ISSPM 2015

4th International Symposium on Statistical Process Monitoring

July 7–9, 2015, Padua, Italy

Abstracts

Dipartimento di Scienze Statistiche

Università degli Studi di Padova

Organizer

- Department of Statistical Sciences, University of Padua, Italy.

Scientific Commitee

- Sotiris Bersimis, University of Piraeus, Piraeus, Greece.
- Giovanna Capizzi, University of Padua, Padua, Italy.
- Philippe Castagliola, Université de Nantes, Nantes, France.
- Eugenio K. Epprecht, PUC-Rio, Rio de Janeiro, Brazil.
- Ross Sparks, CSIRO, Sydney, Australia.
- Kwok Leung Tsui, City University of Hong Kong, Honk Kong.
- William H. Woodall, Virginia Tech, Blacksburg, USA.

Local Organizing Commitee

- Giovanna Capizzi, University of Padua, Padua, Italy.
- Guido Masarotto, University of Padua, Padua, Italy.

Secretariat

- Federica Andolfo, University of Padua, Padua, Italy.
- Monica Boscolo Anzoletti, University of Padua, Padua, Italy.

Contents

Simultaneous Monitoring of Correlated Multivariate Linear and GLM based Regression Profiles in Phase II, *A. Amiri, F. Sogandi, M. Ayoubi* 5

Methods for Interpreting the Out-of-Control Signal of Multivariate Control Chart: A Comparison Study, *S. Bersimis, A. Sgora, S. Psarakis* 5

A New Memory-Type Monitoring Technique for Count Data, *P. Castagliola, A. C. Rakitzis, P. E. Maravelakis* 6

The Performance of the t and the Sign Control Charts in Processes with Finite Number of Inspections, *G. Celano, P. Castagliola, S. Chakraborti, G. Nenes* 7

Nonparametric Control Charts—An Overview, *S. Chakraborti, M. A. Graham* 7

Performance Analysis of Hotelling T^2 under Multivariate Inspection Errors, *W. Chattinnawat, C. B. Green* 8

Multisensor Data Fusion for Surface Metrology, *B. M. Colosimo, L. Pagani* 8

The Use of Two Side-Sensitive $\overline{X}$ Charts to Control Bivariate Processes, *A. F. B. Costa, F. D. Simões, M. A. G. Machado* 9

Identification of Outliers and Influential Observations: an Application, *M. S. De Magalhães, R. S. Von Doellinger, P. N. Silva* 9

Robust CUSUM Control Charts, *R. J. M. M. Does, H. Z. Nazir* 11

Monitoring the Spread of Multivariate Processes Using Projections, *E. K. Epprecht, S. F. Bastos* 12

A First Look at the Performance of a Bayesian Chart to Monitor the Ratio of Two Weibull Percentiles, *P. Erto* 13

Shewhart $\overline{X}$ and S^2 Control Charts with Guaranteed In-Control Performance, *A. Faraz, C. Heuchenne, W. H. Woodall* 13

Acceptance Sampling Plans for Inflated-Pareto Data, *F. O. Figueiredo, A. M. Figueiredo, M. I. Gomes* 14

Improving Biosurveillance System Performance, *R. D. Fricker, Jr.* 15

Process Capability in Multivariate NonLinear Profiles, *R. D. Guevara, J. A. Vargas* 15

SPRT Chart with Optimal Charting Parameters for Monitoring Attributes, *S. Haridy* 16

Monitoring Directional Data, *D. M. Hawkins, F. Lombard* 16

ATTRIVAR: a New Attribute Control Chart to Monitor Process Mean, *L. L. Ho, F. Aparisi* 17

What Can We Learn from Statistical Learning Methods in Process Monitoring?, *L. A. Jones-Farmer, M. Weese, W. Martinez, F. Megahed* 17

Process Monitoring and Diagnosis Via Optimization Methods, *S. B. Kim, Y. H. Kim* 18

Calculating the Steady-State ARL of MEWMA Control Charts, *S. Knoth* 19

Covariance Matrix Control Chart using Support Machines Methods, *E. M. Maboudou* 19

An EWMA Control Chart for the Shape Parameter of the Weibull Distribution for Type II Censored Samples, *P. E. Maravelakis* 20

A Visual Analytics Perspective on the Analysis of Longitudinal Datasets, *F. M. Megahed, Y. Babu, J. George, Z. S. Maman* 20

Phase I Data Set Structure, *F. D. Moura Neto, P. A. S. Souza, M. S. De Magalhães* 21

Pseudo-Precision and Rank Tests, *M. Neuhäuser* 21

Monitoring of the Multi-stage Manufacturing Process Using the Bayesian Structural Equation Modeling Approach, *C. Park* 22

Applications of Multivariate Statistical Process Control in Non-Industrial Processes, *S. Psarakis, A. Sgora, S. Bersimis* 22

Dynamic Screening: An Approach for Expanding the SPC Applications, *P. Qiu, D. Xiang* 23

Controlling Processes of Generally Inflated Poisson Counts, *A. C. Rakitzis, P. Castagliola, P. E. Maravelakis* 23

A Reversible Jump Markov Chain Monte Carlo Approach to the Multivariate Change-point Problem, *S. E. Rigdon, R. Steward, R. Pan* 24

Statistical Process Control From a Sequential Bayesian Standpoint, *G. J. Ross* 25

The c-chart with Bootstrap Adjusted Control Limits to Guarantee Conditional Performance, *A. G. Ryan Driscoll, M. Zhao, W. H. Woodall* 25

Statistical CUSUM Designs with Minimum Sampling Cost, *E. M. Saniga, D. Davis, T. McWilliams, J. Lucas* 26

Surveillance of Non-Stationary Processes, *W. Schmid, T. Lazariv* 28

Monitoring Communication Counts- Detecting Communication Outbreaks for a Sub-Group of a Larger Target Group of People, *R.S. Sparks, W. H. Woodall, J. D. Wilson* 28

Regulating the Bias-Variance Tradeoff in Analysis of Survey Data over Time with Application to a Customer Loyalty Survey, *S. H. Steiner, T. Cooper-Barfoot, J. MacKay* 29

Monitoring Multinomial Attribute Processes Based on Weighted Sum of Chi Squares, *H. Taleb, E. Khediri* 30

A Fully Adaptive Multivariate Control Scheme for Simultaneous Monitoring of Mean Vector and Covariance Matrix of Processes Subject to a Multiplicity of Assignable Causes, *K. A. Tasias, G. Nenes* 30

Bayesian SPC for Count Data, *P. Tsiamyrtzis, D. M. Hawkins* 31

Evolution of Big Data Analytics, *K. L. Tsui* 31

Quality Engineering Faces the Challenges of Big Data and Little Data, *F. Tsung* 32

A New Control Chart for Monitoring Nonlinear Profiles, *J. A. Vargas, R. D. Guevara, N. Acevedo* 32

An Overview and Perspective on Social Network Monitoring, ***W. H. Woodall, M. Zhao, R.S. Sparks, J. D. Wilson*** **33**

Some Thoughts on SPC Applications in Reliability and Maintenance Monitoring, ***M. Xie, Y. Cheng*** **33**

An ARL-Unbiased EWMA Proportion Control Chart, ***S. Yang*** **34**

Alarm Attributes in Systems for Monitoring Quality and Reliability data, ***E. Yashchin*** **34**

High-Dimensional Statistical Process Control, ***C. Zou, L. Feng*** **35**

Another Look at the EWMA Control Chart with Estimated Parameters, ***I. M. Zwetsloot, N. A. Saleh, M.A. Mahmoud, L.A. Jones-Farmer, W.H. Woodall*** **35**

Author Index **37**

Simultaneous Monitoring of Correlated Multivariate Linear and GLM based Regression Profiles in Phase II

A. Amiri (Shahed University, IRN), F. Sogandi (Shahed University, IRN), M. Ayoubi (Tarbiat Modares University, IRN)

In some applications, quality of a process or product is characterized by correlated multivariate linear and generalized linear models (GLM) based regression profiles. Monitoring these profiles separately leads to misleading results because the correlation structure among the multivariate linear and GLM profiles is neglected. In this paper, we specifically concentrate on Phase II and propose two procedures for monitoring multivariate linear and GLM based regression profiles. The performance of the proposed methods is evaluated under different magnitudes of shifts in the regression parameters in terms of average run length (ARL) criterion through simulation studies. The results of simulation studies show the superior performance of the proposed methods rather than monitoring the multivariate linear and GLM profiles separately. Finally, the application of the proposed methods is illustrated through a real case in producing air conditioner in Iran.

Keywords: Multivariate Linear Profile; Generalized Linear Model (GLM); Phase II; Average run length (ARL).

Methods for Interpreting the Out-of-Control Signal of Multivariate Control Chart: A Comparison Study

S. Bersimis (University of Piraeus, GRC), A. Sgora (University of Piraeus, GRC), S. Psarakis (Athens University of Economics and Business, GRC)

Multivariate control charts have proved to be a useful tool in Statistical Process Control (SPC) for identifying an out-of-control process. However, one of the main drawbacks of these charts is that they do not indicate which measured variables have shifted. A number of alternative approaches that aim to diagnose out-of-control conditions and help identify variables responsible for an out-of-control state may be found in the literature. This paper reviews several techniques that are used to diagnose sources responsible for out-of-control conditions and attempt to make a comparative study between them. In particular, we evaluate the performance of each method under different simulation scenarios in terms of successful identification of the out-of-control variables.

Keywords: Multivariate Control Charts; Interpretation; Out-of-control Signal.

A New Memory-Type Monitoring Technique for Count Data

P. Castagliola (Université de Nantes & IRCCyN UMR CNRS 6597, FRA), A. C. Rakitzis (Université de Nantes & IRCCyN UMR CNRS 6597), P. E. Maravelakis (University of Piraeus)

When it is of interest the monitoring of a high-yield process or a health-related process, the considered quality characteristic cannot always be conveniently represented numerically. In such cases, the common practice is to classify each inspected item (or unit) as either conforming or non-conforming, according to the specifications of that quality characteristic. Therefore, for the monitoring of such processes, attributes control charts like the np or the c charts are used. These charts are known to be insensitive in the detection of small and moderate shifts in the parameter(s) under surveillance. A solution to this problem is to incorporate information from the past observations and use, instead of Shewhart-type charts, control charts with memory, such as the exponentially weighted moving average (EWMA) chart.

However, when the observed data are discrete, by applying the classical EWMA control chart for attributes, the values of the EWMA chart statistic are not integers anymore. Moreover, the set of the attainable values changes at each time t. Thus, using the Markov chain method of Brook and Evans (1972) for obtaining the performance of an EWMA control chart for attributes, only leads to an approximation of its actual performance.

Consequently, the goals of this presentation is a) to introduce a new EWMA-type monitoring technique only working with integer values (without any rounding operation), i.e. the values, the computed statistics and the weights are all integer values., b) to evaluate its exact run length properties and c) to compare it with alternative methods like the c, Poisson EWMA, RY EWMA and Poisson CUSUM charts.

Keywords: Average Run Length (ARL); Attributes control charts; Markov chain; Poisson distribution; Poisson CUSUM; Poisson EWMA.

The Performance of the t and the Sign Control Charts in Processes with Finite Number of Inspections

G. Celano (University of Catania, ITA), P. Castagliola (Université de Nantes & IRCCyN UMR CNRS 6597, FRA), S. Chakraborti (University of Alabama, USA), G. Nenes (University of Western Macedonia, GRC)

In many manufacturing environments the production horizon of the same part code between two consecutive set-ups should be limited to a few hours or shifts. When 100% sampling is not possible, on-line quality control should be immediately started by means of a control chart, which is run for a finite number of inspections until the end of production. An extensive simulation study is here conducted to compare the statistical performance of the distribution-free Shewhart SN control chart to the normal theory based Shewhart Student's t control chart for processes with finite horizon. With the implementation of a distribution-free control chart any model assumption about the underlying distribution, such as normality, is needless: this overcomes the important problem of lack of information about the distribution of the observations collected for the quality characteristic to be monitored after each process set-up. Several types of underlying distributions of observations and different numbers of scheduled inspections are considered to show the statistical performance of both the control charts. To carry out the performance comparison for processes having finite production horizon, we use the expected value and the standard deviation of the truncated run length TRL. An illustrative example presents the implementation of both the control charts on a real data set collected from a bottling process of a carbonated drink.

Keywords: Finite Production Horizon; Shewhart Control Chart; Student's t Statistic; Sign Statistic; Truncated Run Length.

Nonparametric Control Charts—An Overview

S. Chakraborti (University of Alabama, USA), M. A. Graham (University of Pretoria, ZAF)

Nonparametric (distribution-free) control charts are useful and applicable in a variety of process monitoring situations, particularly when there is a lack of knowledge about the underlying distribution. These chars have steadily grown in popularity over the last few decades and several nonparametric control charts have been proposed in the literature. The journal Quality and Reliability Engineering International recently published a special issue dedicated to the topic of nonparametric control charts. In this talk we review the basic ideas of nonparametric control charts and provide an overview of some of the literature. Future directions are

indicated.
Keywords: Process Monitoring and Control; Knowledge of Distribution; Robustness; Performance.

Performance Analysis of Hotelling T^2 under Multivariate Inspection Errors

W. Chattinnawat (Chiang Mai University, THA), C. B. Green (North Dakota State University, USA)

This research analyzed the effects of the multivariate normal inspection errors on the performances of the Hotelling for individual observation during phase II. The computed and charted statistic based on the contaminated measurements of Xs may lead to inferior performances in the false alarm rate, power of detection and an Average Run Length (ARL). This research derived explicitly how the multivariate inspection errors are related to the Hotelling statistics. The performance of the MSPC in terms of in-control and out-of-control ARL are reported with respect to different cases/structures of the inspection errors. Given fixed sizes of measurement error covariances, both of the false alarms and detection time can be dramatically affected by just only 10% of the each measurement error. High correlations of the measurement errors are induced larger variances and covariances of the observed process characteristics, leading to inferior performance of the T^2 chart. This research finding has advantages that one can predict how the Hotelling T^2 chart will be inferior with respect to the measurement structure, i.e., %GRR. The lemmas provided in this research yield an explicit expression for numerical analysis of the inferior performance of the T^2 chart.

Keywords: T^2 chart; Measurement Error; MSA; Multivariate SPC.

Multisensor Data Fusion for Surface Metrology

B. M. Colosimo (Politecnico di Milano, ITA), L. Pagani (Politecnico di Milano, ITA)

More and more often, multiple sensors can be used to perform the same measurement task while providing different levels of data density and/or accuracy/precision. A common example is use of contact and contactless sensors to acquire the same manufactured surface. In this scenario, data fusion can be effectively used to improve the reconstruction of the measured surface thus providing a better model for quality inspection (comparison with the target shape) and/or quality monitoring (i.e., detection of unwanted changes from the in-control pattern). This contribu-

tion discusses existing approaches for data fusion in surface metrology in order to outline advantages and challenges of the existing methods. New approaches are then proposed to deal with large sample size of contactless sensors (i.e., the big data issue) and/or uncertainty propagation.

Keywords: Surface; Quality Inspection; Quality Monitoring; SPC; Profile Monitoring; Surface Monitoring; Data Fusion; Gaussian Processes.

The Use of Two Side-Sensitive $\overline{X}$ Charts to Control Bivariate Processes

A. F. B. Costa (São Paulo State University, BRA), F. D. Simões (São Paulo State University, BRA), M. A. G. Machado (São Paulo State University, BRA)

The Hotelling chart was proposed to control p means. Its performance is superior to the joint performance of p $\overline{X}$ charts. Differently of the Hotelling chart, the $\overline{X}$ chart has upper and lower action regions. Because of that, the side-sensitive rule substantially enhances its performance. In our study we considered the bivariate case (p=2) to compare the Hotelling chart with the two side-sensitive $\overline{X}$ charts. The overall performance of the two side-sensitive $\overline{X}$ charts is superior to the performance of the Hotelling chart, in special, when the mean vector parameters don't shift in opposite directions.

Keywords: Bivariate Processes; Side-sensitive Rules; Hotelling Chart; $\overline{X}$ Chart.

Identification of Outliers and Influential Observations: an Application

M. S. De Magalhães (National School of Statistical Sciences / Brazilian Institute of Geography and Statistics, BRA), R. S. Von Doellinger (Division of Methods and Quality / Brazilian Institute of Geography and Statistics, BRA), P. N. Silva (National School of Statistical Sciences / Brazilian Institute of Geography and Statistics, BRA)

One of the most important stages of the planning process and execution of a survey is data editing, in which are identified and eliminated erroneous values, called outliers, i e, observations that deviate from a data model (Barnett and Lewis, 1994; Lee, 1995) and do not reflect the reality of the phenomenon been studied. In certain cases, for good quality statistical information be provided, it may not be necessary to identify all the errors present in the data. It is just sufficient to detect influential observations, that is, those which when included or excluded from the analysis, significantly impact on the estimates of the parameters of interest. The approach generally used to identify influential observations is called

selective editing (Latouche and Berthelot, 1992; Lawrence and McKenzie, 2000). In the methods of selective editing, potentially influential observations are ranked based on values of a score function, which expresses the impact of the error in the estimates of parameters of interest. The observations with scores above a preset threshold are considered critical and should be revised. The definition of the score function implies in determining the probability of the observation to present error (risk component), as well as the magnitude of the error (component of influence). Risk and influence components are used by score functions presented in the literature (Jader and Norberg, 2005). According to Di Zio et al. (2008) the methods commonly employed to obtain the risk and influence components are based on comparison of the observed values of a given variable and the predicted values for a particular model. The differences between observed and predicted values are used in the calculation of scores for identifying observations that generate greater impact on the estimated of parameter of interest. Di Zio et al. (2008) proposed a multivariate model to estimate the probability of error and as well as the error magnitude. The method is based on contaminated normal models (Little, 2008). The data observed are described by a mixture of two multivariate normal distributions that represent the erroneous or contamined data and the data without errors. It is assumed that the distribution of the contaminated data can be obtained by the distribution of the data without errors with an increase in the variance (Ghosh-Dastidar and Schafer, 2006). In this paper, the method of selective editing proposed by Di Zio et. al. (2008) was applied to identify outliers and influential observations in the Household Budget Survey (HBS 2008/2009) of the Brazilian Institute of Geography and Statistics (IBGE) through the use of the following variables, the monthly household income and the annual household expenditure. Moreover, a new method based on a modification of the technique presented by Dizio et al. (2008) is proposed.

References

Barnett V., Lewis T. (1994). Outliers in Statistical Data, New York: Wiley.

Di Zio M., Guarnera U. (2008). A multiple imputation method for non-Gaussian data. Metron LXVI (1):75–90.

Ghosh-Dastidar M., Schafer J.L. (2003). Multiple edit multiple imputation for multivariate continuous data. J Am Stat Assoc 98:807–817.

Latouche M., Berthelot J.M. (1992). Use of a score function to prioritize and limit recontacts in editing business surveys. Journal of Official Statistics, 8, n.3, 389- 400.

Lawrence D., McKenzie R. (2000). The General Application of Significance Editing. Journal of Official Statistics, 16, n. 3, 243-253.

Lee H. (1995). Outliers in Business Surveys, in: Business Survey Methods, Cox B.G., Binder D.A., Chinappa B.N., Christanson A., Colledge M.J. and Kott P.S. (Eds), John Wiley and Sons, Inc. 503-526.

Little, J.A. (1988). Robust estimation of the mean and covariance matrix from data with missing values, J. R. Stat. Soc., Ser. C, Vol. 37, No. 1, 23-38.

Jäder A., Norberg A. (2005). A Selective Editing Method considering both suspicion and potential impact, developed and applied to the Swedish Foreign Trade Statistics, UN/ECE Work Session on Statistical Data Editing, Ottawa (http://www.unece.org/stats/documents/2005.05.sde.htm).

Keywords: Influential Observations; Score Function; Contaminated Data; Household Budget Survey.

Robust CUSUM Control Charts

R. J. M. M. Does (University of Amsterdam, NLD), H. Z. Nazir (University of Sargodha, PAK)

Control charts may be classified into two categories: memoryless control charts and memory control charts. Shewhart-type control charts are termed as memoryless control charts and their main deficiency is that they are less sensitive to small and moderate shifts in the parameters (location and dispersion). The commonly used memory control charts in the literature include cumulative sum (CUSUM) control charts and exponentially weighted moving average (EWMA) control charts. These memory control charts are designed such that they use the past information along with the current information which makes them very sensitive to small and moderate shifts in the process parameters. Most of the evaluations of the existing control charts depend on the assumptions of normality, no contaminations, no outliers and no measurement errors in Phase I for the quality characteristic of interest. In case of violation of these assumptions, the design structures of the charts lose their performance ability and hence are of less practical use. One of the solutions to deal with this is to use control charts which are robust against violations of the basic assumptions, like normality. In this talk we discuss robust CUSUM-type control charts. Particularly, this talk will employ highly robust estimators in Phase I and will study the impact of these estimators on the performance of Phase II CUSUM control charts.

Keywords: Shewhart Control Charts; CUSUM; Phase I; Phase II; Statistical Process Control.

Monitoring the Spread of Multivariate Processes Using Projections

E. K. Epprecht (PUC-Rio, Rio de Janeiro), S. F. Bastos (PUC-Rio, Rio de Janeiro, BRA)

With multivariate processes, the case may be that both the natural process variability and the variability due to special causes are due to independent sources of variation which can be represented by non-observable (or "latent") random variables, independent between themselves. The realizations of the observable variables are in this case largely due to the realizations of the latent variables (of which they are a function), plus random noise. A shift in the mean of one of the latent variables or an increase in its spread result, respectively, in a shift in the mean or an increase in the variability of the vector of observable variables along a specific (or "assignable") direction. In ISSPC'2013 I proposed controlling the mean of this kind of processes with control charts on the estimates of the values of the latent variables, obtained by projections of the vector of observed variables onto the assignable directions. In the present article we propose controlling the spread of such processes with standard deviation control charts (S charts) of the values of the projections of the vector of observed variables on the assignable directions, used in conjunction with a chart on the average of the squared norms of sample residual vectors. The latter is devised to signal process changes due to novel special causes, previously unknown, that increase the process variability along directions that are orthogonal to the subspace of the assignable directions. The performance of the proposed control scheme was compared, via simulation, with other schemes for monitoring the variance-covariance matrix of multivariate processes, namely: the |S| chart, on the the generalized variance; the VMAX chart, on the maximum sample variance; and the VS (variances summation) chart. The proposed scheme has shown to be the most efficient at signalling increases in the spread of this kind of process, with the aditional advantage of directly indicating the variation source over which the special cause is acting.

Keywords: Latent Variables; Projections; Non-observable Variables; Assignable Directions of Variation; Control Charts; Statistical Process Control; Spread; Multivariate Processes; Monitoring.

A First Look at the Performance of a Bayesian Chart to Monitor the Ratio of Two Weibull Percentiles

P. Erto (University of Naples Federico II, ITA)

The aim of the present work is to investigate the performance of a specific Bayesian control chart used to compare two processes. The chart monitors the ratio of the percentiles of a key characteristic associated with the processes. The variability of such a characteristic is modeled via the Weibull distribution and a practical Bayesian approach to deal with Weibull data is adopted. The percentiles of the two monitored processes are assumed to be independent random variables. The Weibull distributions of the key characteristic of both processes are assumed to have the same and stable shape parameter. This is usually experienced in practice because the Weibull shape parameter is a characteristic of the concerned main phenomenon. However, if a change of the shape parameters of the processes is suspected, the involved distributions can be used to monitor their stability too. We first tested the chart when very poor prior information is adopted. To this end, the "prior" values were changed to reflect a 50% shift in both directions from the original values of both the percentile and the shape parameter of the two monitored processes. Then the same above 50% shifts were considered for the "sampling" distributions after the Phase I, with the purpose of estimating the diagnostic ability of the charts to signal an out-of-control state. The traditional approach based on the Average Run Length, empirically computed via a Monte Carlo simulation, was adopted. In this way, the comparison with alternative methods to monitor Weibull processes is made easier.

Keywords: Bayesian Estimators; Control Chart; Weibull Distribution; Average Run Length.

Shewhart $\overline{X}$ and S^2 Control Charts with Guaranteed In-Control Performance

A. Faraz (University of Applied Sciences Upper Austria, AUT), C. Heuchenne (University of Liège, BEL), W. H. Woodall (Virginia Tech, USA)

The in-control performance of the Shewhart $\overline{X}$ and S^2 control charts with estimated in-control parameters has been evaluated by a number of authors. Results indicate an unrealistically large amount of Phase I data is needed to have the desired in-control average run length (ARL) value in Phase II. To overcome this problem, it has been recommended that the control limits be adjusted based on a bootstrap method to guarantee that the in-control ARL is at least a specified value

with a certain specified probability. In our paper we present simple formulas for the required control limits so that practitioners do not have to use the bootstrap method. An assumption of normality is required. The advantage of our proposed method is in its simplicity; there is no bootstrapping and the control chart constants do not depend on the Phase I sample data.

Keywords: Average Run Length (ARL); Bootstrap; Effect of Estimation Error; Statistical Process Control.

Acceptance Sampling Plans for Inflated-Pareto Data

F. O. Figueiredo (FEP, Universidade do Porto & CEAUL, Universidade de Lisboa, PRT), A. M. Figueiredo (FEP & LIAAD-INESC Porto, Universidade do Porto, PRT), M. I. Gomes (FCUL & CEAUL, Universidade de Lisboa, PRT)

In the food industry as well as in other manufacturing industries, it is very important to control the presence of some chemical substances in the raw material that will affect the quality of the final products. Chromatography analyses are performed on samples of items taken from large batches, and on the basis of the obtained measurements we have to conclude about the absence or presence of such substances, and then decide for the acceptance or rejection of the corresponding batches. However, most of the chromatographs have not sufficient precision to detect very low concentrations of such chemical substances, i.e., levels below a certain threshold, and thus, the measurement results of the quality characteristic suggest an underlying inflated continuous distribution, in some cases with a right heavy-tail. In this work we highlight the adequacy of the Inflated-Pareto distribution to model such type of data, and then, we define and evaluate acceptance-sampling plans for variables under this distributional set-up. The development of control charts to monitor processes that generate data of this type will also be of great interest.

Keywords: Acceptance Sampling Plans; Control Charts; Inflated-Pareto Distribution; Statistical Process Control.

Improving Biosurveillance System Performance

R. D. Fricker, Jr. (Virginia Tech, USA)

This talk will describe the problem of biosurveillance which, unlike the typical process monitoring problem, often involves monitoring dozens, hundreds, or thousands of time series. We describe a methodology for optimizing such a biosurveillance system, where the goal is to maximize the system-wide probability of detecting an "event of interest" subject to a constraint on the expected number of false signals. Using this approach, public health officials can "tune" their biosurveillance systems to best detect various threats, thereby allowing practitioners to focus their public health surveillance activities.

Keywords: Biosurveillance; Statistical Process Monitoring.

Process Capability in Multivariate NonLinear Profiles

R. D. Guevara (Universidad Nacional de Colombia, COL), J. A. Vargas (Universidad Nacional de Colombia, COL)

In some cases the quality of a process or product can be characterized by a functional relationship between a response variable and one or more explanatory variables, which is called profile. There are few studies to evaluate the capability of a process with profile quality characteristic, specifically, there is no method in the literature to analyze process capability characterized by multivariate nonlinear profiles. We propose a method to measure the capability of these processes based on a principal component method for multivariate functional data and the concept of functional depth. Performance of the proposed method is evaluated through simulation studies. An example illustrates this method.

Keywords: Functional Depth; Multivariate Functional Data; Multivariate Functional Principal Component Analysis; Nonlinear Profiles; Process Capability Indices.

SPRT Chart with Optimal Charting Parameters for Monitoring Attributes

S. Haridy (Benha University, EGY)

The Sequential Probability Ratio Test (SPRT) chart is considered as one of the most powerful tools for monitoring manufacturing processes due to its tremendous effectiveness. It is highly suitable for the applications where testing is destructive or very expensive, such as the automobile airbags test, ammunition test and uniaxial tensile test. This research studies the effect of the Average Sample Number (ASN) (i.e., the average sample size) on the chart's performance. A design algorithm is proposed to explore the optimal ASN of the SPRT chart for monitoring the fraction nonconforming p. Moreover, the optimal sample sizes of the np and binomial CUSUM charts, which have traditionally been used for monitoring p, are also searched. By optimizing the ASN and other charting parameters of the SPRT chart, the average detection speed is almost doubled. It is also found that the optimal SPRT chart is superior to the optimal np chart and CUSUM chart by 221% and 171%, respectively, in terms of Average Number of Defectives (AND). It is observed that the SPRT chart using a relatively smaller ASN and a smaller sampling interval (h) has a higher overall detection effectiveness.

Keywords: Sequential Probability Ratio Test (SPRT); Control Chart; Average Sample Number (ASN).

Monitoring Directional Data

D. M. Hawkins (University of Minnesota, USA), F. Lombard (University of the Northwest, USA)

Directional data are numbers on the range [0, 2pi). They arrive in traditional settings as angles, but in other less obvious settings such as time of the day, and the fractional parts of real numbers. The most common statistical distribution used for directional data is the von Mises distribution, which has a location and a scale parameter. We outline the optimal cumulative sum control charts for step changes in the location and the scale parameters of von Mises-distributed data, and introduce an effective non-parametric chart with good properties for monitoring location.

Keywords: von Mises; Angular; CUSUM; Nonparametric.

ATTRIVAR: a New Attribute Control Chart to Monitor Process Mean

L. L. Ho (University of São Paulo, BRA), F. Aparisi (Polytechnic University of Valencia, ESP)

Usually attribute control chart presents lower cost (operational and implementation) and less time spent in the sampling inspection than variable control charts. The aim of this paper is to propose a new attribute control chart, namely ATTRIVAR (ATTRIbutes + VARiables), to monitor process mean that has a performance equivalent to an $\overline{X}$ with the benefits of an attribute control chart. The process control starts employing an attribute chart. Each sampled unit is classified as approved or rejected according to some specified criteria. If the number of items classified as rejected is equal or exceeds the control limit than an action (a stoppage of the production process) is taken. Alternatively if the number of non-conforming items is lower than the control limit but equal or exceeds a warning limit then the units of the current sample (ATTRIVAR1) or units of the next sample (ATTRIVAR2) are measured and its average value, ($\overline{X}$) calculated. If the average value is not in the control limit region, an action is also taken otherwise the production process goes on. The parameters of this new control chart are optimized to match a required in-control ARL and to minimize the out-of-control ARL for a given shift. The optimized ATTRIVAR control chart has a performance very similar to the Shewhart's control chart, although the percentage of times that the variables are measured to compute $\overline{X}$ is rather low. Numerical example illustrates the current proposal

Keywords: Attribute Control Chart; X-bar Control Chart; In-control Average Run Length; Out-of-control Average Run Length; Optimization.

What Can We Learn from Statistical Learning Methods in Process Monitoring?

L. A. Jones-Farmer (Miami University, USA), M. Weese (Miami University, USA), W. Martinez (Miami University, USA), F. Megahed (Auburn University, USA)

The increasing availability of high volume, high velocity data sets, often containing variables of different data types, brings an increasing need for monitoring tools that are designed to handle these big data sets. While the research on multivariate statistical process control tools is vast, the application of these tools for big data sets has received less attention. We discuss some of the main directions involving statistical learning and dimension reduction techniques applied to control charts. Although many of these methods show promise for monitoring and surveillance of

large and diverse data sets they have not been widely adopted in practice. Statistical learning is, by nature, an exploratory practice. Thus, the SPC methods based on statistical learning algorithms require practitioners to embrace a different paradigm than traditional model-based or nonparametric control charts. It may be difficult to provide universally applicable design protocols for statistical learning control charts as the application of the methods tend to be very specific to particular samples. Additionally, when using statistical learning-based control charts, it may be difficult to achieve a precise statistical performance in terms of false alarm probabilities or average run lengths. We hope to bring into better focus some of the issues related to monitoring and surveillance methodology informed by data-mining techniques that show promise for monitoring large and diverse data sets.
Keywords: Control Charts; Ensembles; Neural Networks; Regression; Support Vector Machines; Variable Selection.

Process Monitoring and Diagnosis Via Optimization Methods

S. B. Kim (Korea University, KOR), Y. H. Kim (Korea University, KOR)

Recently, process monitoring methods that combine statistical learning algorithms and control charts have been developed to handle the large streams of complex data found in modern systems. In this work, we propose a novel algorithm for one-class classification based on the mixed integer quadratic programming (MIQP) method. The proposed method minimizes the radius of a spherically shaped boundary subject to the number of target data being equal to a constant value specified by users. By modifying this constant value, users can exactly control the proportion of target data described by the spherically shaped boundary. Thus, the proportion of out-of-control observations can be made theoretically equal to a Type I error rate in Phase I process. Moreover, similar to support vector data description, the flexible boundary can be achieved by incorporating kernel functions. New multivariate control chart based on the proposed MIQP method is presented. The usefulness and applicability of the proposed method is demonstrated through experiments with simulated data.
Keywords: Mixed Integer Quadratic Programming; Multiviariate Control Charts; One-class Classification Algorithm; Optimization; Process Monitoring.

Calculating the Steady-State ARL of MEWMA Control Charts

S. Knoth (Helmut Schmidt University Hamburg, DEU)

In Runger's and Prabhu's seminal paper about calculating ARL values for MEWMA (multivariate exponentially weighted moving average) charts deploying the popular Markov chain approximation two issues turned up. The numerical accuracy (measured in number of valid digits) is not good. Second, an unusual type of steady-state ARL was considered. Here, we want to study two types of steady-state ARLs, the conditional and the cyclical one. Both are calculated with higher accuracy than Runger/Prabhu's one.**References**

G.C. Runger, S.S. Prabhu (1996), "A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Chart", JASA 91(436), 1701-1706.

Keywords: Multivariate SPC; Numerical Approximation; Average Run Length.

Covariance Matrix Control Chart using Support Machines Methods

E. M. Maboudou (University of Central Florida, USA)

In real applications, when a change occurs in a multivariate process, it occurs in either location or scale. This talk focuses on monitoring the covariance matrix of multivariate processes when the underlying distribution of the quality characteristics departs from normality. Several methods have been proposed recently to monitor the covariance matrix. For most conventional control charts, such as Alt's charts, the design of the control limits is commonly based on the assumption that the quality characteristics follow a multivariate normal distribution. However, this may not be reasonable in many real-world problems. This talk addresses this issue and suggests a monitoring methodology motivated by statistical machine learning theory. The proposed multivariate control chart is based on the so-call support machines methodology. This chart makes use of information extracted from in-control preliminary samples. A case study demonstrates that the proposed control chart can perform better than conventional charts.

Keywords: Convex Functions; In-control ARL; Mercer Kernel; Reproducible Kernel Hilbert Space; Support Vector Machine.

An EWMA Control Chart for the Shape Parameter of the Weibull Distribution for Type II Censored Samples

P. E. Maravelakis (University of Piraeus, GRC)

In this paper we propose a new EWMA control chart for the shape parameter of the Weibull distribution for type II censored samples. We provide all the necessary details for the design of this control chart and we study its performance. An example is also provided in order to illustrate the use of the new control chart. Finally, some conclusions and recommendations are given along with a few suggestions for future research.

Keywords: EWMA; Weibull; Type II Censoring; Average Run Length.

A Visual Analytics Perspective on the Analysis of Longitudinal Datasets

F. M. Megahed (Auburn University, USA), Y. Babu (Auburn University, USA), J. George (Auburn University, USA), Z. S. Maman (Auburn University, USA)

Graphical methods are often used to visually and qualitatively describe data, as well as depict the relationship between different variables. The use of graphs has been promoted as indispensable component of exploratory data analysis (EDA), which allows statisticians and practitioners to identify outliers, trends, and patterns that merit further study. In addition, graphical methods such as histograms, quantile-quantile plots, and residual plots are often used for assumption checking, an important component of parametric data analysis protocols. It should be noted that these commonly used approaches were proposed prior to the advent of microcomputers as efficient methods to record and visualize data for single (or few) variable(s) processes. As the volume, variety, and velocity of data continues to evolve, there are opportunities to supplement and improve these methods for understanding and visualizing process variation. In this paper, we propose novel graphical tools that are more suitable for larger datasets. We demonstrate how these new tools can be used to better describe the data, using real datasets. Finally, we create a longitudinal visualization toolkit to allow practitioners to implement some of these visualization tools without the need for training, extensive statistical background, and/or specialized statistical software.

Keywords: Animated Graphs; Data Mining; Data Quality; Initial Data Analysis (IDA); Interactive Graphs; Tukey.

Phase I Data Set Structure

F. D. Moura Neto (Rio de Janeiro State University, BRA), P. A. S. Souza (Rio de Janeiro State University, BRA), M. S. De Magalhães (Brazilian Institute of Geography and Statistics, BRA)

Statistical process control is now faced with processes whose quality characteristics are better represented by profiles. For instance, these profiles can be images of a product. Frequently, a profile can be modeled as a continuous functions or, when discretized, as a vector in high dimensional euclidean space. When considering Phase I of a control chart for profile, one can be challenged by the huge amount of high dimensional data points. This dataset has to be analysed in order to establish which products (datapoints) have been produced under an in control stable production process, subject only to common cause small random variation, and which products have been subjected to special cause variation. This distinction, for high dimensional data coming from profiles, requires new techniques of analysis. Here we present a technique based on diffusion maps and Markov chains, using new ideas to analyse such complex data.

Keywords: Structure in Phase I Data Sets; Profiles; Diffusion Maps; Clusterization; Diffusion Distance.

Pseudo-Precision and Rank Tests

M. Neuhäuser (RheinAhrCampus Remagen, DEU)

Nonparametric, or distribution-free, methods can be applied in statistical process control problems. For example, the Wilcoxon rank-sum statistic is used, see e.g. Mukerjee and Chakraborti (Qual. Reliab. Engng. Int. 2012; 28: 335-352). When a variable is measured by automated equipment, or when values are computed, the data are usually stored with pseudo-precision, that is with many more decimal places than justified by the precision of the respective measurement. This pseudo-precision matters when assigning ranks which is a necessary step when the Wilcoxon or any other rank-based statistic is applied. The resulting artificial reduction in the number of ties is a disadvantage because mean ranks give more efficient tests in comparison to randomly broken ties. This finding is demonstrated using asymptotic results and simulations, and illustrated using example data.

Keywords: Nonparametric Process Control; Ties; Wilcoxon Statistic.

Monitoring of the Multi-stage Manufacturing Process Using the Bayesian Structural Equation Modeling Approach

C. Park (Chung-Ang University, KOR)

The structure of the quality at each stage of the multi-stage manufacturing process is expressed as a structural equation model (SEM). In the SEM, the quality level is regarded as the latent variable (LV) which is denoted as a linear function of the previous quality estimates and the current state input variables, and the observed variables of the process are regarded as indicators of the LV. One powerful method for estimating the quality level as well as estimating the regression coefficients is to use the Bayesian SEM approach where the LVs are treated as missing values. The LVs and the structure parameters are estimated by the Bayesian MCMC method using the Gibbs sampler. The monitoring of the process quality can be done in two ways: Direct monitoring of the LVs and Profile monitoring of the regression parameters. This research shows the monitoring procedures for the process quality under the multi-stage process structure. Directing monitoring is conveniently used for out-of-control signal, and the profile monitoring can be used in detecting the source of the variation. The advantage of the Bayesian SEM approach will also be discussed when compared to the state space model (SSM) approach, which is most frequently used in practice so far.

Keywords: Multi-stage Manufacturing Process; Structural Equation Modeling; Latent Variable; Indicator; Bayesian MCMC Method; Profile Monitoring; State Space Model.

Applications of Multivariate Statistical Process Control in Non-Industrial Processes

S. Psarakis (Athens University of Economics and Business, GRC), A. Sgora (University of Piraeus, GRC), S. Bersimis (University of Piraeus, GRC)

Statistical process control (SPC) consists of a set of techniques originated in industry to keep manufacturing processes under control. Because a large number of characteristics or variables are collected, interdependency among variables is expected and hence the variables are correlated. As a result, multivariate statistical process control techniques received increased attention in recent decades. These techniques have been used in a variety of service processes. In this paper we discuss the basic procedures for the implementation of multivariate statistical process control techniques for non-industrial processes. After a brief review of Multivariate SPC techniques we review and classify into groups non-industrial applications of

Multivariate SPC according to the type of the problem to which the technique has been applied.

Keywords: Multivariate Control Charts; Non-Industrial Processes.

Dynamic Screening: An Approach for Expanding the SPC Applications

P. Qiu (University of Florida, USA), D. Xiang (East China Normal University, CHN)

In our daily life, we often need to identify individuals whose longitudinal patterns are different from the patterns of those well-functioning individuals, so that some unpleasant consequences can be avoided. In many such applications, observations of a given individual are obtained sequentially, and it is desirable to have a screening system to give a signal as soon as possible after that individual's longitudinal pattern starts to deviate from the regular pattern so that some adjustments or interventions can be made in a timely manner. For such applications, the conventional SPC charts cannot be applied directly because the in-control distribution of the quality variables would change over time. In this talk, we discuss a dynamic screening system proposed recently for that purpose, using SPC and longitudinal data analysis techniques. This method is demonstrated using a real-data example about the SHARe Framingham Heart Study.

Keywords: Correlation; Dynamic Screening; Longitudinal Data; Process monitoring; Process Screening; Standardization; Statistical Process control; Unequal Sampling Intervals.

Controlling Processes of Generally Inflated Poisson Counts

A. C. Rakitzis (Université de Nantes & IRCCyN UMR CNRS 6597), P. Castagliola (Université de Nantes & IRCCyN UMR CNRS 6597, FRA), P. E. Maravelakis (University of Piraeus, GRC)

In this work, we propose and study a two-parameter extension of the zero-inflated Poisson distribution. The proposed model can be used for modeling count data in which there is an excessive number of values, other than zero. Due to its flexibility, it can be also used for describing various types of unusual count data. A general inflated Poisson distribution is introduced first and then we derive as a special case the two-parameter model. The first parameter is the mean of the ordinary Poisson distribution while the second parameter controls the degree of inflation in the zero and non-zero values of the distribution.

For the monitoring of processes that are described according to the proposed general inflated Poisson distribution, we propose three control charts: (a) a Shewhart-type one, (b) a runs-rules based one and (c) a CUSUM-type one. All schemes are upper-sided and suitable for the detection of increasing shifts in the expected number of counts. Aspects of their statistical design are discussed and their performance is compared under various out-of-control situations. Moreover, changes in both parameters of the process are considered. Finally, a practical example based on real data is also provided.

Keywords: Average Run Length (ARL); CUSUM Control Charts; Inflated Distribution; Poisson Distribution; Runs Rules; Statistical Process Control.

A Reversible Jump Markov Chain Monte Carlo Approach to the Multivariate Change-point Problem

S. E. Rigdon (Saint Louis University, USA), R. Steward (Saint Louis University, USA), R. Pan (Arizona State University, USA)

We consider a p dimensional stochastic process of length N where the random component follows a multivariate normal distribution with an unknown but constant covariance structure. Furthermore, we assume the first T observations are distributed around a common mean vector. At this point the mean vector then undergoes a shift in exactly d dimensions where d is unknown. The remaining N - T observations are then normally distributed around this new mean vector. The goal of our analysis is to simultaneously estimate (1) the set of variables whose means have changed, (2) the change-point location, (3) the mean vectors both before and after the change-point, and (4) the covariance matrix. We employ MCMC methods in a Bayesian analysis that subsumes all of the above problems. This method requires us to simultaneously account for the detailed balance equation while allowing comparison between models undergoing different dimensional changes in the mean vector. We demonstrate how a reversible jump Markov Chain Monte Carlo (RJMCMC) approach satisfies these requirements and under suitable conditions provides accurate parameter estimation.

Keywords: Change-point; Bayesian Analysis; Markov Chain Monte Carlo; Multivariate Statistical Process Control.

Statistical Process Control From a Sequential Bayesian Standpoint

G. J. Ross (University College, GBR)

Statistical process control (SPC) is concerned with detecting unexpected changes in the behavior of a monitored process, so that corrective action can be taken. SPC techniques often make the following assumptions a) that sufficient observations are available from the process to allow all model parameters to be estimated accurately prior to the start of monitoring b) that the user of SPC techniques has the ability to stop the process for investigation once a potential change has been found, and c) that all changes should be considered equally important regardless of their magnitude. Although these assumptions are satisfied in some contexts, there are many others where they are not, particularly in recent attempts to apply SPC techniques in domains outside the traditional setting of industrial manufacturing processes. We develop a Bayesian framework for SPC which allows it to be carried out without the need for any of the above assumptions. By phrasing SPC in terms of multiple change point detection, we show how modern techniques for Bayesian computation can be used to carry out sequential inference which generalizes traditional Phase II SPC analysis. A variety of different algorithms are provided corresponding to different problem settings, and a number of examples are given to illustrate their use.

Keywords: Phase II; Change Detection; Bayesian Methods.

The c-chart with Bootstrap Adjusted Control Limits to Guarantee Conditional Performance

A. G. Ryan Driscoll (Virginia Tech, USA), M. Zhao (Virginia Tech, USA), W. H. Woodall (Virginia Tech, USA)

The integrity of Phase II control charting depends on the accuracy of Phase I estimation. Studies have shown that extremely large sample sizes are needed in Phase I to ensure that performance of control charts with estimated parameters is comparable to the performance of charts with known parameters. The sample size recommendations can be impractical for attribute control charts. We assess the in-control performance of the c-chart with an estimated number of non-conforming items. We show that the sampling variability associated with estimation results in a high percentage of control charts with in-control average run lengths well below that of corresponding control charts with known parameters. This sampling variability can be thought of as between-practitioner variability. To overcome the variability in performance, a c-chart with bootstrapped control limits is recom-

mended. A simulation study reveals that these adjusted bootstrapped control limits guarantee that the conditional average run length performance of the control chart meets or exceeds a predetermined level of performance with a specified probability. The out-of-control performance of the c-chart with adjusted limits is also discussed.
Keywords: AARL; Average of ARL; Bootstrap; c-Chart; Estimation Effect; SDARL; Standard Deviation of Average Run Length; Statistical Process Control.

Statistical CUSUM Designs with Minimum Sampling Cost

E. M. Saniga (University of Delaware, USA), D. Davis (University of Delaware, USA), T. McWilliams (Drexel University, USA), J. Lucas(J.Lucas and Associates, USA)

We design a CUSUM control chart for controlling a process mean so that the cost of sampling is minimized. Constraints are placed on the average time to signal a process shift when an assignable cause occurs and when the process remains in control. These are called the average time to signal out of control and average time to signal in control, respectively, or ATS0 and ATS1. Statistical control chart design with minimum sampling cost is a design method which can be thought to be between economic statistical design and statistical design. In economic statistical design we find the control chart design such that costs are minimized while statistical and other constraints are met. One practical shortcoming of economic statistical design is that a number of cost and system parameters must be estimated. This design method dominates statistical design in that an economic statistical design is less expensive and still meets specified statistical constraints. Further, Saniga (1989) has shown that one can "tighten" statistical designs and reduce cost, a counterintuitive result. One shortcoming of statistical design is the determination of the average time to signal and the sample size. In a CUSUM statistical design, one usually specifies n and the in and out of control average run lengths (ARL0 and ARL1) Then a search is made for k and h such that specified ARL constraints are met. Theoretically, Reynolds and Stombous (2004) have suggested that an optimal procedure for controlling a sample mean is a CUSUM (or EWMA) chart with n=1. Practically speaking, though, some n=1 designs may not be feasible in meeting ARL constraints. Further, costs of n=1 designs may be far from optimal (see, e.g. Saniga, et al (2006a, 2006b)). Finally, the determination of an optimal intersample interval is complex. A solution to the shortcomings of economic statistical design and statistical design is a design that minimizes sampling cost and meets the ATS (and other) constraints. This design may be employed in situations where one may not be able to estimate the cost and system parameters to determine an economic statistical design. This design that minimizes sampling cost dominates a pure statistical design in the same way

that an economic statistical design dominates a pure statistical design. That is, a design that minimizes sampling cost meets statistical and other constraints at minimum sampling cost. In addition, as in economic statistical design, we find examples where "tightened" statistical designs are cost minimal. Designs that address the sampling cost issue have been proposed by Weiler(1952) and by Page (1954) . In the former study a design is found with a fixed critical value that finds n such that nARL0 is minimum. The latter finds the same design except that the critical value can vary as well as n. Note that in each case one finds a design that minimizes the number sampled until a true signal occurs. Also, note that neither addresses the issue of true sampling cost, which has fixed and variable components We consider a number of problems based upon an example of Lorenzen and Vance (1986) and for each of these examples we find the minimum sampling cost design, the statistical design assuming n=1 and the optimal intersample interval can be found such that cost is minimized, and the economic statistical design. Our results indicate that a minimum sampling cost design dominates pure statistical design much as economic statistical design dominates pure statistical design. Of course, the minimum sampling cost design is dominated by the economic statistical design but it has the advantage of not requiring estimation of a large number of cost and system parameters. One other interesting finding is that as the fixed cost of sampling increases the optimal k and h of the minimum sampling cost design approach values that are similar to that of an optimal X bar chart design (k=3, h=0), a finding similar to that of Saniga, et al (2006, 2006, 2012). The implication is that in these situations one can employ an easier to implement X bar control chart rather than a CUSUM chart with minimal additional cost.

References

Lorenzen T. J. and L.C. Vance(1986), "The Economic Design of Control Charts: A Unified Approach", Technometrics 28, pp. 3-10.

Page, E. S., (1954) " Control charts for the Mean of a Normal Population", Journal of the Royal Statistical Society, Series B, Vol. 16, No.1, pp. 131-135

Reynolds, M. Jr. and Z Stoumbos, (2004)" Control Charts and the Efficient Allocation of Sampling Resources", Technometrics, Vol.46, 200-214

Saniga, E. M. (1989). " Economic Statistical Design of Control Charts with an Application to X bar and R Charts", Technometrics 31, 313-320.

Saniga, E. M., Thomas McWilliams, Darwin Davis and James Lucas, (2006), " Economic advantages of CUSUM Control Charts for Variables, Frontiers in Statistical Quality Control, Vol. 8, H-J Lenz and P-T Wilrich, eds. Physica Verlag pubs,

Saniga, E.,T. McWilliams, D.Davis and J. Lucas , (2006b)"Economic Control Chart Policies for Monitoring Variables", Int.J. Productivity and Quality Management, , Vol.1, No.1, 2006, , pp. 116-138

Erwin Saniga, James Lucas , Darwin Davis, Thomas McWilliams, (2012)"Economic Control Chart Policies for Monitoring Variables when there are Two Components of Variance", Frontiers in Statistical Quality Control., eds. Lenz, Wilrich and Schmid, Volume 10, pp. 85-95,

Weiler, H.,(1952) “ On the Most Economical Sample Size for Controlling the Mean of a Population” , The Annals of Mathematical Statistics, Vol. 23, No.2, PP. 247-254
Keywords: CUSUM Control Charts; Minimum Sampling Cost; Statistical Design.

Surveillance of Non-Stationary Processes

W. Schmid (European University Frankfurt (O), DEU), T. Lazariv (European University Frankfurt (O), DEU)

In nearly all papers on process control for time dependent data it is assumed that the underlying in-control process is stationary. In the present paper the in- control process is modelled by a multivariate state-space model which may be non- stationary. The parameter of interest is its mean. The likelihood ratio method, the sequential probability ratio test and the Shiryaev-Roberts procedure are applied to derive control charts. These procedures depend on certain reference values. The corresponding generalized approaches are considered as well and generalized control charts are calculated. They do not have a further design parameter. In an extensive simulation study the behaviour of the introduced schemes is compared with each other using various performance criteria like the average run length, the average delay, and the probability of a successful detection.
Keywords: Control Charts; Statistical Process Control; Change-Point Detection; Time Series; State-Space Model.

Monitoring Communication Counts- Detecting Communication Outbreaks for a Sub-Group of a Larger Target Group of People

R.S. Sparks (CSIRO, AUS), W. H. Woodall (Virginia Tech, USA), J. D. Wilson (University of San Francisco, USA)

We discuss scalable methods of monitoring a target group ranging from 100 to 1000 people, with the aim of detecting small sub-groups (4 to 10 people) with sudden significantly elevated communication levels. The aim is to detect the increased activity as early as possible for a given in-control false alarm rate as well as identify which sub-group of people are involved with the increased activity. The approach will differentiate between two types of outbreaks. The first is where all members in the sub-group increase the number of communications with each other and

there is no communication out-of-control "sparsity" within the sub-group to exploit (no members fail to increase their communication counts). The second is when the sub-group has a dominant leader and most of the communication increase for the team is between the dominant leader and sub-group members while only a few sub-group members simultaneously increase their level of communication. The majority of sub-group members don't change their communication patterns with others in the sub-group so here there are sparse communications within the subgroup that are worth exploiting in developing an efficient surveillance plan. The approach will be first to start deriving plans when we assume that in-control communication counts between any two individuals have the same distribution (the homogeneous case), before using this to derive a plan for the more realistic case when in-control communication counts are inhomogeneous. The detection properties of the two approaches will be discussed and compared under various scenarios and recommendations will be given on their appropriate implementation.

Keywords: Crime; Dominant Leader; EWMA; ARL.

Regulating the Bias-Variance Tradeoff in Analysis of Survey Data over Time with Application to a Customer Loyalty Survey

S. H. Steiner (University of Waterloo, CAN), T. Cooper-Barfoot (University of Waterloo, CAN), J. MacKay (University of Waterloo, CAN)

In the analysis of survey data, common objectives include estimating the population average, tracking time trends, and comparing population subgroups. When samples are taken over time, we can estimate using only the present time data or also include historical data. However, when the characteristic is drifting over time and sample sizes are small, the decision to include historical data trades precision for bias. We propose regulating the bias-variance tradeoff using Weighted Estimating Equations based on a suitable Generalized Linear Model that incorporates covariates. A customer loyalty survey for a smartphone vendor will be presented and resulting present time estimates of Net Promoter Score will be compared across various approaches applied to example data and simulated data.

Keywords: Bias-variance Tradeoff; Weighted Estimating Equations; EWMA; Net Promoter Score.

Monitoring Multinomial Attribute Processes Based on Weighted Sum of Chi Squares

H. Taleb (University of Gafsa, TUN), E. Khediri (University of Tunis, TUN)

When monitoring multinomial processes using chi-square control charts, an-out-of Control situation doesn't give an idea about of which category is responsible. A new approach is proposed to monitor multinomial processes based on general gamma distribution. The proposed statistic is a weighted sum showing the relative importance of all categories. The proposed statistic can be approximated by the gamma distribution. The proposed procedure is applied to a real example and compared to chi square control chart. A sensitivity analysis of the proposed control chart is given.

Keywords: Multinomial; Chi-square; Average Run Length.

A Fully Adaptive Multivariate Control Scheme for Simultaneous Monitoring of Mean Vector and Covariance Matrix of Processes Subject to a Multiplicity of Assignable Causes

K. A. Tasias (University of Western Macedonia, GRC), G. Nenes (University of Western Macedonia, GRC)

In today's process applications, the simultaneous monitoring of multiple correlated quality characteristics is crucial. This paper presents a new fully adaptive Phase II multivariate statistical process control (m-SPC) scheme for monitoring processes where multiple assignable causes may occur. The assignable causes are independent and affect both the mean vector and the covariance matrix, which are monitored by a T^2 control chart and a multivariate Shewhart control chart based on differential entropy, respectively. A model based on Markov chain theory is developed for the stochastic operation of the proposed scheme in order to find its economic-statistical design (ESD). A wide benchmark of examples has been generated to demonstrate the economic and statistical superiority of the proposed model against simpler approaches.

Keywords: Multivariate Processes; Correlation; Multiple Assignable Causes; Hotelling's T^2 control chart; Differential Entropy.

Bayesian SPC for Count Data

P. Tsiamyrtzis (Athens University of Economics and Business, GRC), D. M. Hawkins (University of Minnesota, USA)

Count data (usually modeled by a Poisson distribution), are traditionally monitored by frequentist u/c chart, CUSUM and EWMA. These charts require the gathering of a large phase I sample to obtain a reliable "in-control" parameter estimate. Self-starting proposals that ameliorate the need for a large phase I sample have also proposed. All these methods are frequentist, they allow retrospective inference during phase I and they have no coherent way to incorporate available prior information about the process. In this work, we introduce a Bayesian self-starting procedure, which can incorporate prior information, allow online inference and is particularly attractive for short-run settings and phase I type analysis. Details about the modeling will be presented and the new proposal will be tested against frequentist-based alternatives in a real data example.

Keywords: Bayesian SPC by Attributes; Self-starting; Gamma Mixture; Change-point.

Evolution of Big Data Analytics

K. L. Tsui (City University of Hong Kong, HKG)

Due to the advancement of computation power and data storage and collection technologies, the field of data modelling and applications have been evolving rapidly over the last two decades, with different buzz words as machine learning, knowledge discovery in databases (KDD), data mining (DM), business analytics, big data analytic, ... There are tremendous opportunities in interdisciplinary research and education in data science, system informatics, and big data analytics; as well as in complex systems optimization and management in various industries of finance, healthcare, transportation, and energy, etc. In this talk we will present our views and experience in the evolution of big data analytics, challenges and opportunities, as well as applications in various industries.

Keywords: Big Data Analytics; System Informatics; Business Analytics.

Quality Engineering Faces the Challenges of Big Data and Little Data

F. Tsung (Hong Kong University of Science and Technology, HKG)

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

Keywords: Quality Engineering; Data Analytics; 3D Printing.

A New Control Chart for Monitoring Nonlinear Profiles

J. A. Vargas (Universidad Nacional de Colombia, COL), R. D. Guevara (Universidad Nacional de Colombia, COL), N. Acevedo (Universidad Nacional de Colombia, COL)

Profile monitoring has received considerable attention in a growing number of statistical process control applications. We propose a control chart based on a depth measure, for monitoring nonlinear profiles. Control charts based on depth measures offer more flexibility than traditional nonlinear profile monitoring methods, because it is not necessary to impose some restrictive assumptions. A comparison study with other established control charts is carried out via simulations.

Keywords: Depth measures ; Functional data; Profile monitoring.

An Overview and Perspective on Social Network Monitoring

W. H. Woodall (Virginia Tech, USA), M. Zhao (Virginia Tech, USA), R.S. Sparks (CSIRO, AUS), J. D. Wilson (University of San Francisco, USA)

In this expository paper we give an overview of statistical methods for the monitoring of social networks. We discuss the advantages and limitations of various methods and discuss some relevant issues. Relationships are given between network monitoring methods and related monitoring methods in engineering statistics and public health surveillance. We encourage researchers in the industrial process monitoring area to work on developing and comparing the performance of social network monitoring methods. We give a number of research ideas.

Keywords: Bayesian Process Monitoring; Scan Methods; Social Network Analysis; Social Network Change Detection; Statistical Process Monitoring.

Some Thoughts on SPC Applications in Reliability and Maintenance Monitoring

M. Xie (City University of Hong Kong, HKG), Y. Cheng (City Univ of Hong Kong, HKG)

Recently there has been an increasing use of SPC in reliability related studies. In this talk, we will describe some of our recent works on SPC for time-between-events. The methods have been applied to various reliability related monitoring. For failure monitoring in reliability engineering, both the occurrence frequency and failure severity/magnitude should also be modelled and studied. The dependence is a difficult issue. There are also a number of other challenges, especially when the maintenance is incorporated into the modelling, and they will be discussed in this talk as well.

Keywords: Control charts; Reliability; Maintenance; Process Monitoring; Weibull Distribution; Gamma Distribution.

An ARL-Unbiased EWMA Proportion Control Chart

S. Yang (National Chengchi University, TWN)

Control charts are effective tools for out-of-control detection in both manufacturing and service processes. As much of the data in industries come from processes having non-normal or unknown distributions, the commonly used Shewhart variable control charts cannot be appropriately used, because they depend heavily on the normality assumption. The Average Run Length (ARL) is generally used to measure the detection performance of a process when using a control chart, but it is biased for the monitoring statistic with an asymmetric distribution - that is, the ARL-biased control chart leads to some out-of-control ARL values that are larger than the in-control ARL, hence taking longer to detect the shifts in parameter than to trigger a false alarm. To overcome this problem we herein propose an ARL-unbiased EWMA proportion chart to monitor the process mean for process data with non-normal or unknown distributions. We further explore the procedure to determine the control limits and to investigate the out-of-control mean detection performance of the ARL-unbiased EWMA proportion chart. Further, we compare the out-of-control detection performance of the ARL-unbiased EWMA proportion chart, the arcsin transformed symmetric EWMA sign chart and other existing mean charts. The proposed ARL-unbiased EWMA proportion chart shows superior mean detection performance. Thus, we recommend the ARL-unbiased EWMA proportion chart for process data with non-normal or unknown distributions.

Keywords: Mean chart; Non-normal Distribution; Unbiased Average Run Length.

Alarm Attributes in Systems for Monitoring Quality and Reliability data

E. Yashchin (IBM Research, USA)

Modern monitoring systems involve massive analysis of data streams geared towards speedy detection of unfavorable process conditions, while maintaining a low rate of false alarms. Even for well-designed systems, however, one can expect to see, at any point in time, a large number of alarm conditions. A key question is one of alarm prioritization, which is needed to direct attention of various groups of users towards problems that are most relevant from their perspective. Alarm attributes serve as instruments for such prioritization. In this paper we introduce several types of alarm attributes, discuss their properties and illustrate their application in early warning systems for supply chain data.

Keywords: Statistical Process Control; Change-Point Theory; Sequential Analysis; Likelihood Ratio tests; Filtering.

High-Dimensional Statistical Process Control

C. Zou (Nankai University, CHN), L. Feng (Northeast Normal Univerisity, CHN)

Monitoring high-dimensional data streams has become increasingly important for real-time detection of abnormal activities in many statistical process control (SPC) applications. Although the multivariate SPC has been extensively studied in the literature, the challenges associated with designing a practical monitoring scheme for high-dimensional processes when between-streams correlation exists are yet to be addressed well. Classical T-square-based schemes do not work well because the contamination bias in estimating the covariance matrix grows rapidly with dimension. We propose a test statistic which is based on the "divide-and-conquer" strategy, and integrate this statistic into the multivariate EWMA charting scheme for Phase II process monitoring. The key idea is to calculate the T-square statistics on low-dimensional sub-vectors and combine them together. The proposed procedure is essentially distribution-free and computation and storage efficient. The control limit is obtained through the asymptotic distribution of the test statistic under some mild conditions on the dependence structure of the stream observations. Our asymptotic results also shed light on quantify the size of a reference sample requires. Both (asymptotically) theoretical analysis and numerical results show that the proposed method is able to control the false alarm rate and deliver robust change detection.

Keywords: Asymptotic Normality; Exponentially Weighted Moving Average; High-dimensional Data Streams; Weak Dependence Structure.

Another Look at the EWMA Control Chart with Estimated Parameters

I. M. Zwetsloot (University of Amsterdam, NLD), N. A. Saleh (Cairo University, EGY), M.A. Mahmoud (Cairo University, EGY), L.A. Jones-Farmer (Miami University, USA), W.H. Woodall (Virginia Tech, USA)

When in-control process parameters are estimated, Phase II control chart performance will vary among practitioners due to the use of different Phase I data sets. The typical measure of Phase II control chart performance, the average run length (ARL), becomes a random variable due to the selection of a Phase I data set for estimation. Aspects of the ARL distribution such as the standard deviation of the average run length (SDARL) can be used to quantify the between-practitioner variability in control chart performance. In this article/presentation, we assess the in-control performance of the exponentially weighted moving average (EWMA) control chart in terms of the SDARL and percentiles of the ARL distribution when the process parameters are estimated. Our results show that the EWMA chart requires a much larger amount of Phase I data than previously recommended in the literature in order to sufficiently reduce the variation in the chart performance. Since it could be extremely difficult to lower the variation in the in-control ARL values sufficiently due to practical limitations on the amount of the Phase I data, we recommend an alternative design criterion and a procedure based on the bootstrap approach.

Keywords: Bootstrap; Estimation Effect; SDARL.

Author Index

Acevedo, N. , 32
Amiri, A. , 5
Aparisi, F. , 17
Ayoubi, M. , 5

Babu, Y. , 20
Bastos, S. F. , 12
Bersimis, S. , 5, 22

Castagliola, P. , 6, 7, 23
Celano, G. , 7
Chakraborti, S. , 7
Chattinnawat, W. , 8
Cheng, Y. , 33
Colosimo, B. M. , 8
Cooper-Barfoot, T. , 29
Costa, A. F. B. , 9

Davis, D. , 26
De Magalhães, M. S. , 9, 21
Does, R. J. M. M. , 11

Epprecht, E. K. , 12
Erto, P. , 13

Faraz, A. , 13
Feng, L. , 35
Figueiredo, A. M. , 14
Figueiredo, F. O. , 14
Fricker Jr., R. D. , 15

George, J. , 20
Gomes, M. I. , 14
Graham, M. A. , 7
Green, C. B. , 8
Guevara, R. D. , 15, 32

Haridy, S. , 16
Hawkins, D. M. , 16, 31
Heuchenne, C. , 13
Ho, L. L. , 17

Jones-Farmer, L. A. , 17, 35

Khediri, E. , 30
Kim, S. B. , 18
Kim, Y. H. , 18
Knoth, S. , 19

Lazariv, T. , 28
Lombard, F. , 16
Lucas, J. , 26

Maboudou, E. M. , 19
Machado, M. A. G. , 9
MacKay, J. , 29
Mahmoud, M. A. , 35
Maman, Z. S. , 20
Maravelakis, P. E. , 6, 20, 23
Martinez, W. , 17
McWilliams, T. , 26
Megahed, F. , 17
Megahed, F. M. , 20
Moura Neto, F. D. , 21

Nazir, H. Z. , 11
Nenes, G. , 7, 30
Neuhäuser, M. , 21

Pagani, L. , 8
Pan, R. , 24
Park, C. , 22
Psarakis, S. , 5, 22

Qiu, P. , 23

Rakitzis, A. C. , 6, 23
Rigdon, S. E. , 24
Ross, G. J. , 25
Ryan Driscoll, A. G. , 25

Saleh, N. A. , 35
Saniga, E. M. , 26
Schmid, W. , 28
Sgora, A. , 5, 22
Silva, P. N. , 9

Simões, F. D. , 9
Sogandi, F. , 5
Souza, P. A. S. , 21
Sparks, R. S. , 28, 33
Steiner, S. H. , 29
Steward, R. , 24

Taleb, H. , 30
Tasias, K. A. , 30
Tsiamyrtzis, P. , 31
Tsui, K. L. , 31
Tsung, F. , 32

Vargas, J. A. , 15, 32
Von Doellinger, R. S. , 9

Weese, M. , 17
Wilson, J. D. , 28, 33
Woodall, W. H. , 13, 25, 28, 33, 35

Xiang, D. , 23
Xie, M. , 33

Yang, S. , 34
Yashchin, E. , 34

Zhao, M. , 25, 33
Zou, C. , 35
Zwetsloot, I. M. , 35