Preface

The XIth International Workshop on *Intelligent Statistical Quality Control* took place in Sydney, Australia from August 20 to August 23, 2013. It was hosted by Professor Ross Sparks, CSIRO Mathematics, Informatics and Statistics, North Ryde, Australia. The invitational workshop was jointly organized by Professors S. Knoth, W. Schmid, and R. Sparks. The 23 papers in this volume were carefully selected by the scientific program committee, reviewed by its members, revised by the authors and, finally, adapted by the editors for this volume.

The focus of the book lies on three major areas of statistical quality control: statistical process control (SPC), acceptance sampling and design of experiments. The majority of the papers deal with statistical process control while acceptance sampling, and design of experiments are treated to a lesser extent.

The book is divided into four parts. Subject of Part I is statistical process control. Part II is devoted to acceptance sampling. Part III covers the design of experiments, while in Part IV related fields are considered.

Part I: Statistical Process Control

Social networks are increasingly attracting the attention of academic and industry researchers. Monitoring communications between clusters of suspicious individuals is important in flagging potential planning activities for terrorism events or crime. Governments are interested in methodology that can forewarn them of future terrorist attacks or social uprisings in disenchanted groups of their populations. In the paper of **Sparks** a range of approaches is examined that could be used to monitor communication levels between suspicious individuals.

Woodall and **Driscoll** deal with the monitoring of a rare event. A review of some recent results is given and several new approaches are offered. Because some of the competing monitoring procedure have implicit headstart features, there are compelling arguments for the use of steady-state performance metrics. The choice of an appropriate performance metrics is discussed in detail. The strong adverse effect of Phase I parameter estimation on Phase II performance of various charts is summarized. In addition, the important practical issue of the effect of aggregation of counts over time, some generalizations of standard methods, and some promising research ideas are discussed.

Big data is a popular term that is used to describe the large, diverse, complex, and/or longitudinal datasets generated from a variety of instruments, sensors, and/or computer-based transactions. In **Megahed** and **Jones-Farmer** several big data applications are discussed to highlight the opportunities and challenges for applied statisticians interested in surveillance and statistical process control. The goal of the authors is to bring the research issues into better focus and encourage methodological developments for big data analysis in these areas.

Epprecht presents a survey of the research on techniques for the statistical control of industrial multiple-stream processes. These are processes in which the same type of item is manufactured in several streams of output in parallel, or still continuous processes in which several measures are taken at a cross section of the product. This paper seems to be the first literature review on this topic. Essential differences in the underlying models are stressed and issues for further research are pointed out.

Yashchin considers a unified methodology based on the use of likelihood ratio tests to monitor processes. This approach leads to control schemes that provide good statistical performance and are easy to implement. They depend on just one design parameter and require a limited computational effort that is dynamically adjusted based on the process conditions. An example pertaining to multivariate control of the normal mean is discussed in detail.

Lazariv and **Schmid** give an overview about variance control charts for time dependent processes. In their paper, they consider charts based on the likelihood ratio approach and the generalized likelihood ratio approach, the sequential probability ratio method and the generalized sequential probability ratio procedure, the Shiryaev–Roberts procedure and a generalized Shiryaev–Roberts approach, and different types of exponentially weighted moving average (EWMA) charts. Within an extensive simulation study, these schemes are compared with each other. In order to measure the performance of the schemes, the average run length and the average delay are used.

Controlling both increases and decreases in a parameter by using a control statistic with an asymmetrical distribution, frequently leads to an ARL-biased chart. This means that some out-of-control ARL values are larger than the in-control ARL. **Knoth** and **Morais** discuss the idea of ARL-unbiased charts, provide instructive illustrations of ARL-(un)biased charts, relate ARL-unbiased Shewhart charts with the notions of unbiased and uniformly most powerful unbiased tests, and briefly discuss the design of EWMA charts not based on ARL(-unbiasedness).

The paper of **Su**, **Gan**, and **Tang** concerns with cumulative sum (CUSUM) charts. If the density of the in-control process is unknown, they propose to estimate its density by a kernel density estimator. The performance of this chart is investigated for unimodal distributions. The obtained results reveal that this chart works well if a sufficient number of observations of the in-control process are available.

In **Yang** and **Arnold**, EWMA control charts based on the process proportions and an arcsin transformation are proposed to monitor the process mean and variance simultaneously. The sampling properties of the new monitoring statistics are analyzed. EWMA recursions are applied to these quantities. The behavior of the new schemes is analyzed by making use of its average run lengths.

Capizzi and **Masarotto** concern with the application of variable selection (VS) algorithms for monitoring multivariate data. These charts share the common idea that process faults usually affect a small fraction of the monitored quality characteristics. Hence, VS methods can be used to identify the subset of the variables for which the shift may have occurred. However, the suggested VS-based control charts differ in many aspects such as the particular VS algorithm and the type of the control statistic. In this paper, some VS-based control charts are compared with each other in a variety of out-of-control scenarios.

Göb and **Lurz** use an extension of the Camp-Meidell inequality to determine the control limits of a Shewhart chart. This procedure does not make use of any distributional assumption but only needs the existence of moments higher than 2. It is discussed how the moments in the bounds can be estimated from a Phase I sample. Appropriate estimators, their properties, and the effect of estimation on the properties of the process monitoring charts are investigated. In particular, the use of empirical Camp-Meidell bounds in quantile control charts is studied.

When simultaneous schemes are used, the quality characteristic is deemed to be out of control whenever a signal is triggered by either individual chart. **Morais**, **Ramos**, and **Pacheco** deal with the problem of misleading signals (MS), meaning that a shift in the process mean can be misinterpreted as a shift in the process variance and vice versa. Conditions are discussed to achieve values for the probability of a misleading signal smaller than or equal to 0.5 and alternative simultaneous Shewhart-type schemes are explored.

Saniga, **Davis**, **Faraz**, **McWilliams**, and **Lucas** investigate the characteristics of economic control chart designs for both Shewhart and CUSUM control charts. Authors in the past have made some suggestions regarding the design of these charts, where design is defined as finding the values of sample size, intersample interval and control limit (Shewhart), or control parameters for the CUSUM chart. In the present paper, the authors run a large number of experiments consisting of many configurations of the parameters and describe and model the results in terms of the actual economic designs.

Hryniewicz analyzes SPC procedures when the quality parameters of interest can be hardly directly monitored. Training data are used to build a model that is used for the prediction of the value of an unobservable variable of interest. In the paper, a model of a process is considered when traditionally applied assumptions are violated. In such a case, it is shown that some non-statistical prediction models proposed in the area of data-mining perform better than popular linear prediction models. However, new problems have to be considered when shifts in the levels of process parameters may influence the performance of applied classification algorithms.

Nowadays, there are many applications of SPC outside engineering, e.g., in public health and finance. **Di Bucchianicco** and **van den Heuvel** investigate to which extent modern statistical theory like hypothesis testing, prediction intervals, and tolerance intervals may be used to extend Shewhart's ideas in the above-mentioned wider range of application domains. Alternative settings of statistical control proposed in the literature including Bayesian settings are discussed as well.

Part II: Acceptance Sampling

Wilrich considers sampling plans which have besides the specification limit an additional limit. Such extended sampling plans are, e.g., used for the evaluation of bacterial contamination in foods, the amount of active ingredient used in formulating drug products, and the strength of concrete. The operating characteristic function of these extended sampling plans for inspection by variables is derived, and the advantages/disadvantages in comparison with unextended sampling plans are discussed.

The concept of a fractional acceptance number is particularly useful for short-run food manufacturing processes involving a measurable quality characteristic such as the percentage sugar or fat content. **Govindaraju** and **Jones** propose a new fractional acceptance number sampling plan, which is a mix of attribute and variables methods. The operating characteristics of the proposed plan are evaluated using common error distributions and the incomplete beta function.

Steland extends the methodology of acceptance sampling for variables with unknown distributions when additional sampling information is available to such settings. Based on appropriate approximations of the operating characteristic, new acceptance sampling plans are derived that control the overall operating characteristic. The results cover the case of independent sampling as well as the case of dependent sampling. In particular, a modified panel sampling design and the case of spatial batch sampling are studied. The latter allows to detect and analyze local clusters of degraded or damaged modules in photovoltaics.

Part III: Design of Experiments

Vining, **Freeman**, and **Kensler** point out that the reliability of products and processes will become increasingly important in the near future. The paper begins with a review of the current practice for planning reliability experiments. It then reviews some recent work that takes into proper account the experimental protocol.

A basic issue is that most reliability engineers have little training in planning experiments while most experimental design experts have little background in reliability data.

Hassler, **Silvestrini**, and **Montgomery** analyze Bayesian D-optimal designs, which have become computationally feasible to construct for simple prior distributions. They identify several concerns for DB-optimal designs. It is shown that some parameter values give rise to models that have little utility to the practitioner for effect screening. For some generalized linear models such as the binomial, inclusion of such models can cause the optimal design to spread out toward the boundary of the design space. This can reduce the D-efficiency of the design over much of the parameter space and result in the Bayesian D-optimal criterion's divergence from the concerns of a practitioner designing a screening experiment.

The Bayesian Lasso is a variable selection method that can be applied in situations where there are more variables than observations. Thus both main effects and interaction effects can be considered in screening experiments. To apply the Bayesian framework to experiments involving the effect heredity principle, which governs the relationships between interactions and their corresponding main effects, several initial tunings of the Bayesian framework are required. **Noguchi**, **Ojima**, and **Yasui** propose models that do not require the initial tuning values to be specified in advance.

Part IV: Related Areas

The time scale need not be chronological if a failure occurs based on the cumulative damages suffered from its usage or exposure to some risks. The true time scale need not be supported by the observed field reliability data. **Yamamoto** and **Takeshita** use simulations to investigate the properties of time scale models and the sample properties of their estimates. The estimator and the time scale functions are applied to a problem of finding a suitable time scale for field reliability data.

Lenz analyzes the phenomenon that the naive Bayesian classifier may dominate the proper one as happened in clinical studies. The reason for the dominance relation lies in a mix of an a-priori not fixed dimension of the state-space (symptom space) given a disease, the feature selection procedure, and the parameter estimation. Estimating conditional probabilities in high dimensions when using a proper Bayesian model can lead to an "over fitting," a missing value problem, and, consequently, to a loss of classification accuracy.

The level of a workshop on *Intelligent Statistical Quality Control* is determined by the quality of its papers. We believe that this volume truly represents the frontiers of statistical quality control. The editors would like to express their deep gratitude to the members of the scientific program committee, who carefully invited researchers from around the world and refereed all submitted papers:

Olgierd Hryniewicz, Poland
Sven Knoth, Germany
Yoshikazu Ojima, Japan
Wolfgang Schmid, Germany
Ross Sparks, Australia
Peter-Th. Wilrich, Germany
William H. Woodall, U.S.A.
Emmanuel Yashchin, U.S.A.

We would like to cordially thank our host at Sydney, Ross Sparks. He created a perfect atmosphere for inspiring and fruitful discussions. Moreover, we thank Springer, Heidelberg, for the continuing collaboration.

Hamburg, Germany — Sven Knoth
Frankfurt (Oder), Germany — Wolfgang Schmid
July 2014

http://www.springer.com/978-3-319-12354-7

Frontiers in Statistical Quality Control 11
Knoth, S.; Schmid, W. (Eds.)
2015, XXI, 393 p. 75 illus., 41 illus. in color., Hardcover
ISBN: 978-3-319-12354-7