classification from the regional perspective of logistics	International Logistics
	Urban Logistics
	Rural logistics
Classification according to the particularity of logistics	General Logistics
	Special logistics
Classification according to the perspective of logistics operators	Self-operated Logistics
	Third party logistics
	Crowdsourcing Logistics
Classification according to the fields involved in Logistics	Logistics of daily necessities
	Pharmaceutical logistics
	Catering logistics
	Dangerous goods logistics
	Automobile logistics
	Home appliance logistics
	Retail logistics
	Agricultural products logistics
Classification according to the timeliness of logistics	Standard delivery
	Immediate delivery
	Limited time delivery
	Precise delivery

5. Results and Discussion

5.1 Data collection environment

The first step of logistics statistics is to collect dynamic data accurately, quickly and timely, which can also become the perception of data. The perception level includes recognition system, location system and tracking system. Sensing technology equipment (RFID, barcode gun, sensor, etc.), combined with GPS, positioning and tracking, can collect the information of logistics system unit in real time and automatically, and then process the information according to the system structure and operation logic, so as to realize the real-time and process information palm of objects and other objects.

5.2 Data management environment

Data management is divided into transmission layer and storage layer. The transmission layer uses various transmission networks and communication technologies to transmit the information collected by sensing devices in a timely and safe manner. Transmission media include Internet, mobile communication network, cluster communication technology, etc.

5.3 Data analysis and application environment

After using high-end technology to store, process and mine data, it is applied to exchange interface, public service platform and user application. It directly provides users with required information, provides data support for their decision-making, provides users with customized services, reduces application costs, and improves processing efficiency. At the same time, it also realizes commodity traceability, waybill tracking, intelligent sorting and distribution, forecast and early warning. In addition, the intelligent decision system can also make scientific decisions, put forward optimization suggestions for distribution routes, and provide decision-making reference for enterprises, transportation departments and government departments.

6. Reference

[1] Wu Ping. New momentum, trend and path of intelligent logistics development in the context of "Internet +" [J]. *Business economy research*, 2018 (4)

[2] Zhou Dingbo, Deng song. Transformation strategy of traditional logistics to intelligent logistics [J]. *Open guide*, 2017 (12)

[3] Jinjiangjun. Research on Development Strategies of intelligent industry [J]. *Research on technology economy and management,* 2012 (11)

[4] Chen Lisu. Research on the development of smart industry and the transformation and upgrading of China's industry [J]. *Reform and strategy*, 2016 (10)

[5] Wu Ningjie. Discussion on the transformation path of China's intelligent logistics under the background of "Internet +" [J]. *Business economy research,* 2018 (6)

[6] 6. Li Guanyi. Internet thinking on re innovation of e-commerce logistics and transformation of traditional logistics -- Thinking Based on value connection [J]. *Science and technology management research*, 2016 (9)

StEering
Inter-University Research Center:
Design, Quality and Reliability

Sampling Survey of Industrial Enterprises and Study on Related Issues

Yongjin, Jin[1)2)3)]

[1)] Center for Applied Statistics, Renmin University of China, Beijing 100872, China.
[2)] School of Statistics, Renmin University of China, Beijing 100872, China.
[3)] Institute of Survey Technology, Renmin University of China, Beijing 100872, China.

*jinyongj_519@aliyun.com

Keywords: Industrial Sample Survey under a Designated Size; the Standard of Scale; Sampling Frame

1. ABSTRACT

An industrial survey under a designated size is one of the most important components of the industrial statistical survey and a primary tool for collecting industrial data under a designated size. Based on the hackle of the history of industrial surveys under designated size, this paper summarizes the current sample survey designs in China. Then combined with field surveys, from the perspective of data, several related problems in the industrial surveys under the designated size were studied, such as the standard of scale to categorize industrial size and sampling frame. Finally, corresponding improvement suggestions to the above problems are proposed.

2. INTRODUCTION

The sampling survey of industrial enterprises is also known as the "sampling survey of industrial enterprises below designated size" and is one of the statistical sampling survey projects of the Chinese government. The industrial output value is an important component of GDP. The data comes from two parts. One is the above-scale enterprises. These enterprises report the output value and other related data to the National Bureau of Statistics through “online reporting”. The other is the sub-scale enterprises. The number of these enterprises is large, but the scale is small. The relevant can be obtained through sampling surveys [1].

The sample survey of industrial enterprises under a designated size began in 1997. After 20 years of exploration and development, a scientific and effective survey plan has been formed. The current scale standard is 20 million RMB of annual operating income. This standard was 5 million from 1997 to 2010 and adjusted to 20 million in 2011[2][3].

The situation of sub-scale enterprises is very complicated and can be divided into three types: A, enterprises in the catalogue. The names of the enterprises come from the Economic Census conducted every 5 years; B, enterprises not in the catalogue. It is mainly the newly founded enterprises after the census; C, individual enterprises. They did not register in the relevant departments, but carry out production activities. Sampling is carried out for sub-scale enterprises, and it is divided into two parts here, one is catalogue sampling for A and the other is region sampling for B and C [4].

3. METHODOLOGY

Catalogue sampling is used to investigate the population A. A One-stage stratified random sampling is used. The sampling design is based on the province, and all sub-scale enterprises in the province are stratified by industry (41 categories). It can be divided into 5 steps:

(1) Constructing Sampling Frame. The scope of the sampling frame is to include all the sub-scale industrial in the list. The content of the sampling frame includes the enterprise code, address, industry classification, main business types, etc.
(2) Determining Sample Size. The sample size needs to meet the estimation accuracy requirements of the province level and industry level. It is determined by The National Bureau of Statistics.
(3) Making Stratification. It is based on industry. If some stratum has a large number of enterprises, it is necessary to further stratify according to the size of the enterprise.
(4) Selecting Samples. Each unit in the sampling frame is given a “permanent random number”. Selecting n_h enterprises with the smallest “permanent random number” are selected as the sample of the h-th stratum.
(5) Calculating the Weights. The basic weights are the reciprocal of the sampling probability. Some adjustments are needed to obtain the final weights.

Region sampling is used to investigate the population B and C. The region frame can only be used here. It can be divided into 3 steps:

(1) Constructing Sampling Frame. The frame is based

on province level and is the list of the village/ neighborhood committee.

(2) Calculating Sample Size. Cluster sampling is used to survey. All units in the selected committee are investigate. The sample size is the number of village/neighborhood committees to be drawn and is calculated according to accuracy requirements.

(3) Making Stratification. It is based on the number of individual industries in the village/neighborhood committee.

4. Results and Discussion

The sampling survey of sub-scale industrial enterprises is shown in the figure below:

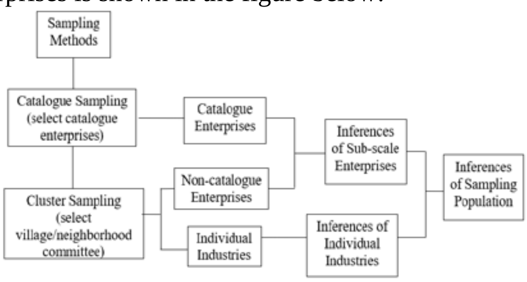

Figure 1 The sampling survey of sub-scale industrial enterprises

Two of the issues in industrial surveys under a designated size are discussed here.

(1) The Standard of Scale

The current scale standard is 20 million RMB of annual operating income. According to the calculation, the output value of above-scale enterprises now accounts for 92-93% of the total industrial output value, and the sampling estimate of the proportion of output value is less than 10%. Considering the unbalanced regional development, if we use uniform standards, the coverage of enterprises varies greatly between eastern and western regions. If the standard is not raised, the sampling section will be meaningless. Quantities of above-scale enterprises will lead to the quality problems of data. It is difficult to monitor a large number of companies.

What indicators are used as the basis for scale division? At present, it is operating income, but the operating income changes too quickly, and the regions are not balanced. Whether the number of employees can be used as the scale standard? It is related to output value and the intensity of the industry. Large companies have many employees and can be incorporated into the coverage of above-scale

What is the better Scale Standard? First, the number of enterprises above this standard is not very large, while account for a large proportion of the total output value. Second, after the adjustment, the sampling population will expand, and the larger the population is, the higher the sampling efficiency is. Third, the reduction in the number of companies over scale standard is conducive to data quality monitoring.

(2) Sampling frame

At present, there are two sampling frames in the sampling part, one is catalogue frame and the other is region frame. Catalogue frame can be obtained from the Economic Censususes conducted every five years. But it is difficult to grasp the appearance and disappearance of enterprises in a timely manner, especially for small enterprises. At the same time, the sampling frame is extremely expensive to maintain and difficult to operate. Region frame is a list of village/neighborhood committees. But the population and the number of individual industries vary from village to village. At the same time, the stratification is not precise enough, both sampling and non-sampling errors are large.

Here are two suggestions for improvement. First, merge the catalogue frame and the region frame into a single region frame to avoid the problems of the great changes of catalogue enterprises. Second, refine the stratification of the village /neighborhood committees in the region frame. Stratification factors include the number of catalogue enterprises, the number of individual industries, the size of the population, etc. The distribution of the industries should also be considered.

5. CONCLUSIONS

Since the pilot survey conducted in 1996, there have been many adjustments in industrial surveys under a designated size. A scientific and effective survey plan has been formed, but there are still many issues. Based on the hackle of the history of industrial surveys under designated size, we summary the current sample survey designs in China. Besides, two problems are studied and related suggestions are given.

6. REFERENCES

[1] P.J.Lei, “Design for the Programme of the Regular Industrial Statistics Pilot Sample Survey,” Statistical Research[J],vol. 14, Issue (5), pp.69-74, 1997.

[2] ‘The Sampling Frame Of Industry Under Scale’ Research Group, “A Study On Sampling Frame Of Under Scale Industry Firms In China,” Statistical Research[J], no.1, pp.41–45,2003.

[3] National Bureau of Statistics of China, Statistical report system of sampling survey for Industry Under Scale,”[EB/OL].http://www.stats.gov.cn/tjsj/tjzd/gjtjzd/201701/t20170109_1451439.html,2017-01-09.

[4] Y.J.Jin, T.Y.Jiang, “A Study on Some Issues in Industrial Surveys under a Designated Size,” Statistical Research[J], vol.36, no. 3, pp. 44-52,2019.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Evaluation of food safety traceability information transfer efficiency ——Based on PSR-FAHP model

Zhongdong Li[1], Xiaojuan Yuchi [2]

[1)] Shandong Technology and Business University Institute of Peninsula Economic Research.
[2)] Shandong Technology and Business University Institute of Economics.

[*]Corresponding author - e-mail address: 1531364336@qq.com

Keywords: food safety; information transfer efficiency; PSR-FAHP

ABSTRACT – The food safety traceability system has become the main measure to strengthen food safety management and prevent food safety risks. However, current food safety incidents indicate that food safety information is not converted into reliable and effective signals through the food safety traceability system. The key to improving the efficiency of food safety traceability information is to build an effective evaluation mechanism. Based on the analysis of the role of various stakeholders in the traceability information transmission from the perspective of food enterprises, based on the principle of index system construction, the pressure-state-response (PSR) model is used to establish an index evaluation system. Aiming at the complexity and ambiguity of food safety traceability information transfer, the gray fuzzy analytic hierarchy process (FAHP) is used as the theoretical basis to construct a comprehensive evaluation model, which analyzes the three internal and external pressures, states and response measures. And in each aspect, the specific evaluation indicators are formulated and the weights of each indicator are calculated. Finally, the overall efficiency level of current food safety traceability information transmission is obtained and targeted system improvement measures and suggestions are proposed.

1. INTRODUCTION

Interest conflicts and reconciliation between different interest subjects leads to higher transaction costs and information asymmetry as previous study[1].

The key point of the implementation of food safety traceability information transmission efficiency evaluation is to evaluate the information transmission efficiency of each interest subject in the food quality and safety traceability process. The key point is to select the effective efficiency evaluation index and the weight ratio of the evaluation index. Yusheng Chen etc., Jian Yu and Hongxia Zhang stated the food safety issues facing the reputational damage and economic loss, the enterprise to the will of the food safety quality control, enterprise operation costs and benefits as well as the management mode and the enterprise itself information equipment and personnel management ability, enterprise brand price tags such as certification, enterprise risk emergency response mechanism, etc. [2-4]. Wang, JN et al. proved that the macro and micro environment will become an important basis for enterprises to make cooperative strategic decisions in the food supply chain [5]. Golan et al. think that food-related enterprises can confirm that in the consumption process, if consumers tend to pay extra product prices for traceable products, the food enterprises will have a stronger investment tendency [6]. According to Starbird, the main reason for some manufacturers to establish a supply chain traceability system is the pressure from downstream enterprises [7].

PSR model was proposed by Tony Friend and David Rapport [8].It can better reflect the causal relationship among natural, economic and social interactions in the ecosystems. This method not only consider the efficiency of the system itself but also the external environment of pressure and response ability. It is the manifestation of the essence of the sustainable development evaluation model, because the process of food traceability information is based on food safety traceability system, so choose the PSR model evaluation of food safety traceability information transmission efficiency..

Through the literature review found that the fuzzy comprehensive evaluation method and analytic hierarchy process (FAHP) is widly used, two kinds of supply chain performance evaluation system by using FAHP to solve food quality traceability system of the construction of the advantage is able to differentiate between different elements and specific analysis of the importance of each influence factor in the system.

2. METHODOLOGY

Methodology include PSR evaluation system model and FAHP evaluation method.

2.1 Construction of Food Safety Traceability Information Transmission Efficiency Index system

Table 1 Food safety traceability information transmission efficiency index system based on PSR model

Target layer	System layer	Indicator layer
Food safety traceability information transmission efficiency index system based on PSR model	Pressure(P)	
	State(S)	
Food safety traceability information transmission efficiency index system based on PSR model	Response(R)	

2.2 The grey fuzzy hierarchy evaluation method is used to evaluate the efficiency

1.Establishing the factor set of efficiency evaluation index according to Table 1.

2.Establishing fuzzy judgment matrix:

$$A = \left(a_{ij}\right)_{n\times n} \quad (1)$$

3.Calculating the weight of each index in the fuzzy judgment matrix.

$$W_i = \frac{-1+\frac{n}{2}+\sum_{j=1}^{n} a_{ij}}{n(n-1)} (i = 1,2,\cdots,n) \quad (2)$$

4.Consistency test of fuzzy complementary judgment matrix

$$I\left(Y_1,W_1^*\right) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\left|a_{ij} + w_{ij} - 1\right| \quad (3)$$

5.Establishing conclusion set of grey fuzzy evaluation.

6.Establishing the evaluation sample matrix.

7.Determining the evaluation gray class.

8.The grey evaluation system is calculated and the evaluation weight matrix is established.

9.The grey fuzzy synthesis vector is calculated and evaluated comprehensively.

3. RESULTS AND DISCUSSION

According to the above calculation results, according to the principle of maximum membership degree, it can be seen from the matrix C_1 that the pressure of internal and external environmental forces facing the current food enterprises on the efficiency of traceable information transmission is at a general level, and the pressure of enterprises on food safety issues effectively promotes the traceable information transmission. C_2 matrix represents the current state of information transmission efficiency of the enterprise, which is shown as low efficiency, indicating that the construction of traceable information transmission of most enterprises is slow and low. According to the matrix C_2, it can be found that the response index is a general response, indicating e enterprises have not established a perfect information transfer mechanism. The final evaluation value is 5.286, that is, the gray class grade produced is normal, indicating that the current food safety traceability information transmission efficiency belongs to the "general efficiency" level.

4. CONCLUTIONS

According to the above results, food enterprises are generally effective in coping with internal and external pressures. Most of them fail to pay attention to the importance of food traceability information to stakeholders and the important role of their own development needs, and accordingly fail to make effective responses to ensure food production safety. Therefore, the current food safety traceability efficiency level is low. Relevant measures should be taken to improve the efficiency of food safety traceability information transmission by combining efficiency levels, internal and external pressures faced by enterprises, factors affecting stakeholders and information transmission status.

5. REFERENCES

[1] Shi Sheng, Zhou Deyi, Wang Puqing. Research on information transmission efficiency and government governance strategy of traceable food safety system [J]. Journal of agricultural economics, 2008(5): 20-26, 110.

[2] Zhang Hongxia. Game analysis of online food quality and safety signal transmission [J]. Journal of agricultural technology and economy, 2017(9): 116-123.

[3] Yu jian, Yang bin. Quality ethics, information transmission and model change: strategic innovation of food enterprises from the perspective of social responsibility [J]. China soft science, 2016, (1): 184-192.

[4] Chen Yusheng, Fang Ruijing, Yin Shijiu, et al. Game analysis of traceability information transmission between food enterprises and consumers [J]. Journal of ocean university of China (social science edition), 2014, (5): 68-72.

[5] Wang J, Chen T, Wang J. Research on cooperation strategy of enterprises' quality and safety in food supply chain[J]. Discrete Dynamics in Nature and Society, 2015.

[6] Golan E. D, Barry K. W. Food Trace ability: One Ingredient in a Safe and Efficient Food Supply[J]. Prepared Foods,2014,37 (4):59-70;

[7] Starbird A. L., Sohet D. A. The demand for food safety: market imperfections and the role of government[J]. Food Policy,2016,18 (2): 152-162;

[8] RAWER W. Development of environmental indicator systems: experiences from Germany. Environmental Management， 2000， 25 (6):613-623.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Measurement and Convergence Analysis of Coordinated Development Level of Industry and Employment under the High-quality Economic Development in China

Dingping, Liu [1]
Pei, Liu [2]

[1)] College of Statistics and Big Data, Henan University of Economics and Law, Zhengzhou, Henan, China 450046,
[2)] Economics School, Zhongnan University of Economics and Law, Wuhan, Hubei, China 430073

[*] Corresponding author: Dingping Liu
e-mail address: zzldp@126.com

Keywords: High-quality economic development; coordinated development of industry and employment; convergence

ABSTRACT – This paper measured the industrial and employment coordinated development level in China and studied its spatial correlation, regional heterogeneity and convergence characteristics. This study found that the average national coordinated development index during 2004-2015 shows a rising trend. Regionally the eastern and central demonstrate relatively high coordinated development, but the high growth areas are mainly located in the western and northeastern. The alpha convergence coefficient is in downward trend and tends to be steady. This study is conducive to promoting the coordinated and unbalanced development of regional industry and employment, and realizing high-quality economic development.

1. INTRODUCTION

The differentiated and unbalanced state of China's regional development has made industrial upgrading and industrial regional transfer a normal state in the process of economic structural transformation. How to achieve the coordination between industry and employment structure in the process of industrial transfer, and achieve sustained and coordinated development of regional population, resources, and environmental protection has become the key to high-quality economic development in China. The coordinated development of industry and employment under high-quality development needs to be based on the "Five Development Concepts of innovation, coordination, green, openness, and sharing" [1], "Theory of Market Resource Allocation" [2], "Theory of Regional Integrated Development" [3], "Harmonious Development between People and Nature" [4], and "urban and rural integrated development" [5], etc. This study aims to investigate regional coordinated development levels and convergence features by measuring and analyzing the industrial and employment coordinated development level.

2. METHODOLOGY

This study constructed an industrial development index system (industrial structure, industrial innovation, industrial investment, industrial opening, industrial consumption, industrial pollution, industrial efficiency) and employment development index system (employment structure, employment density, employment income, employment pressure, employment security, job market, employment environment) in 7 dimensions system and 12 subdivisions for 279 prefecture-level cities. The coordinated development index of industry and employment is calculated based on multiple coupling degrees: first the value of the subdivided index is standardized (0-1 standardized); then the industrial and the employment development index are calculated according to the entropy method; and finally, the multiple coupling coordination index CT is obtained.

$$C = \left\{\frac{u_1 u_2 + u_2 u_1}{(u_1 + u_2)^2}\right\}^3$$

$$\mathrm{T} = u_1 * w_1 + u_2 * w_2$$

$$\mathrm{CT} = \sqrt{C * T} \qquad (1)$$

Convergence and absolute convergence in the method of convergence are both absolute convergences.

Different coefficients are used to indicate the degree of dispersion. The specific formula is as follows:

$$S = \sqrt{\frac{\sum_{i=1}^{N}\left(index-\overline{index}\right)^2}{N}}$$

$$CV = \frac{S}{\overline{index}}$$

$$\alpha = \sqrt{\frac{\sum_{i=1}^{N}\left(\ln index-\overline{\ln index}\right)^2}{N}} \quad (2)$$

Where i is the region, S is the standard deviation, CV is the coefficient of variation, index is the coordinated development index, and N is the total of all regions.

3. RESULTS AND DISCUSSION

At the national level, the average value of the coordinated development index of industry and employment between 2004 and 2015 showed an upward trend. The index increased from 0.1 in 2004 to 0.145 in 2015, a growth rate of 45%. By region, Figure 1 shows that the coordinated development index of the eastern, central, western, and northeastern regions in the range of 2004 to 2015 was 0.132-0.149, 0.095-0.133, 0.076-0.103, and 0.087-0.119, with growth rates in order 12.87%, 40.00%, 35.52%, and 36.78%. The overall coordination level is in the order of high to low is east, middle, northeast, and west. The order of growth rate from high to low is central, northeast, western, and eastern. It can be seen that the overall level of the eastern region is the highest but the growth rate is the lowest.

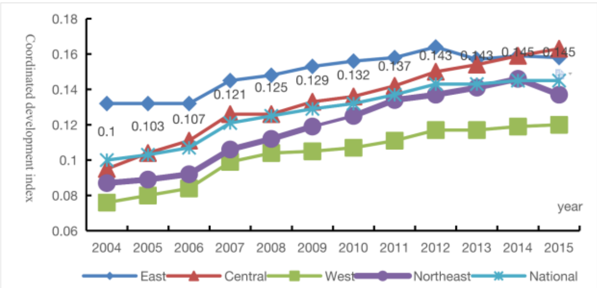

Figure 1 Changes in the coordinated development index of industry and employment in different regions of China

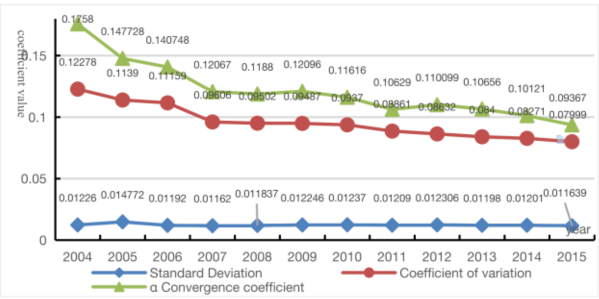

Figure 2 Convergence of the overall national coordinated development coefficient

Figure 2 shows the change of the convergence coefficient of the coordinated development coefficient. The standard deviation coefficient is basically below 0.002 and remains at a stable level. The coefficient of variation is in a downward trend, from 0.1228 in 2004 to 2015. The α convergence coefficient is also in a downward trend. The α convergence coefficient decreased from 0.1758 in 2004 to 0.0937 in 2015. The α convergence coefficient in 2015 was close to half that in 2004, and the decline rate reached 46.70% compared to 2004. It shows three things: that the overall coordinated development of industry and employment in the country is at a near-steady level; the regional coordinated development strategy has been implemented; and the level of high-quality economic development has been improved to a certain level.

4. CONCLUSION

This paper uses the entropy method and multiple coupling coordination indicators to measure and analyze the level of coordinated development of industry and employment. The study found that the average value of the coordinated development index from 2004 to 2015 showed an upward trend. In terms of sub-regions, the level of coordinated development is higher in the east and central regions, however the regions with high growth rates are located in the west and northeast regions; the convergence coefficient is declining and is approaching steady levels. The findings provide evidence and reference for the coordinated development of high-quality economic development.

5. ACKNOWLEDGEMENT

This study is supported by the Special Fund for Fundamental Scientific Research from Zhongnan University of Economics and Law (No.201810339).

6. REFERENCES

[1] M. Wei., and S.H. Li, "Research on the Measurement of the Quality of China's Economic Development in the New Era," *The Journal of Quantitative & Technical Economics*, pp.3-18, 2018.（in Chinese）

[2] L. Zhang, G.S. Zhu, X.Y. Li., and X.B., Wang, "Key Industrial Policies and Local Government Resource Allocation," *China Industrial Economy*, PP. 63-80., 2017.（in Chinese）

[3] X.F. Huang, and H.Y. Zheng, "Analysis of regional integration, regional specialization and convergence - based on empirical analysis of 9 cities in the Pearl River Delta region," *Statistical Research*, PP.90-96., 2010.（in Chinese）

[4] H. Zhang., S.Y. Tang., and Z.R., Geng, "The level of synergistic development, dynamic trajectory and convergence of the integration of livable industry and ecological livability in the Yangtze River Delta urban agglomeration," *Quantitative Economics, Technology and Economy Research*, PP.91-96, 2019.（in Chinese）

[5] Y. Yi, and A.M. Zhang, "Research on the Coordinated Development Model of Urban-Rural Industrial Integration under the Background of Urban-Rural Integration," *Soft Science*, PP.105-109, 2018.（in Chinese）

StEering
Inter-University Research Center:
Design, Quality and Reliability

Development of Model-Robust Designs for Accelerated Failure Time Models with Right-Censored Data

M. J. Rivas López[1], R. Martín Martín[2*], I. García-Camacha Gutiérrez[2]

[1)] Department of Statistics, University of Salamanca, Plaza de los Caídos s/n, 37008, Salamanca, Spain.
[2)] Department of Mathematics, University of Castilla-La Mancha, Avda. Carlos III s\n, 45071, Toledo, Spain.

[*]Corresponding author - e-mail address: Raul.MMartin@uclm.es

Keywords: Acceleration factor, Right-Censoring, Asymptotic mean squared error

ABSTRACT – Accelerated Failure Time (AFT) models are commonly used in the field of manufacturing. These models are defined through the survival function of the time-to-event variable. This work deals with the construction of robust designs for AFT models with misspecified acceleration factor. We consider the class of contamination functions proposed by Huber (1975). In this case, the estimate is subject to both 'bias error' due to the inadequacy of the model, as well as 'variance error' due to sampling. We obtained the asymptotic mean squared error matrix of the parameter estimates for right censored observations. Finally, a typical distribution in AFT, Log-Logistic distribution, was considered to illustrate the results.

1. INTRODUCTION

The AFT models look for the parameters of a probability distribution for a time-to-event, T, that is affected by some covariates $\boldsymbol{x} = (x_1, x_2, \dots, x_k)$ giving the conditions under the experiment runs. Let T_u be the random variable T observed in the usual levels of the covariates, following a distribution given by a baseline survival function, S_u. The AFT model is

$$S(t|\boldsymbol{x}, \boldsymbol{\theta}) = S_u(t/AF(\boldsymbol{x}, \boldsymbol{\theta})), \tag{1}$$

where $AF(x, \theta)$ is an Acceleration Factor (AF), which speeds up the effect of the survival time in the survival function. The AF is typically the exponential of a linear relationship,

$$AF(\boldsymbol{x}, \boldsymbol{\theta}) = \exp(\boldsymbol{\theta}'\boldsymbol{x}). \tag{2}$$

After some algebra, the corresponding log-linear form of (1) with respect to time may be reduced to

$$\log(T) = \mu + \boldsymbol{\theta}'\boldsymbol{x} + \sigma\varepsilon, \tag{3}$$

where $\mu \equiv E(\log(T_u))$, $\sigma \equiv Var(\log(T_u))$ and ε is a random variable with mean 0 and variance 1.

For each distribution of T, there is a corresponding distribution for ε. Thus, the survival and probability density functions of T can be expressed by the ones of ε

$$S(t|\boldsymbol{x}; \mu, \boldsymbol{\theta}) = S_\varepsilon((\log(t) - \mu - \boldsymbol{\theta}'\boldsymbol{x})/\sigma) = S_\varepsilon(\mathrm{z_t}),$$

$$f(t|\boldsymbol{x}; \mu, \boldsymbol{\theta}) = \frac{1}{\sigma t} f_\varepsilon(\mathrm{z_t}).$$

2. METHODOLOGY

If the experimenter assumes the form of the acceleration factor is (2) but the true accelerator factor is

$$AF(\boldsymbol{x}, \boldsymbol{\theta}) = \exp\left(\boldsymbol{\theta}'\boldsymbol{x} + \frac{g(x)}{\sqrt{n}}\right),$$

where the function $g(x) \in \mathcal{F}$ represents uncertainty about the exact AF form, the true log-linear model (3) is

$$\log(T) = \mu + \boldsymbol{\theta}'\boldsymbol{x} + \sigma\tilde{\varepsilon},$$

where $\tilde{\varepsilon} = \varepsilon + \frac{g(x)}{\sigma\sqrt{n}}$ and $f(t|\boldsymbol{x}; \mu, \boldsymbol{\theta}) = \frac{1}{\sigma t} f_{\tilde{\varepsilon}}(\mathrm{z_t})$.

When bias occurs, the Mean Squared Error of the parameter estimates $MSE(\widehat{\boldsymbol{\theta}})$ rather than variance becomes a suitable measure of the design quality. Then, a function must be derived from $MSE(\widehat{\boldsymbol{\theta}})$ in order to optimize some aspect of model. In this work, we consider the loss function according to the D-optimality criterion

$$\mathcal{L}_D(g, \xi) = \det(MSE(\widehat{\boldsymbol{\theta}}))^{1/p} \tag{4}$$

where p is the number of model parameters and ξ is the selected design. Therefore, our aim will be to achieve the design optimizing (4) adopting a minimax approach

$$\xi^* = \arg min_{\xi\in\Xi}\ max_{g\in\mathcal{F}}\ \mathcal{L}_D(g, \xi). \tag{5}$$

3. RESULTS

Let us consider all the units under study are set at time 0 and there is a follow-up until time d. The end of the study is the only censoring considered here. Let us consider known $\sigma \equiv Var(\log(T))$.

Let $\delta \equiv 1_{\{T<d\}}$ be an indicator variable being 1 if the observation corresponds to an event and 0 if it corresponds to a censored time. Thus, the log-likelihood function for model (3) and just one observation at value $\boldsymbol{x}$ is

$LL(\mu, \boldsymbol{\theta}; t, \boldsymbol{x}) = -\delta \log(\sigma t) + \delta \log\big(f_\varepsilon(z_t)\big) + (1-\delta) \cdot \log(S_\varepsilon(z_t))$

Let us consider an approximate design

$$\xi = \begin{Bmatrix} x_1 & x_2 & \dots & x_n \\ \omega_1 & \omega_2 & \dots & \omega_n \end{Bmatrix},$$

so that $0 \le \omega_i \le 1$ and $\sum_{i=1}^n \omega_i = 1$, and let $\boldsymbol{\beta} = (\mu, \boldsymbol{\theta})$ be the vector of parameters and $\boldsymbol{\beta_0} = (\mu_0, \boldsymbol{\theta_0})$ their true values.

Theorem 1. *The asymptotic distribution of the Maximum Likelihood Estimator (MLE) is*

$$\sqrt{n}\big(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}\big) \sim AN(M^{-1}(\xi)b(\xi, g), M^{-1}(\xi)), \qquad (6)$$

then

$$MSE(g, \xi) = M^{-1}(\xi)\big(b(\xi, g)b'(\xi, g) + M(\xi)\big)M^{-1}(\xi),$$

where $M(\xi) = \lim_{n\to\infty} \frac{1}{n} E_{(2)}\left[-\sum_{j=1}^n \frac{\partial^2 LL(\boldsymbol{\beta}; t, x_j)}{\partial \boldsymbol{\beta}^2} |_{\boldsymbol{\beta_0}}\right] =$

$$= \sum_{i=1}^n \omega_i \cdot I(\boldsymbol{\beta_0}; t, x_i) \begin{pmatrix} 1 & x_i \\ x_i & x_i x_i' \end{pmatrix} \textit{ with}$$

$$I(\boldsymbol{\beta_0}; t, x_i) = \frac{1}{\sigma^2}\left([f_\varepsilon'(z)]_{-\infty} + \frac{f_\varepsilon(z_d)^2}{S_\varepsilon(z_d)} + \int_{-\infty}^{z_d} \frac{f_\varepsilon'(z)^2}{f_\varepsilon(z)} dz\right),$$

and

$$\tilde{b}(\xi, g) = \lim_{n\to\infty} \frac{1}{\sqrt{n}} E_{(2)}\left[\sqrt{n}\sum_{j=1}^n \frac{\partial^2 LL(\boldsymbol{\beta}; t, x_j)}{\partial \boldsymbol{\beta}^2} |_{\boldsymbol{\beta_0}}\right] =$$

$$= \frac{1}{\sigma^2\sqrt{n}} \sum_{i=1}^n \omega_i g(x_i) \left[\int_{-\infty}^{z_d} \frac{f_\varepsilon'(z)^2}{f_\varepsilon(z)} dz - \frac{f_\varepsilon(z_d)^2}{S_\varepsilon(z_d)} \cdot ([f_\varepsilon(z)]_\infty - f_\varepsilon(z_d))\right] \begin{pmatrix} 1 \\ x_i \end{pmatrix} := \frac{1}{\sigma^2\sqrt{n}} b(\xi, g).$$

3.1 Log-logistic AFT model

Let us assume that the variable time-to-event, *T*, follows a Log-Logistic distribution. Then, ε follows a Logistic distribution so that $f_\varepsilon(z) = \frac{e^z}{(1+e^z)^2}$ and $S_\varepsilon(z) = \frac{1}{1+e^z}$ with $-\infty < z < \infty$.

Functions involved in (6) for this model are

$I(\beta_0; t, x_i) = \frac{1}{3\sigma^2}(1 - S_\varepsilon(z_d)^3)$ and $\tilde{b}(\xi, g) = \frac{1}{3\sigma^2\sqrt{n}} \sum_{i=1}^n \omega_i g(x_i)[1 - S_\varepsilon(z_d)^3] \begin{pmatrix} 1 \\ x_i \end{pmatrix}$.

Case 1.

Let $\xi = \begin{Bmatrix} 0 & 1 \\ \omega & 1-\omega \end{Bmatrix}$ be the design to be sought. Let us consider the class of contamination function

$$\mathcal{F} = \left\{g \colon \max_{x \in \{0,1\}} |g(x)| \le c \textit{ and } g(0) = 0\right\},$$

being *c* a constant, and common variance $\sigma = 1$. Then,

Theorem 2. *The minimax D-optimal robust design (5) on* $\chi = \{0,1\}$ *allocates a proportion of*

$$\omega^* = \frac{A(\theta, \mu, c, d) - \sqrt{A(\theta, \mu, c, d)B(\theta, \mu, c, d)}}{A(\theta, \mu, c, d) - B(\theta, \mu, c, d)}$$

observations at point $x = 0$*, where*

$$A(\theta, \mu, c, d) = (c^2 + 3)\big(e^{(\theta+\mu)} + d\big)^3 - c^2 e^{3(\theta+\mu)},$$
$$B(\theta, \mu, c, d) = 3\sigma^2\big(e^{(\theta+\mu)} + d\big)^3.$$

Case 2.

Let ξ be a continuous design and $m(\boldsymbol{x})$ the density of ξ, $\boldsymbol{x} \in \chi$. Consider the continuous version of the matrices and functions involved in (6). Let $h(\boldsymbol{x}) = (1, \boldsymbol{x})$ be the regressor vector, $a(\boldsymbol{x}) = 1 - S_\varepsilon(z_d)^3$ and define

$$K(\xi) = \int_\chi a^2(\boldsymbol{x}) m^2(\boldsymbol{x}) h(\boldsymbol{x}) h'(\boldsymbol{x}) d\boldsymbol{x},$$
$$G(\xi) = K(\xi) - M(\xi)A^{-1}M(\xi).$$

We denote

$$\mathcal{F} = \left\{g \colon \int_\chi h(\boldsymbol{x}) g(\boldsymbol{x}) d\boldsymbol{x} = 0, \int_\chi g^2(\boldsymbol{x}) d\boldsymbol{x} \le c^2 < \infty, g(0) = 0\right\}$$

the class of contamination functions for c a positive constant. Then,

Theorem 3*. The maximum of (5) on* $\mathcal{F}$ *as above is*

$$max_{g\in\mathcal{F}}\ \mathcal{L}_D(g, \xi)$$
$$= (1-\nu)\left(\frac{1 + \frac{\nu}{1-\nu}\lambda_{max}[M^{-1}(\xi)G(\xi)]}{|M(\xi)|}\right)^{1/p},$$

where $\nu \in [0,1]$ *is a parameter representing the relative importance, to the experimenter, of the error due to bias rather than to variance.*

4. REFERENCES

[1] Huber, "Robustness and designs", *In J.N. Srivastava Ed., A survey of Statistival Design and Linear Models*, 1975, pp. 287-303.

[2] M.J. Rivas-López, J. López Fidalgo, and R. Del Campo, "Optimal experimental designs for accelerated failure time with Type I and random censoring", Biometrical Journal, vo. 56, no. 5, pp. 819-837, 2014.

[3] X. Xu, "Robust prediction and extrapolation designs for censored data", Journal of Statistical Planning and Inference, vo. 139, no. 2, pp. 486-502, 2009.

Acknowledgements: This work was funded by the Ministerio de Economía, Industria y Competitividad (Project MTM2016-80539-C2-1-R), Junta de Castilla y León (Project SA080P17), the Consejería de Educación, Cultura y Deportes of Junta de Comunidades de Castilla-La Mancha (Project SBPLY/000380), and the European Regional Development Fund (FONDOS FEDER).

StEering
Inter-University Research Center:
Design, Quality and Reliability

Computing exact and approximate D-optimal designs for electrical impedance measurements

R. Martín Martín[1*], J. L. Polo Sanz[2], À. Sebastià Bargues[1]

[1)] Department of Mathematics, University of Castilla-La Mancha, Avda. Carlos III s\n, 45071, Toledo, Spain.
[2)] Escuela de Ingeniería Industrial y Aeroespacial de Toledo, University of Castilla-La Mancha, Avda. Carlos III s\n, 45071, Toledo, Spain.

[*]Corresponding author - e-mail address: Raul.MMartin@uclm.es

Keywords: Optimal Experimental Design, Impedance , Algorithm

ABSTRACT - The electrical behaviour of a great variety of systems can be characterized using impedance measurements [1,2], that is, the opposition that a specific system under study exhibits to the flow of electric current as a function of frequency. Figure 1 shows a simple circuit model in which the parallel combination of the capacitor *C* and the resistance R_1 describes the impedance of systems such as the simplest form of a corroding electrode interface or a biological cell membrane. We formulate the problem of estimating parameters in the linear single-input single-output (SISO) system in frequency domain using the theory of optimal experimental design analysis. In order to obtain the optimal input signals for parameter estimation, we compute approximate and exact optimal designs optimizing the determinant of the information matrix by adapting two of the most algorithms (randomized exchange algorithm REX, and KL-exchange algorithm) that are routinely used nowadays.

1. INTRODUCTION

Let us consider a sinusoidal input current *i*(*t*) to the electrical circuit shown in Figure 1. The resistance R_2 will carry all the current. Depending on the angular frequency ($\omega=2\pi f$, where *f* is frequency), the current *i*(*t*) will divide between the capacitor *C* and the resistance R_1. Let us also define the voltage *v*(*t*) at the terminals of the circuit as the output signal.

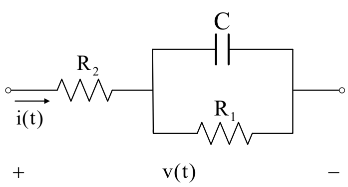

Figure. 1. Equivalent circuit model.

The linear time-invariant (LTI) circuit shown in Figure 1 can be described by the ordinary differential equation:

$$R_1 C \frac{dv(t)}{dt} + v(t) = R_1 R_2 C \frac{di(t)}{dt} + (R_1 + R_2) i(t) \qquad (1)$$

with zero initial conditions and $t \in [0, T_E]$. Let us define $\theta = (a, b, c)$ for $a, b, c \in \mathbb{R}$ the unknown vector parameters with $a = R_1 R_2 C$, $b = R_1 + R_2$ and $c = R_1 C$. The solution $v(t;\, \theta)$ of (1) depends of the unknown parameters. The vector θ is to be estimated from measurements of the form:

$$V(t) = v(t,\theta) + \varepsilon(t), \qquad t \in [0,, T_E] \qquad (2)$$

where the errors $\varepsilon(t)$ are i.i.d. $N(0, \sigma^2)$. It can be shown that the Fisher information matrix has the form:

$$M_t(i,\theta) = \int_0^{T_E} \nabla_\theta V(t,\theta)|_{\theta=\theta_0} (\nabla_\theta\, V(t,\theta))^T|_{\theta=\theta_0} dt.$$

where T is the transposition.

2. METHODOLOGY

Because the input current is a scalar stationary function, using Parseval's theorem [3] we can express the Fisher Information Matrix as:

$$M(S_i(\omega), \theta) = 2 \int_0^\infty \Re \left\{ \left(\frac{\partial Z(\omega,\theta)}{\partial \theta} \right)\bigg|_{\theta=\theta_0} \left(\frac{\partial Z(\omega,\theta)}{\partial \theta} \right)^H \bigg|_{\theta=\theta_0} \right\} S_i(\omega) d\omega$$

with $S_i(\omega)$ the normalised power spectral density of *i(t)*, $Z(\omega,\theta) = \frac{j\omega a + b}{j\omega c + 1}$ the transfer function, that is, the driving point impedance function of the electrical equivalent circuit (EEC) of Figure 1, $\Re\{\cdot\}$ is the real part of the expression and the superscript H denotes the conjugate transpose operator. The input spectrum can be confined to some finite interval $[\omega_I, \omega_E]$ where ω_I and ω_E are the lowest and highest frequencies, respectively, in which the impedance has been measured.

Under this scenario, identifying S_i with the design measure, ξ, we are in the same framework as for optimal experimental design theory.

In general, numerical techniques are needed to determine optimal designs, be they exact or approximate. Different techniques have been developed for each. Many of them are based on exchange schemes and have been proposed to construct D-optimal designs. Theoretical foundations for them are provided from the Equivalence Theorem.

Algorithms for the construction of optimal input are *identical* to those. In this work we proposed a modification of the randomized exchange algorithm (REX) for computing D-optimal approximate design of experiments. According to their authors [4], this algorithm outperforms all existing algorithms. With regard to the computing of D-optimal exact design, a modification of KL-exchange algorithm of Atkinson and Donev [5] has been used.

3. RESULTS

The modified algorithms were applied to this kind of circuit (Figure 1). Let $R_1 = 1000$, $R_2 = 10$ and $C = 10^{-5}$ be the nominal values of the parameters and $\omega \in [2\pi 10^{-2}, 2\pi 10^{4}]$. For numerical study, we consider the original frequency design space to be a set of grid points spread equidistantly by 0.01.

For comparison purposes, KL-algorithm was executed considering the same number of support points that the design used for experimenters. They used to measure at a number of discrete points (5 per decade) logarithmically spaced.

Figure 2 illustrates the support points for D-optimal designs obtained after running the algorithms and the design employed by the experimenter. Table 1 displays D-optimal criterion values for the designs obtained from REX, KL and practitioners.

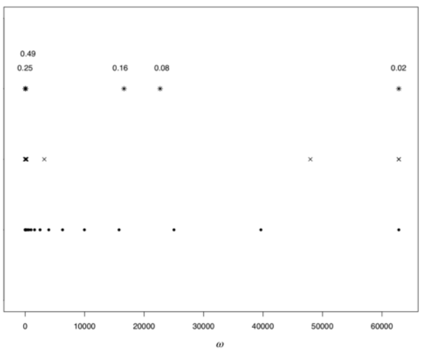

Figure. 2. Support points for D-optimal designs obtained by: (*) REX modified algorithm and associated weights, (x) KL modified algorithm and the experimental design points (·)

Table 1. Criterion values (m = number of parameters)

Design	$\phi_D(\xi) = ([detM(\xi^*)]^{\frac{1}{m}}$	Eff_{ϕ_D}
REX	11603.97	
KL	11603.58	
Experimental	7588.28	0.65

4. CONCLUSIONS

Optimum designs are a valuable tool for measuring the worth of an experimental design, through the efficiency, $Eff_{\phi_D} = \frac{\phi_D(\xi_{exp})}{\phi_D(\xi_D^*)}$. This efficiency is the percentage of observations which the optimum design would need to achieve the accuracy of the experimental design being compared. The result of efficiency of the *traditional design* is presented in Table 1. This shows how the accuracy of parameter estimates can be increased by the use of optimal inputs.

5. ACKNOWLEDGEMENTS

This work was funded by the Ministerio de Economía, Industria y Competitividad (Project MTM2016-80539-C2-1-R), the Consejería de Educación, Cultura y Deportes of Junta de Comunidades de Castilla-La Mancha (Project SBPLY/000380), and the European Regional Development Fund (FONDOS FEDER).

6. REFERENCES

[1] E. Barsoukov and J. R. Macdonald, *Impedance Spectroscopy: Theory, Experiment, and Applications*, 3rd ed. New Jersey: Wiley; 2018.

[2] S. Grimnes and. Ø. G Martinsen, *Bioimpedance and Bioelectricity Basics*, 3rd ed. London: Academic Press; 2015.

[3] Graham C. Goodwin, Robert L. Payne, *Dynamic System Identifications: Experiment design and data analysis*, Mathematics in science and engineering: Volume 136. Academic Press; 1977.

[4] R. Harman, L. Filová and P. Richtárik, "A Randomized Exchange Algorithm for Computing Optimal Approximate Designs of Experiments", *Journal of the American Statistical Association*, DOI:10.1080/01621459.2018.1546588, 2019.

[5] A.C. Atkinson, A.N. Donev and R.D. Tobias, *Optimum Experimental Designs with SAS*. Oxford: University Press; 2007.

Heuristic algorithm to reduce the number of measurements in a quality control

Jose Moler[1*], Blanca Palacios[1]

[1)] Universidad Pública de Navarra. Spain

[*]Corresponding author - e-mail address:jmoler@unavarra.es

Keywords: Regression analysis, quality control

ABSTRACT – In an industrial process aluminum pieces are hung in an industrial frame in groups of 120. The pieces are anodized before the coloring treatment. At present, quality control team measures the whole set of pieces to obtain color and roughness measurements in order to elucidate the validity of the production. The quality team looks for the smallest number of pieces to be controlled and their location in the frame in order to obtain a reliable information about the quality of the whole set of pieces in the frame.

1. INTRODUCTION

An industrial process consists in coloring pieces previously anodized. Pieces are initially hung in groups of 120 in an industrial frame. This frame is submerged in a chemical solution with two main steps: anodizing and coloring process. After drying the pieces, the quality team does two measurements of each piece of the frame:

a) Color: the variables a, b, L are observed. Their absolute value measures the intensity of a color as follows:

   a: >0 red color; <0 green color

   b: >0 yellow color, <0 blue color

   L: >0 intensity of white color, =0 black color

b) Roughness: a micrometer provides the deepness of the surface in the central point of the piece. This variable is denoted as *micraje.*

So that, the final response is (L, a, b, micraje).
Each measurement requires a different measurement device. This procedure is time consuming and the workload avoids to the quality employees the calibration of the devices as frequently as it should be done to guarantee the quality of the measurements.
Managers in the company suspect that a smaller number of measurements would not reduce significantly the information in the quality control and the failure of quality standards would be detected by the quality team with the same success rate as they do now.
To handle this problem the company provides a data set with the measurements of 76 industrial frames. Observe that each frame have 120 pieces and from each piece four measures (a, b, L, micraje) are obtained.

2. Methodology

The frame has two parallel rectangular subframes that we denote B and S. In each subframe we distinguish interior side (I) and exterior side (E). Each side has two parallel rows of hooks, one at the top (U) and the other one at the bottom (D). By combining this three divisions we have 8 zones, for instance, BED is the bottom row (D) in the exterior side (E) of the subframe B.

We follow three steps:

A. Factorial analysis to classify zones by similarity
B. Heuristic algorithm:
   B1. Fix n the number of measurement points
   B2. Consider all the possible groups of size *n* among the 120 positions
   B3· For each group of size *n* and each variable (a, b, L, micraje) we adjust a linear regression model for the measurement in a position out of the group (dependent variable) with respect to the n measurements in the group (regressors).
   B4. In each regression, save R^2 for each variable: R^2 (a), R^2 (b), R^2 (L), R^2 (Micraje)
   B5. Choose the group with the smallest mean R^2.
C. Calibration of the procedure:
   C1. Chose randomly 65 frames to apply B
   C2. Compare the real measures of the 11 remaining frames with the estimates obtained with B.

C3. Calculate U-theil to measure the advantage of using "n" points with respect to "m>n" points

C4. Tests: coherence between the values of the descriptive values of the calibration group and the predicted ones.

3. Results and discussion

Due to the symmetry of the frame, the study is reduced to one of the sides of the frame (B or S) but the results are not good. The reason is that the recipient where the frames are submerged has a division that separates subframes B and S, so small changes between the liquid on both parts of the recipient can affect the color and anodizing process. We will consider both sides of the frame to select the optimal location points.

First we consider n=1. We obtain that the best location is the number 70 for both, color and roughness. The calibration study shows U-theils smaller than 0.11 for color variables and smaller than 0.51 for roughness variable. In table 1 we observe that there is no coherence, in general, between the estimated results and the real values.

Table 1 Calibration for n=1. P-values of the coherence between the global mean (variance and quartiles) and the estimated one.

	mean	variance	P25	P50	P75
a	***	***	***	**	*
b	***	***	0.3908	***	***
L	***	***	***	***	***
Micraje	*	***	***	***	***

***<0.001; **<0.01, *<0.05

The best results are obtained when n=4. In this case, U-theil values are smaller than 0.1 regardless the variable studied. Besides this, p-values of the tests for the coherence between real and estimated values are all over 0.1 (see Table 2). Observe that the variance for the Micraje variable is significantly not coherent. In general, variability tends to be more difficult to reproduce with estimations. Regression estimations are expected values so, smoother values than the real ones are obtained. In Figure 1 box-plots for estimated and real Micraje values in the 8 zones of the frame are shown. Observe again that estimated values are less variable than real ones which have wider boxes than their partner. However, estimated and real values move in the same range regardless the zone and, as quartiles are not significantly different, only extremal values are not reproducible via estimation.

Table 2 Calibration for n=4. P-values of the coherence between the global mean (variance and quartiles) and the estimated one.

	mean	variance	P25	P50	P75
a	0.5495	0.1187	0.9731	0.1370	0.8663
b	0.5298	0.0934	0.7619	0.1370	0.2966
L	0.7546	0.6070	0.6616	0.6618	0.3999
Micraje	0.2416	***	0.6252	0.6830	0.2065

***<0.001; **<0.01, *<0.05

The range of R^2 values obtained to choose location points grows as the number of points n grow, but the improvement is negligible when n is greater than 5.

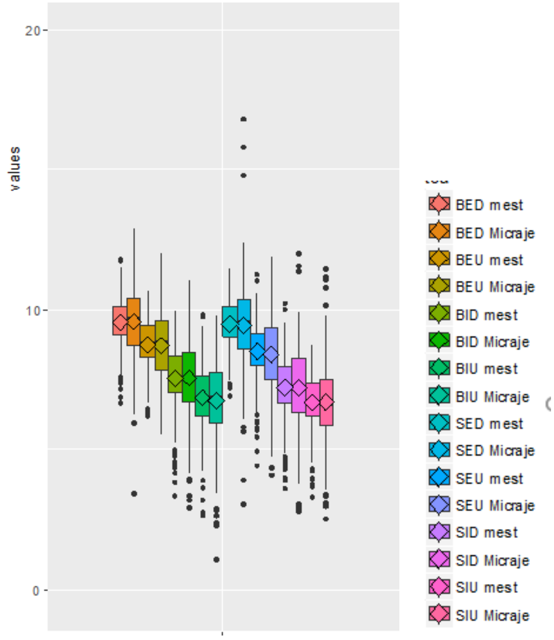

Figure 1 Estimate (mest) and real Micraje values for the 11 frames not used in the regression.

4. Conclusions

This study is a goal of a project which also includes the search, via experimental designs, of the optimal composition of the liquid where anodizing and coloring have place in order to avoid fatal productions.

The solution presented here considers that there is no significant loss of information with the measurement in 4 points for the color variables and 4 different points for the roughness instead that measuring all the pieces in the frame. This procedure not only will reduce the time in the quality control phase but also will allow the calibration of sensible devices within measurements. This will give a more reliable information to validate the whole set of pieces in each frame.

References

[1] Serebrennikova, I., Vanýsek, P. & Birss, V. I. Characterization of porous aluminum oxide films by metal electrodeposition. *Electrochim. Acta* **42,** 145–151 1997.

[2] J.L. Mo, Z.G. Wang, G.X. Chen, T.M. Shao, M.H. Zhu, and Z.R. Zhou, "The effect of groove-textured surface on friction and wear and friction-induced vibration and noise," *Wear*, vol. 301, no. 1–2, pp. 671–681, 2013.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Restricted experimental designs to optimize quality standards in an electro-color process

M. Blanca Palacios[1*], José Moler[1]

[1)] Universidad Pública de Navarra. Spain

[*]Corresponding author - e-mail address: blanca.palacios@unavarra.es

Keywords: Experimental design, quality control

ABSTRACT – An industrial process consists in coloring aluminum pieces, which have been previously anodized. The process submerges an industrial frame with 120 pieces in different recipients that contain specific chemical solutions. We focus on the chemical solutions in the recipient where the anodizing is carried out and where the coloring of the pieces is done. The main goal is to design an experiment, without stopping the industrial process, in order to establish the influence of the chemical components in the final color of the piece.

1. INTRODUCTION

The final color of an aluminum piece is the main quality standard in an industrial process. The quality team measures in each piece 4 variables denoted by a, b, L, Micraje.

a: >0 intensity of red color; <0 intensity of green color

b: >0 intensity of yellow color, <0 intensity of blue color

L: >0 intensity of white color, =0 black color

Micraje: roughness of the anodizing layer in microns.

The higher the absolute value of a, b and L, the more intensity of color.

120 aluminum pieces are set in an industrial frame. There are two main steps in the industrial process, which are the immersion of the frame in the anodizing and Electro-color containers, respectively. The chemical engineers consider several factors in each container that influence the values of the observed variables a, b, L and Micraje.

Anodizing container: level of sulfuric acid (AN_sulf), level of aluminum (AN_al), amperage (AN_Amp), voltage (AN_V).

Electro-color container: level of sulfuric acid (COL_sulf), level of tin (COL_sn), direct current (COL_VCont), alternating current (COL_Valt).

Besides this, temperature of the liquid and time with electricity in the anodizing or electro-color containers are fixed. These factors are fixed during the experimentation because any change would destroy the pieces in the frame, that is, would be a fatal production. Moreover, the values of the controlled factors in the anodizing container and electro-color container move in a range fixed beforehand in order to avoid fatal productions, this is called the security range.

Unfortunately, fatal productions happen although controlled factors move in the security range. This is why the design of a set of experiments that combine the controlled factors when they take extremal values in the security ranges can give information to avoid fatal production and, then, to update the security ranges.

2. Methodology

The company produces three finished products (different final colors in the aluminum pieces) that are called N, V and B which are ordered from less to more darkness of the color. 61 experiments have been designed and, so, 61 frames have been treated following the specifications of each experiment. Finally, the values of a, b, L and Micraje for the 120 pieces in each frame have been collected in a data set.

The methodology followed in this work is as follows:

A. State of the art.

B. Linear Discriminant Analysis: this procedure allows to know which controlled factors characterize finished products V, N and B.

C. Full and fractional factorial designs

C1. Design of the set of experiments: for each frame, we consider all the 2^k combinations of extreme values in the security range, where k is the number of controlled factors in the container.

C2. Search of full 2^k factorial designs: k=3; k=4 and study of interactions and main effects.

C3. Search for fractional design in the set of performed experiments.

3. **Results and discussion**

The effects of the anodizing process in the color of pieces have been deeply studied in the literature. In [1], [2] and [3] many interesting relationships between controlled factors and the response variables are established and we summarize them in Table 1. For instance, we observe that larger values of sulfuric acid in the anodizing container will produce darker pieces and larger diameter porous. It is also observed that the Micraje is mainly influenced by the factor in the anodizing container.

Table 1. Effects in color and roughness of larger values of sulfuric acid and voltage in the anodizing and electro-color containers

	Color	Roughness
+ AN_sulf	+ dark	+ porous diameter
+ AN_V	+ dark	+ barrier layer, + porous diameter, -porous density.
+ COL_sulf	- dark	
COL_Valt/Vcont	15-16 optimal	Valt>Vcont + porous diameter

In Figure 1, we observe a quite schematic representation of the discriminant analysis.

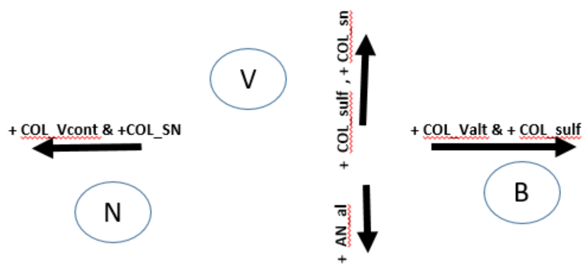

Figure 1. Linear discriminant analysis, classification of the three finished pieces depending on the controlled factor values.

Finally, an exhaustive study of all the full 2^k factorial designs provides the following results:

a) AN_al: the presence of this variable is the result of the anodizing procedure and only indicates the use of the same liquid in the container and does not have influence in the response variables.
b) AN_sulf: does not have an important influence in the process.
c) AN_V and AN_amp: they are really important in the roughness variable. In fact, each unity more in voltage produces an expected raise of the roughness variable of 0.34μ.

On the other hand, we have that the values of the controlled factors in the electro-color container have influence in the response color variables but not in the roughness variable. Table 2 shows the expected changes in the color response variables, when one more unit of the controlled factor is added.

Table 2. Expected changes in the response color variables for each unity that the corresponding controlled factor changes.

	a	b	L
COL_Sn		*0.43***	*-0.34***
COL_Sulf	*-0.14**	*-0.28***	*0.51***
COL_Vcon	*-0.32**	*-0.28***	

p-values: *: <0.05; **:<0.001

4. **Conclusions**

The theoretical results in Table 1 were helpful for the chemical engineers because they allow them to standardize temperature, time and electrical current to control darkness in the finished piece.

Even though a design of experiments without restrictions in the controlled factor value range had been more informative, several conclusions were obtained:

4.1 Aluminum is not an influential factor in the roughness and the color of the pieces.
4.2 When Col_sulf decreases, darker colors are obtained (a and L decrease).
4.3 When Col_Sn increases creases, the yellow intensity grows (b is larger).
4.4 Controlled factors in the anodizing recipient affect only the roughness; in particular, more voltage produces more roughness (diameter of porous).
4.5 Dark pieces only require much more time of immersion in anodizing and electro-color recipients, provided that controlled factors move on the security range.

References

[1] T. Sato & S. Sakai. Electrolytic Colouring of Anodized Aluminum with Nickel Sulphate, *Transactions of the IMF*, 57:1, 43-47, 1979

[2] H.J. Gohausen and G.C. Schoener. Electrolitic coloring of anodized aluminum using tin electrolytes. Plating and Surface Finishing, 71, 56-63, 1984.

[3] J.W. Diggle, T.C. Downie y C.W. Goulding. Anodic oxide films on aluminum. Chemical Reviews, 69, 3, 365-405, 1968.

StEering
Inter-University Research Center:
Design, Quality and Reliability

Empirical variograms

Giovanni Pistone

de Castro Statistics, Collegio Carlo Alberto, Piazza Vincenzo Arbarello 8, 10122 Torino, Italy

email: giovanni.pistone@carloalberto.org
home page: https://www.giannidiorestino.it

Keywords: Engineering Statistics, Gini Methods, Probability, Simulation, Variogram.

ABSTRACT – In this paper, I discuss some methodological issues raised by a non parametric-method for assessing the technological characteristics of a surface. The method was originally suggested by Grazia Vicario (DISMA Politecnico di Torino) on the occasion of a research project in metrology.

1. INTRODUCTION

Consider a response surface on a given domain (usually a rectangle). Some level measurement are available at given testing points. We are interested in assessing the conformity of the shape of the response surface with respect to a given norm. For example: “is the surface bended in some direction?” Or: “Is there a bump in the middle?” Or: “Is there a waviness of a type associate to a specific technology?” These are possible defects that cannot be specified in a parametric way. Hence, usual parametric statistical methods are not usable. See Vicario [7], Vicario *et al*. [8].

A very popular modeling method relies on the assumption that the surface under study is the realization of a random field, for example, a Gaussian random field. In such a case, the observed characteristics of the surface will in fact depend on the auto-covariance of the random field. More specifically, under the stationarity assumptions that the mean value is constant and the covariance depends on the distance between the test points only, the characteristic of the surface are given by the random field *variogram.* The variogram is a function defined on possible distances between two test point in the domain. Its value equals one half the expected squared norm of the difference of response values at two test points. This approach is sometimes called Kriging method or Matheron method and can be used both in a parametric or non-parametric approach. See Cressie [1], Matheron [3], Pistone and Vicario [4].

G. Vicario has suggested in [9] to use the variogram as a non-parametric method, without assuming any randomness of the surface, but using instead the idea of a sistematic or random sampling of couples of test points on the given domain. The rational of this approach will be based on general principles of the theory of statistical designed experiments.

The present paper discusses this idea as an application of a Gini methodology in the sense of Yitzhaki, and Schechtman [10]. Only toy examples are considered.

2. METHODOLOGY

Consider a (deterministic) real function f on a multivariate real domain D and two independent equally distributed random variables X and Y with values in the domain D . Half the expected value of the squared difference between $f(X)$ and $f(Y)$ equals the common variance of both responses. This is a very crude index of the shape of f. For example, if X and Y are uniform on the unit interval, then the index is the squared quadratic norm of f.

A better estimate would be obtained by considering the conditional expected value of half of the squared difference between $f(X)$ and $f(Y)$, given the distance between X and Y. Such a conditional expected value will be a random variable which is a function of the random distance $d(X,Y)$. We will call such a function the *empirical variogram,* or, simply, *varioram,* of the response surface.

In this modelling, the empirical variogram is an expectation computed in the conditional distribution of the couple (X,Y) with respect ot the value of the distance between X and Y. The computation of such conditional distribution is feasible in closed form in some very simple cases. More complex cases can be dealt with by using numerical methods.

The interest of the activity described above is the possibility to inspect in such cases the shape of the empirical variogram of typical shapes of the response function. The hope is to became able to assess if some characteristics of interest of the response surface *f* can be revealed this way.

The simplest application of this set-up arises in the case were the random variables X and Y assume all the combinations of points of a uniform grid on a rectangular domain. And the distance is computed along the grid. In this case, the procedure described above reduces to the standard non-parametric estimation of the variogram due to Matheron, see [3]. However, when the number of available measured points on the grid is very high, it is interesting to consider a sampling of such points. This leads naturally to the consideration of our proposal.

The sampling version of the methodology requires the use of methods for the sampling of conditional distribution, which represents a delicate issue, see Diaconis *et al*. [2]. In turn, the derivation of a-priory bounds on the precision of the sample estimate of the empirical variogram requires technical assumptions on the random variables X and Y. Such methods are described, for example, in Vershynin [6] and Wainwright [11].

A final remark concerns the assumption of independence between the two random trial points X and Y. Other types of joint distribution are conceivable and some could be useful. From this point of view, it is interesting to remark that lower and upper bounds on the conditional variance we are computing are available from methods in Optimal Transport theory, see Santambrogio [5].

3. CONCLUSIONS

We have performed some preliminary computation regarding a possible way to evaluate characteristic of interest of technological shapes. Notice that we are not modeling here the measurement errors. In fact we are assuming that the precision of measurement and the number of points that can be tested are both very high so that all classical concerns are not of interest. In fact, some modern measurement technologies are so quick and precise that experimenter is concerned by the required computing time and by the risk of overfitting more then by the risk of measurement errors. See some examples in Vicario [7] and Vicario and Pistone [9].

ACKNOWLEDGMENTS

The author was supported by de Castro Statistics and Collegio Carlo Alberto. He is a member of INdAM-GNAMPA.

REFERENCES

[1] N.A.C. Cressie, *Statistics for spatial data*, New York: John Wiley & Son 1993. Revised reprint of the 1991 edition.

[2] P. Diaconis, S. Holmes, M. Shahshahani, "Sampling from a manifold," in *Advances in modern statistical theory and applications: a Festschrift in honor of Morris L. Eaton*, Beachwood, OH: Inst. Math. Statist. 2013, pp. 102--125.

[3] G. Matheron, *Traité de géostatistique appliqué*, Mem. Bur. Rech.e Geog. Minieres, no. 14, Editions Technip 1962

[4] G. Pistone, G. Vicario. “A note on Semivariogram.” In *Topics on Methodological and Applied Statistical Inference* (T. Di Battista, E. Moreno, W. Racugno Eds), Springer 2016, pp. 181-190.

[5] F. Santambrogio, *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling,* Birkhäuser 2015.

[6] R. Vershynin, *High-dimensional probability: an introduction with applications in data science*, Cambridge: Cambridge University Press 2018.

[7] G. Vicario. “Inference On Errors In Industrial Parts: Kriging And Variograms Versus Geometrical Product Specifications (GPS) Standard,” *This Conference,* 2020.

[8] G. Vicario, G. Craparotta, and G. Pistone, “Meta-models in computer experiments: Kriging vs artificial neural networks,” *Quality and Reliability Engineering International*, vol. 32, pp. 2055–2065, 2016.

[9] G. Vicario, G. Pistone, “Simulated variogram-based error inspection of manufactured parts," *Statistical Papers*, vol. 59, 1411–1423, 2018.

[10] S. Yitzhaki, E. Schechtman, *The Gini Methodology: A Primer on a Statistical Methodology*, Springer 2013

[11] M.J. Wainwright, *High-dimensional statistics: a non-asymptotic viewpoint*, Cambridge: Cambridge University Press 2019

StEering
Inter-University Research Center:
Design, Quality and Reliability

StEering
Inter-University Research Center:
Design, Quality and Reliability

An Iterative Method for Tuning Complex Simulation Code

Yonggwan Shin[1], Juyoung Hong[1], Yun Am Seo[2], Youngsaeng Lee[3],
Jeong-Soo Park[1,*]

[1)] Department of Statistics, Chonnam National University, Gwangju 500-757, Korea.
[2)] AI Weather Forecast Research Team, National Institute of Meteorological Science (NIMS), Korea.
[3)] Digital Transformation Department, Korea Electric Power Corporation, Korea.

[*]Corresponding author - e-mail address: jspark@jnu.ac.kr

Keywords: Gaussian process; Calibration; Computer experiments;

ABSTRACT – Tuning a complex simulation code is the process of improving the agreement of a code calculation with respect to a set of experimental data by adjusting parameters implemented in the code. For this problem, the approximated nonlinear least squares (ANLS) method based on Gaussian process metamodel has been used by some researchers. A potential drawback of the ANLS method is that the metamodel is built only once and not updated thereafter, and thus, the computer data are no longer used. To address this difficulty, in this study, an iterative algorithm is proposed, in which the parameters of the simulation code and Gaussian process metamodel are alternatively re-estimated and updated by the maximum likelihood estimation and the ANLS method.

1. INTRODUCTION

Modern computer simulation codes have various unknown parameters. Assuming the validity of the simulation code, we can adjust or estimate such parameters using the nonlinear least squares estimation (NLSE) method. This procedure is called calibration or code tuning. It is defined as the process of improving the agreement of a code calculation or set of code calculations with respect to a chosen and fixed set of experimental data by adjusting parameters implemented in the code[1,2]. Han et al. [3] differentiated the tuning parameter and calibration parameters; however, in this study, the two parameters are treated as the same (henceforth, tuning parameter).

2. METHODOLOGY

2.1 Gaussian process model as a surrogate

For the metamodel of simulation code, we use a Gaussian process (GP) model or a spatial regression model that treats response $y(\underline{x})$ as a realization of a random function superimposed on a regression model:

$$Y(\underline{x}) = \sum_{j=1}^{k} \beta_j f_j(\underline{x}) + Z(\underline{x}) + \epsilon, \quad (1)$$

where f s are known functions and β s are unknown regression coefficients. Here, random process $Z(.)$ is assumed to be a Gaussian process with mean zero and covariance

$$cov(\underline{t},\underline{u}) = V(\underline{t},\underline{u}) = \sigma^2 R(\underline{t},\underline{u}), \quad (2)$$

between $Z(\underline{t})$ and $Z(\underline{u})$ for $\underline{t} = (t_1,\dots,t_d), \underline{u} = (u_1,\dots,u_d)$, where σ^2 is the process variance and $R(\underline{t},\underline{u})$ is the correlation function. Some possible choices for the correlation function are taken from the Gaussian correlations denoted by

$$R(\underline{t},\underline{u}) = exp[-\theta \sum_{i=1}^{d} |t_i - u_i|^2], \quad (3)$$

where $\theta \geq 0$. One may consider a different version of (3) by taking several θ value as follows:

$$R(\underline{t},\underline{u}) = exp[-\sum_{i=1}^{d} \theta_i |t_i - u_i|^2], \quad (4)$$

where $\theta_i \geq 0$, for $i = 1,2,\dots,d$. Some combinations of βs and θs determine the model, but the following simple GP models are considered at first:

$$y(\underline{x}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_d x_d + Z(\underline{x}) + \epsilon, \quad (5)$$

with the correlation function (3) or (4). When a *common* θ of (3) is used, we call (5) as "Model 1" in this study. When the *several* θ_i's of (4) are used, we call (5) as "Model 2".

2.1 Max-min algorithm

Following are the steps for the proposed tuning method. We call it a Max-min algorithm because it uses maximization and minimization iteratively.

Algorithm Max-min: iterate Step 3 and Step 4 until convergence.

Step 1 (model building): build a surrogate (5) using the MLE for the given *computer data only*.

Step 2 (initial solution): set iteration $i = 1$, and find parameters by minimizing RSS_p using a surrogate (5). RSS_p is the residual sum of squares with predictor.

Step 3 (maximization): build a new surrogate (5), using the MLE for the *combined data* with the fixed parameters obtained in the previous step.

Step 4 (minimization): set iteration $i = i + 1$, and find parameters by minimizing RSS_p using the surrogate built in Step 3. If parameters satisfies the stopping rule, then stop; otherwise, go to Step 3.

2.1 Application to a nuclear fusion model

A simple measure of energy efficiency in a nuclear fusion device is the global energy confinement time, τ_E. The theoretically based confinement model can be written as follows (Kay and Goldston [4]):

$$y_E = f(\tau, P, I, N, B), \qquad (6)$$

where f is a known function calculated using a complex simulation code called Baldur; P is the total input power; I is the plasma current; N is the electron density; B is the magnetic field; and $\tau = (\tau_1, \tau_2, \tau_3, \tau_4)$ are the following adjustable parameters determining energy transfer, that is, drift waves, rippling, resistive ballooning, and the critical value of η_i, respectively.

Note that the experimental data consists of only P, I, N, B and the real observation y_E, whereas the computer data comprises eight independent variables ($\underline{T}$, P, I, N, B) and computer response y_C obtained using Baldur. The experimental data were drawn from the database of S. Kaye: 42 observations from the PDX (Poloidal Divertor Experiment) tokamak in Princeton and 64 and from the Baldur simulator (Singer et al. [5]). Because the Baldur simulator needs five CPU minutes on a Cray supercomputer for one execution, a careful selection of input points is required, which is statistical design problem for a complex simulation code.

3. RESULT and DISCUSSION

Table 1 provides the result of τ estimation by using ANLS, SMLE, and Max-min algorithm on the basis of Model 1. The last column in Table 1 shows the value of RSS_p at the convergence of the algorithms. RSS_p for SMLE is obtained by the calculation of RSS_p for the estimated τ. Note that RSS_p for the Max-min with Model 1 is the smallest among three methods. Figure 1 show the residual plot of the tuning parameters for the Tokamak data, which were obtained by the Max-min algorithm with Model 1. These results indicate that we are fitting the computer data relatively well, and that we are getting good coverage of the range of experimental observations.

Max-min algorithm requires more computing time than the ANLS method. This is because Max-min uses the combined data, resulting in larger correlation matrix that must be inverted. It also needs more iterative minimizations of RSS_p compared to just one time RSS_p minimization in the ANLS method.

Table 1 Tuning results from nuclear fusion example where $\hat{\tau}$ are estimates of tuning parameters.

Method	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	RSS_p
ANLS	1.012	2.035	1.110	1.308	0.0015
SMLE	1.120	2.055	0.118	1.303	0.0010
Max-min	0.667	1.053	0.477	1.823	0.0007

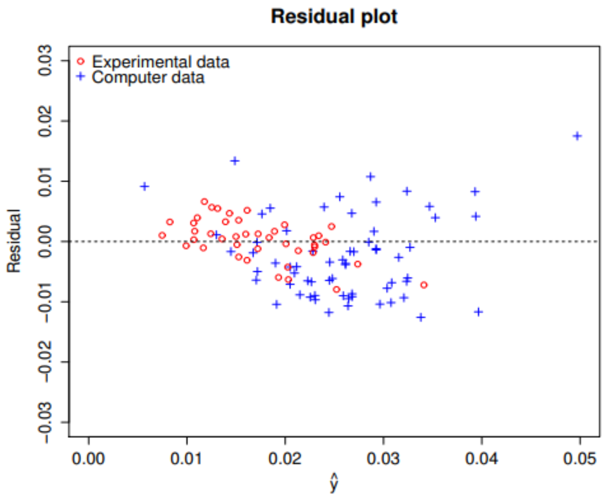

Figure 1 Residual plot for nuclear fusion example in which the computer code was tuned by Max-min algorithm a Gaussian Process Model

There are basic limitations related to the experimental design to tune computer code to data. We found that the performance of tuning methods is much dependent on the design for computer experiments. Further research on relevant designs under sequential tuning approach must be helpful.

4. REFERENCES

[1] D.D Cox, J.S. Park, C.E Singer, "A statistical method for tuning a computer code to a data base," *Computational Statistics & Data Analysis*, vol. 37, no. 1, pp. 77-92, 2001.

[2] F. Cailliez, A. Bourasseau, and P. Pernot, "Calibration of Forcefileds for Molecular Simulation: Sequential design of Computer Experiments for Building Cost-Efficient kriging Metamodels," *Journal of Computational Chemistry*, vol. 301, no. 1–2, pp. 671–681, 2013.

[3] G. Han, T.J. Santner, J.J. Rawlinson, "Simultaneous determination of tuning and calibration parameters for computer experiments," *Technometrics*, Vol. 51, no. 4, pp. 464-474, 2009.

[4] S.M. Kaye, G.C. Goldston, "Global energy confinement scaling for neutral-beam-heated tokama," *Nuclear Fusion*, Vol.25, no. 1, pp. 65-69, 1985.

[5] C. Singer, D. Post, D. Mikkelsen, M. Redi, A. Mckenney, A. Silverman, F.G.P. Seidl, P.H. Rutherford, R.J. Hawryluk, W.D. Langer, L. Foote, D.B Heifetz, W.A. Houlberg, M.H. Hughes, R.V. Jensen, G. Lister, J. Ogden, "BALDUR: A one-dimensional plasma transport code," *Computer Physics Communications*, Vol. 49, no. 2, pp. 275-398, 2018.

[6] T.J. Santner, B.J. Williams, W.I. Notz, *The Design and Analysis of Computer Experiments*, 2nd ed. Springer-Verlag New York; 2018.

A Study on Brand Construction Strategies of World-renown Brands and the Inspirations

[1)] Jiazhen,Yang*.
[2)] Yanfeng, Li.
[3)] Wenlong, Tang.
Shandong Technology and Business University Institute of Confucian Business, Yantai, Shandong, China

*Jiazhen, Yang - e-mail address: yangjz3868@126.com
Keywords: International brands; Brand construction; Inspirations

Abstract –This study investigates the successful cases in brand construction of six world-renown corporations. The result indicates that, despite the fact of differing in trade, products, industrial scale, and company history, the same ideological pursuits are shared in management philosophy, quality awareness, technological innovation, market cultivation and employee devotion. From this study it concludes that (1) Brand awareness and quality priority play fundamental roles; (2) The key of brand construction is to set a clear brand orientation by conducting thorough market investigation; (3) The core is to enable brand localization and keep meeting customers' satisfaction; (4) Identifying market orientation and conducting continuous reform and innovation are the sustaining momentum.

- **Introduction**

As the worldwide economy is still struggling with a slow recovering, and the economic globalization is facing new challenges and opportunities. Countries and industries are in keen competition to seize a more important position and bigger share in the international market, and to get stronger influence and a more powerful voice, with an ultimate purpose of advancing their own domestic economy and prosperity. In order to achieve the goal of enhancing market position through improving quality of products and service, and to facilitate domestic and global economic development, it is crucial to understand and strengthen the efforts of brand construct for building world-known brands. This study aims to identify an effective strategy of brand construction by way of analyzing the successful cases of world-renown brands of different trades with different business models from the diversified international market. Results from this study might serve as a valuable reference to inspire and benefit the companies that have plans in their perspective business to build their own brands of products and services, and enable them to globalize these brand of products and services.

- **Methodology**

This research studies the cases of six international brands: Starbucks, Siemens, KFC, IBM, P&G and Swiss watches for their strategies and the management of brand construction.

Three key methods have been used in this study: case analyses, analytical comparisons and logical deduction.

- **Results and Discussions**

For survey purpose, the top 10 countries with most top 500 world brands from 2017 to 2019 are shown in Table 1**. USA continues to lead the pack, followed by traditional European developed countries such as Britain and France, and Japan, China, Germany, Switzerland and Italy are catching up closely.

Also for the same survey purpose, the top 10 industries with most of 500 world brands from 2017 to 2019 are shown in Table 2**. The top three perennial leading industries, i.e. automobile and parts, food and beverage, and communication media, show a decreasing or saturating trend, whereas retails, computers and communications and telecommunications show an upward trend.

Results of the analyses of the selected six brands are summarized in the following:

1. Starbucks—To build the brand from the scratch.

In the last 40 years, Starbucks adopted an effective and goal-orientated brand strategy to upgrade the once unknown brand to an international famous one. The success of Starbucks lies in the practice of setting its brand position before drawing up detailed and targeted strategic plan of product development, store design, management models, marketing strategies, supplying and franchise etc. which make sure Starbucks' shooting success.

2. Siemens — To preserve in the spirit of craftsmanship.

The spirit of artisan requires Siemens to strive for excellence and upholds its established reputation in the era of "German industry 4.0".

3. KFC in China—To achieve brand localization

The strategy of brand localization adopted by KFC is the key factor in its maintaining an excellent growth momentum in China and enjoying the leading position in fast food industry. The success of KFC inspires famous brands to follow the local economic situation, culture and lifestyle to cultivate an international market and to develop new products that suit local customers.

4. IBM—To seize the opportunity of transformation in trend.

Established brands are unavoidable in the face of new challenges to maintain their leading position in trade. It is crucial for them to research markets in order to grasp customers' tastes and waves of industrial development. It requires great courage to achieve a prompt transformation by way of redirecting market position of current products and taking decisive measures to cut spare business lines.

5. Swiss Watches—To plunge itself in innovation

After four years' restructuring, a new brand — Swatch emerged. In no more than 30 years, the newly emerged brand was crowned with the title of "modern antique". The market orientation targets at young generations aged 18-30 with massages of "fashionable, exciting, tasty, commemorative, high quality, reasonable price", which boosts the impact of well-crafted brand design.

6. P&G — To stick to the brand differentiation strategy

P&G adopts strategy of "one brand, more varieties" to cultivate a variety of markets from the aspects of function, price and package to meet customers' expectation of different levels and to foster a brand preference, customer loyalty thus encouraged and enhanced.

From these analyses, at least three points have become obvious in the course of successful brand name construction:

- The sheer quantity of globally accepted brand names indicates the standard and level of economic and social development of the country and her advancement of civilization;
- The pursuit of brand name construction promotes the product and service quality of the industry and its overall competitiveness;
- There are patterns and regularities in the pursuit of brand name construction, the analyses of the cases chosen in this study have revealed these regularities in their successful practices.

• Conclusions

The survey and case analyses in this study indicate that the long journey of brand construction of world brands, established or newly emerged, is loaded with the vision and wisdom of the industry management. One can reasonably concludes that (1) Brand awareness and quality priority play fundamental roles; (2) The key of brand construction is to set a clear brand orientation by conducting thorough market investigation; (3) The core is to enable brand localization and keep meeting customers' satisfaction; (4) Identifying market orientation and conducting continuous reform and innovation are the sustaining momentum.

With an aim of building successful world brands, Chinese companies can learn from these successful practices. It is expected to enhance brand awareness, to set a clear brand position and meticulous strategies, to achieve brand localization by following local customers' tastes, and to maintain the persistence of reform and innovation.

• References

[1] L.Y.Qian, "An analysis of the Marketing of P&G Brand Strategy," Marketing Weekly, vol.5, pp82-83 ,2005.

[2] X.X.Shen, "A Study on Brand Construction of KFC in China — Consequences in the Publicity Localization Context," *Jiannan Literature*, vol.12, pp.204-207,2011.

[3] Y.L.Tong, "An analysis of P&G Brand Strategy," *Heilongjiang International Economics*, vol.9, pp.66-67, 2006.

[4] J.J.Liu, *A Study on Management Strategy of Siemens in China*, Beijing: Beijing University of Technology; 2014.

[5] X.R.Zha, *A Study on The Bran Strategy of Starbucks*, Beijing:University of International Business and Economics; 2006.

[6] H.Y.Zhu, *An analysis of the Localization of KFC in China*, Hangzhou:Zhejiang University of Technology; 2012.

Table 1. 2017-2019 Top 10 Countries with Top 500 World Brands**

Rank	Country	The Number of Top 500 World Brands			Trend
		2017	2018	2019	
1	USA	233	223	208	↓
2	UK	39	42	44	↑
3	France	40	43	43	→
4	Japan	38	39	42	↑
5	China	37	38	40	↑
6	Germany	26	26	27	↑
7	Switzerland	21	21	21	→
8	Italy	14	15	14	↓
9	Denmark	8	7	9	↑
10	Sweden	7	7	7	→

Table 2. 2017-2019 Top 10 Countries with Top 500 World Brands**

Rank	Industry	The Number of Top 500 World Brands			Trend
		2017	2018	2019	
1	Automobile and Parts	38	36	35	↓
2	Food and beverage	35	33	33	→
3	Communication Media	33	35	30	↓
4	Retail	23	22	24	↑
5	Energy	25	23	23	→
6	Internet	23	25	22	↓
7	Banking	17	17	22	↑
8	Computers and Communication	19	20	21	↑
9	Telecommunication	15	15	19	↑
10	Insurance	14	15	17	↑

**** Tables 1 and 2 Data from "Top 500 World Brands" by World Brand Lab.**

StEering
Inter-University Research Center:
Design, Quality and Reliability

The magic of danmaku: a social interaction perspective of gift sending on live streaming platforms

Jilei Zhou[1], Jing Zhou[2*], Ying Ding[3], Hansheng Wang[1]

[1)] Guanghua School of Management, Peking University, China
[2)] School of Statistics, Renmin University of China, China
[3)] Renmin Business School, Renmin University of China, China

[*]zhoujing_89@126.com

Keywords: Live streaming; paid gifting; social interaction

ABSTRACT –This paper investigates the role of viewers' social interaction in paid gifting on live streaming platforms. We argue that viewer-viewer interaction can prompt paid gifting by affecting viewers' arousal level through stimuli extracted from danmaku. Types of danmaku-related stimuli are presence of others, social competition, and emotional stimuli. Specifically, presence of others is measured by total number of words; social competition by debate level; and emotional stimuli by similarity of danmaku, number of excitement-related words, and number of emoji. Empirical results show that except for number of emoji, the other four variables positively affect paid gifting.

1. INTRODUCTION

In the past few years, live streaming has attracted numerous users in China. However, Chinese live streaming firms derive only a small proportion of their revenue from advertising. Most of their revenue comes from a practice called paid gifting, which was invented by Chinese companies. Since paid gifting affects the revenue of a firm, understanding the factors that affect paid gifting on live streaming platforms becomes important for practitioners. While many studies have investigated the factors that impact gifting in the real world [1-2], investigation of the influence of paid gifting in the virtual community, especially in live steaming, has just begun. In this paper, we aim to examine the impact of social interaction on paid gifting rather than status seeking. The social interaction in broadcast media can be simply classified into two categories: broadcaster-viewer interaction and viewer-viewer interaction. We notice that few researchers have already shown interest in examining the relationship between broadcaster-viewer interaction and viewers' support or consumption behavior in live streaming platforms [3-4]. However, little is known about the role of viewer-viewer interaction.

In this work, we argue that viewer-viewer interaction can play an important role in prompting paid gifting by affecting viewers' arousal level. In the context of live streaming, danmaku is a major way for viewers to interact with others. We then claim that arousal level can be influenced by different stimuli extracted from danmaku. We focus on three types of danmaku-related stimuli: presence of others, social competition, and emotional stimuli. Specifically, the presence of others is measured by total number of words. Social competition is measured by debate level. Emotional stimuli are measured by the level of similarity of danmaku, number of excitement-related words, and number of emoji. According to arousal theory, we argue that viewer-viewer interaction is positively related to the amount of paid gifting because of elevated arousal level. Data used in this paper were collected by crawling DOUYU.COM from September 1 to 17, 2017. We find several interest results, which has both theoretical and pratical implications.

2. METHDOLOGY

We argue that viewer-viewer interaction can play an important role in prompting paid gifting by affecting viewers' arousal level. In the context of live streaming, the interaction between viewers can be seen as sensory stimuli. A more active interaction indicates high-intense stimuli, which intuitively induces a high level of arousal. Such high arousal level at a given point can impair a calm, careful decision-making process, which usually results in hedonic purchase behavior [5] or overbidding in auction [6]. The paid gifting in a live streaming platform can be seen as a hedonic product. Therefore, we suggest that viewer-viewer interaction can increase arousal level, which can prompt paid gifting. Specifically, we have the following hypothesis:

Hypothesis 1: The number of gifts given by viewers in a period is positively related to the total number of words in the same period.

Hypothesis 2: The number of gifts given by viewers in a period is positively related to the level of debate in the same period.

Hypothesis 3: The number of gifts given by viewers in a period is positively related to the similarity level of danmaku in the same period.

Hypothesis 4: The number of gifts given by viewers in a period is positively related to the number of excitement related words in the same period.

Hypothesis 5: The number of gifts given by viewers in a period is positively related to the number of emoji in the same period.

To examine the hypotheses proposed earlier, we conducted a linear regression.

3. RESULTS AND DISCUSSION

Empirical results are reported in Table 1. The proposed model is statistically significant (F = 14520) and suitable (adjusted R^2 = 66.46%). First, Table 1 reports a significant positive effect for the total number of words (β = 0.009, p < 0.001). This means that the higher the number of total words is, the more gifts a broadcaster will receive. Second, the level of debate also has a significant positive (β = 4.083, p < 0.001) relationship with the dependent variable. That is, a heated discussion will lead to a higher tendency to send gifts. Therefore, hypotheses 1 and 2 are supported. Next, the similarity of danmaku (β= 28.659, p < 0.001) and the number of excitement-related words (β= 3.237, p < 0.001) also have a positive effect on gift sending. These variables measure the interaction between viewers and broadcasters. Therefore, hypotheses 3 and 4 are supported. However, the number of emoji (β = -0.056, p = 0.152) shows no significant effect on gifting. Hypothesis 5 is not supported. This is because in this case we can only extract emoji symbols from danmaku, but cannot know whether each is positive or negative. Therefore, the effect of emoji may not be detected. We summarize our hypotheses testing results in Table 4.

In addition to the main hypotheses, we also derived several other findings from control variables, most of which are lagged terms. First, we obtain a significant positive effect for the lagged term of number of viewers entered (β = 2.312, p < 0.001). This result implies the role of status seeking, which is consistent with prior studies of virtual gifting. The larger the number of entered viewers is, the higher level of status-based utility they derive from their peers. Second, we also find a significant positive effect for the lagged term of number of gifts gained (β = 0.751, p < 0.001). This reflects a herding effect, which is also widely discussed in prior studies. Finally, the lagged terms of number of viewers left (β = -3.871, p = 0.036) and number of gifts lost (β = -0.772, p < 0.001) are negatively associated with gift sending. This indicates that different channels have a competitive relationship.

Table 1 Results of linear regression

Variable	Estimate	Sd
Constant	1.863$^{\cdot}$	0.975
Total number of words	0.009***	0.000
Level of debate	4.083***	1.029
Similarity of *danmaku*	28.659***	2.156
Number of excitement-	3.237***	0.521
Number of emoji	-0.056	0.007
Last number of viewers	2.312***	0.366
Last number of viewers	-3.871*	1.844
Last number of gifts	0.751***	0.025
Last number of gifts lost	-0.772***	0.103

* p<0.05, ** p<0.01, *** p<0.001

4. CONCLUSIONS

We study the impact of viewer-viewer interaction on paid gifting in live streaming. We focus on three types of viewer-viewer interaction in various scenarios; they are presence of others, social competition, and emotional stimuli. In the end, we found several interesting results. This paper has several theoretical implications compared with previous literature. First of all, we contribute to the literature on social media by investigating live streaming, which is an emerging social medium equipped with a novel function called paid gifting. Second, we contribute to the literature on gifting by considering paid gifting on live streaming platforms. Previous literature mainly focuses on gifting in the real world.; Toubia and Stephen, 2013). Third, we contribute to the literature on broadcast media consumption by considering interaction between viewers. Besides, this study gives the live streaming platform insight into designing danmaku exposure.

5. REFERENCES

[1] Baskin, E., Wakslak, C.J., Trope, Y., Novemsky, N., 2014. “Why feasibility matters more to gift receivers than to givers: A construal-level approach to gift giving,” *Journal of Consumer Research*, vol. 41, pp. 169-182.

[2] Yang, A., Urminsky, O., 2015. Smile-seeking givers and value-seeking recipients: Why gift choices and recipient preferences diverge. Working Paper.

[3] Hamari, J., Sjöblom, M., 2017. What is eSports and why do people watch it? Internet research 27, 211-232.

[4] Payne, K., Keith, M.J., Schuetzler, R.M., Giboney, J.S., 2017. Examining the learning effects of live streaming video game instruction over Twitch. Computers in Human Behavior 77, 95-109.

[5] Fedorikhin, A., Patrick, V.M., 2010. Positive mood and resistance to temptation: The interfering influence of elevated arousal. Journal of Consumer Research 37, 698-711.

[6] Ku, G., Malhotra, D., Murnighan, J.K., 2005. Towards a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions. Organizational Behavior and Human decision processes 96, 89-103.

Social Dinner at Restaurant "Finisterrae"

Tuesday February 20th, 2020, 8.30pm

Address: Piazza Santa Croce, 12, 50122 Firenze.

Tel: +39 0552638675

List of participants

Arcidiacono, Gabriele	g.arcidiacono@unimarconi.it
Bacherini, Stefano	stefano.bacherini@alstomgroup.com
Bargues, Ángela Sebástia	angela.sbargues@uclm.es
Bertocci, Francesco	francesco.bertocci@esaote.com
Baumanns, Sascha	sascha.baumanns@gmail.com
Belingardi, Giovanni	giovanni.belingardi@polito.it
Berni, Rossella	rossella.berni@unifi.it
Bollani, Luigi	luigi.bollani@unito.it
Bonora, Nicola	nbonora@unicas.it
Cafarelli, Antonio	cafarelli.an@pg.com
Cantone, Luciano	luciano.cantone@uniroma2.it
Catelani, Marcantonio	marcantonio.catelani@unifi.it
Ceccato, Riccardo	ceccato@gest.unipd.it
Celegato, Alessandro	alessandro.celegato@alice.it
Citti, Paolo	p.citti@unimarconi.it
Cristaldi, Loredana	loredana.cristaldi@polimi.it
De Ketelaere, Bart	bart.deketelaere@kuleuven.be
Ferrari, Guido	guido.ferrari@unifi.it
Frieri, Rosamarie	rosamarie.frieri2@unibo.it
Galimberti, Chiara	c.galimberti19@campus.unimib.it
Galvis, Leonardo	leogalro@hotmail.com
García-Camacha Gutiérrez, Irene	irene.garciacamacha@uclm.es
Guinot, Christiane	christianeguinot@gmail.com
Guinot, Philippe	christianeguinot@gmail.com
Guo, Qian	guoqian561@126.com
Haselgruber, Nikolaus	nh@cis-on.com
Hu, Kingpei	qingpeihu@amss.ac.cn
Jin, Yongjin	jinyongj_519@aliyun.com
Lanzerath, Maria	mlanzera@wlgore.com
Leardi, Carlo	carlo.leardi@tetrapak.com
Li, Zhongdong	1531364336@qq.com
Liu, Dingping	zzldp@126.com
Lopez-Fidalgo, Jesus Fernando	fidalgo@unav.es

Magrini, Alessandro	alessandro.magrini@unifi.it
Martín Martín, Raúl	raul.mmartin@uclm.es
Mestres, Marina Vives	marina.vives@udg.edu
Moler, Jose Antonio	jmoler@unavarra.es
Nikiforova, Nedka	n.nikiforova@unifi.it
Palacios, Blanca	blanca.palacios@unavarra.es
Pegoraro, Luca	pegoraro@gest.unipd.it
Pietavolo Antonio	antonio.pievatolo@mi.imati.cnr.it
Pistone, Giovanni	giovanni.pistone@carloalberto.org
Ranfagni, Silvia	silvia.ranfagni@unifi.it
Ruggiero, Andrew	a.ruggiero@unicas.it
Shin, Yonggwan	syg.stat@gmail.com
Stefanini, Federico Mattia	federicomattia.stefanini@unifi.it
Tan, Matthias Hwai Yong	matthtan@cityu.edu.hk
Tian, Yubin	tianyb@bit.edu.cn
Trotta, Iacopo	iacopo.trotta@alstomgroup.com
Tsung, Fugee	season@ust.hk
Vicario, Grazia	grazia.vicario@polito.it
Vining, G. Geoff	vining@vt.edu
Wang, Dianpeng	wdp@bit.edu.cn
Wang, Kang	wk0536@126.com
Wilson, Ozuem	nh@cis-on.com
Xiong, Schifeng	xiong@amss.ac.cn
Yang, Jiazhen	yangjz3868@126.com
Yuchi, Xiaojuan	1531364336@qq.com
Zhao, Yanyun	cas-kriu@ruc.edu.cn
Zhou, Jing	zhoujing_89@126.com