9th Workshop on Quality Improvement Methods
at the Universitätskolleg Bommerholz

Bommerholzer Str. 60, 58456 Witten-Bommerholz
tel.: (0 23 02) 39 60 fax.: (0 23 02) 39 63 20

# ABSTRACTS

## A General Construction Method for Mixed-Level Supersaturated Designs

*Kalliopi Mylona (National Technical University of Athens)*

Abstract Supersaturated designs are a large class of factorial designs which can be used for screening out the important factors from a large set of potentially active variables. The huge advantage of these designs is that they reduce the experimental cost drastically, but their critical disadvantage is the high degree of confounding among factorial effects. Mixed-level factorial designs are experimental designs whose factors have different numbers of levels. These designs are often useful for experiments involving both qualitative and quantitative factors. The $E(f_{NOD})$-criterion is often applied for the construction and evaluation of supersaturated designs as it can be used as a measure of non-orthogonality. A general method for constructing $E(f_{NOD})$-optimal mixed-level supersaturated designs possessing the equal occurrence property is presented. The essential feature of this method is the use of infinite classes of supplementary difference sets, a fact which allow the construction of optimal mixed-level designs with good properties. Keywords and phrases: Resolvable design, supersaturated design, supplementary difference sets, uniform design.

## Geometric representations for random effects models

*Tim Holland-Letz (Ruhr-Universität Bochum))*

Clinical experiments in the area of pharmacokinetics generally try to estimate parameters of a non linear time-concentration curve. Usually, a random effects model is assumed, the so called population model with repeated measurements on several patients each. This situation poses methodological challenges for optimal experimental design, as the setup results both in partly correlated observations and a parameter dependant variance, thus violating two key assumptions of classical design theory. It is the aim of this talk to adapt and extend the existing design methods to cover some of these situations. As the most important pharmacokinetic parameters are summary statistics of several model parameters (e.g. the area under the concentration curve) the focus will be on c-optimal designs which allow optimal estimations of these quantities. In particular, the equivalence theory by Kiefer(1974) and Pukelsheim(1993) is applied to the specific situation and the geometric representation proposed by Elfving(1952) is generalized to allow for the situation described above.

# Determine support points of locally optimal designs for nonlinear models

*Min Yang (University of Missouri)*

Deriving optimal designs for nonlinear models is in general challenging. One crucial step is to determine the number of support points needed. Current tools handle this on a case-by-case basis. Each combination of model, optimality criterion and objective requires its own proof. The celebrated de la Garza Phenomenon states that under a $(p-1)$th-degree polynomial regression model, any optimal design can be based on at most $p$ design points, the minimum number of support points such that all parameters are estimable. Does this conclusion also hold for nonlinear models? In this talk, a new approach is introduced to address this question. With this new approach, it can be easily shown that such phenomenon exists for many commonly studied nonlinear models, such as the logistic and probit regression model, Emax model, exponential model, three- and four-parameter log-linear models, Emax-PK1 model, as well as many classical polynomial regression models. The proposed approach unifies and extends many well-known results in the optimal design literature. It has four advantages: (i) it can be applied to many forms of nonlinear models; to continuous or discrete data; to data with homogeneous or non-homogeneous errors; (ii) it can be applied to any design region; (iii) it can be applied to multiple-stage optimal design; and (iv) it can be easily implemented.

# EMO Stopping Criteria: Detecting Stagnation in Multi–Objective Optimization

*Luis Martí (Universidad Carlos III de Madrid)*

Most soft–computing, heuristic or non-deterministic methods need a stopping criterion. The stopping criterion, which is usually a heuristic itself, is responsible for minimizing the wastage of computational resources by detecting scenarios where it makes no sense to continue executing the method. Evolutionary multi–objective optimization (EMO) approaches also need a stopping criterion. Paradoxically, this is a matter that has often been overlooked by the community, probably because it plays a supporting part. Consequently, the theoretical and practical implications concerning this topic have not yet been properly explored. Indeed, many real–world applications of theoretically outstanding methods may have underperformed due to an incorrect algorithm termination scheme. The formulation of an effective criterion is particularly complex in the EMO case, as judging progress can turn out to be as complex as the optimization itself. In other types of problems, such as function approximation, pattern recognition or single–objective optimization, on the other hand, the axis can be used as a "zero" reference for progress. That approach is unviable in EMO since in this case a solution consists of a set of points. Therefore, progress must be assessed in a relative manner rather than using the optimal solution set. While there has been little theoretical research dealing with EMO convergence, there have been even fewer attempts to deal with the stopping issue. In this talk we deal with the issue of EMO stopping criteria. The problem in question is presented, along with an in–depth description of the requirements imposed to stopping criteria for making them usable from a practical point of view. We also examine current approaches and their building blocks and discuss the connections of this issue with related matters, like performance indicators, convergence analysis and parameter optimization. We hope to prompt a debate on how to carry on the development of this issue.

# Robust optimization of nonlinear discrete-time processes with parametrically uncertain models

*Martin Mönnigmann (Ruhr-Universität Bochum)*

Various engineering problems involve finding a steady state of a dynamical system that is optimal with respect to a profit function. Since models of technical systems are often imprecise, any practically relevant optimization method must systematically cope with model uncertainty. This contribution presents an approach that is based on nonlinear programming and bifurcation theory. The proposed method is based on measuring the distance between a candidate optimal point and a critical boundary in the model parameter space. This distance to a critical boundary can be interpreted as a measure of parametric robustness of the candidate point. In a typical engineering application, the critical boundary of interest is the stability boundary of the dynamical system. In this case, the parametric distance between the candidate point and the critical boundary is a measure for the parametric robustness of stability.

The method has originally been developed for parametrically uncertain ordinary differential (or differential-algebraic) equations [1]. This contribution focuses on the extension to parametrically uncertain discrete time systems [2]. In this case, flip, fold, and Neimark-Sacker bifurcation points constitute the stability boundary. Simple extensions to parametrically robust feasibility boundaries and parametrically robust disturbance rejection are also discussed, and extensions to higher-order bifurcations and singularities are briefly mentioned.

The method is illustrated with several engineering examples.

[1] M. Mönnigmann, W. Marquardt, C. H. Bischof, T. Beelitz, B. Lang, P. Willems, A hybrid approach for efficient robust design of dynamic systems, SIAM Review 49 (2), 236-254, 2007.

[2] D. Kastsian, M. Mönnigmann, Robust Optimization of Fixed Points of Nonlinear Discrete Time Systems with Uncertain Parameters, SIAM J. Appl. Dyn. Sys., in print.

# Some Results Concerning the Run Rules, Synthetic and VSS Xbar charts with Estimated Parameters

*Philippe Castagliola (Université de Nantes)*

When monitoring a process using control charts, it is a common practice that a Phase I data set is used to estimate both the unknown in-control process mean 0 and in-control process standard-deviation 0. Once the process is considered to be in-control, these estimated control limits are assumed as fixed. This common practice totally ignores the effect of estimating 0 and 0 and the strong impact in terms of Run Length properties (cdf, ARL, SDRL, ...). In this presentation, we will focus on 3 different (Run Rules, Synthetic and VSS Xbar) charts and show how much the RL properties are different in the case where the parameters are supposed known and in the case where the parameters are supposed unknown and estimated. We will also show that the use of specific constants depending on the number m of Phase I samples allows to decrease this difference in practice. This talk is a follow up of the introductory talk presented last year during the 9th workshop on Quality Improvement in Bommerholz.

# Control charting under drift

*Sven Knoth (HSU Hamburg)*

Typically, control charts (surveillance or change point detection schemes) such as Page's CUSUM, Roberts' EWMA, and the Shiryaev-Roberts procedures are setup to detect a step change as quickly as possible while maintaining a low false alarm rate. In practical applications, a change may happen gradually. A linear drift (or trend) would be the first idea to model it.
Except Chang and Fricker Jr. (1999), most of the literature published about control charting under or for drift considers linear drift (even more, linear in the observation number). The early papers like Davis and Woodall (1988), Aerne et al. (1991) and Davis and Krehbiel (2002) consider the classical Shewhart chart with and without runs rules. Gan (1991 and 1992) provides numerical algorithms that allow accurate calculation of the Average Run Length for EWMA and CUSUM under drift. It is surprising that his methods are not frequently deployed. Most of the papers about control charting under drift rely on Monte Carlo studies and Markov chain approximations. Generally, one has to distinguish between procedures, which are especially designed for detecting drifts, and the usual step change detection schemes that do also a good job for identifying drifts. A recent paper with special drift detection procedures is Zou et al. (2009), where besides the usual EWMA and CUSUM and the more sophisticated GEWMA, special GLR methods are constructed. They provide both asymptotic and simulation results. Contrary to Chang and Fricker Jr. (1999) they indicate that the step change EWMA and CUSUM are really outperformed. In Fahmy and Elsayed (2006) a regression model on finite windows is applied for identifying potential drifts. Eventually, in Reynolds Jr. and Stoumbos, Z. G. (2001) the drift performance of control charts simultaneously monitoring mean and variance is evaluated. Again, all numerical results are fetched by Monte-Carlo studies. Additionally, Sweet (1988) discusses linear trend detection on the basis of EWMA schemes and the Winter-Holt model for forecasting linear trends. The most curious paper is Divoky and Taylor (1995) who studied 613 different trend rules in order to find an optimal trend detection scheme.
This work will give some new insights based on numerical results that are based on Gan (1991) and extensions. The performance measures under consideration are the zero- and steady-state Average Run Length. Both one- and two-sided setups are evaluated. The numerical treatment of the difference based schemes is comparable to the evaluation of control charts of dependent data and is done with Monte-Carlo simulations. Eventually, some recommendations for the application of control chart types are given.

# Kernel K-means clustering based Local Support Vector Domain Description control chart

*Issam Ben Khediri (Technische Universität Dortmund)*

As industrial technologies become more complex, modelling and detection of potential faults become more and more difficult. Indeed, usually modern industrial processes do not satisfy classical control chart assumptions, such as normality, independency or linearity. To overcome these limitations many non-parametric and non-linear charts have been recently proposed. One of these methods is the Support Vector Domain Description (SVDD) control chart. However, despite the success of applying SVDD method to fault detection of several processes, there was not enough research concerning applications of such methods to processes that run under multiple operating modes, because of product changes, set-point changes and manufacturing strategies. Unfortunately, direct application of the current methods to such processes tends to produce unsatisfactory performance due to the adopted assumption of only one nominal operating region for the underlying process. To address the above problem, this study develops a process monitoring scheme based on separate local models based on combination of SVDD method and kernel K-mean clustering in order to develop a new control chart able to monitor processes that are non-linear and that have different operating modes. In order to assess this monitoring strategy two different simulation studies as well as a real case study of an Etch Metal process are performed. Results show that the proposed control chart provides efficient fault detection performance with reduced false alarm rates.

# A comparison of sampling methods for approximating integrals in the optimal design of conjoint choice experiments

*Peter Goos (University of Antwerp)*

The semi-Bayesian approach for constructing efficient conjoint choice designs requires the evaluation of the design selection criterion value over numerous draws taken from the prior parameter distributions assumed in generating the design. The semi-Bayesian D-criterion value of a design is then calculated as the expected value of the D-error over all the draws taken. The traditional way to take draws from a distribution is to use the Pseudo-Monte Carlo approach. However, other sampling approaches are available as well. Examples are Quasi-Monte Carlo approaches using Halton sequences, Faure sequences, modified Latin hypercube sampling and extensible shifted lattice points, a Gauss-Hermite quadrature approach and a method using spherical-radial transformations. Not much is known in general about which sampling scheme is most efficient for calculating semi-Bayesian D-errors when constructing efficient conjoint choice designs. In this study, we compare the performance of these approaches under various scenarios and identify the most efficient sampling scheme for each situation.

# Orthogonal arrays in statistical software

*Ulrike Grömping (BHT Berlin)*

Traditionally, a few orthogonal arrays have been proposed and studied intensely, e.g. the L12, L18 or L36; additionally, the regular fractional factorial designs and screening designs by Plackett and Burman (1946) have been frequently used. However, researchers often have research needs that go beyond those simple designs. If a more general request is formulated, it is often accommodated using a non-orthogonal design obtained using a specified model and some optimality criterion (e.g. D-optimality), or it is forced into an existing orthogonal scheme by omitting some factors or factor levels.
Regular fractional factorial designs for pure 2-level factors have been intensively studied, and minimum aberration is an established criterion in selecting an appropriate design among them. About a decade ago, a generalization of this approach to non-regular 2-level designs (Tang and Deng 1999) and more general mixed level designs (Xu and Wu 2001) has been introduced. This generalized approach allows an overall assessment of the goodness of non-regular and mixed-level orthogonal arrays based on generalized word length patterns. Together with published catalogues of useful orthogonal arrays, these results can be used in software for implementing an automatic creation of experimental designs according to user requests based on orthogonal arrays. However, the state of knowledge in this area is far less advanced than for the regular designs, and software implementations which allow creation of orthogonal arrays that satisfy specific estimation needs are scarce. SAS software has a lone standing as THE software that can generate the most flexible orthogonal arrays, based on the %mktex macro provided by Warren Kuhfeld (cf. Kuhfeld and Tobias 2005, Kuhfeld 2009).
The talk discusses the theoretical concepts and presents an attempt at implementing general (mixed-level) orthogonal arrays into R software as part of the R-package DoE.base - with its function oa.design. The package makes use of a large data base of orthogonal arrays assembled by Warren Kuhfeld (2009) and can - within a particular design - automatically assign columns such that a design's properties are optimized in terms of generalized word length pattern. In principle, the procedure is able to find orthogonal arrays of generalized resolution IV or even V. However, resources are a severe limitation, and the space of non-isomorphic mixed-level candidate designs is so large that the Kuhfeld catalogue only covers a very small portion. The talk is a presentation of work in progress. The workshop participants' expertise in Design of Experiments and statistics for quality improvement will hopefully generate useful feedback to be accounted for in the R-software implementation.

## Numerical Sound Synthesis

*Stefan Bilbao (University of Edinburgh)*

Sound synthesis, based on physical modelling principles, has developed enormously recently, thanks in part to great increases in computational power on readily available computer hardware. At the heart of physical modelling is, as the name suggests, a strict adherence to a physical description of a musical object—such an object could be a musical instrument, either of conventional type, or a wholly imaginary construction, various electronic and electro-acoustic audio effects, and full models of acoustic spaces. The great benefits of a physical approach are a) very high-quality sound rendering which has an innately acoustic quality, b) great flexibility (i.e., no reliance on a database of recorded sound fragments, and c) simplified control for the eventual user.
Most techniques which have been used for these purposes have developed directly from earlier antecedents in sound synthesis, and are descendants of signal processing techniques involving digital filter designs, and banks of oscillators. Though sometimes very efficient, such methods are isolated from mainstream numerical methods used in other simulation applications, which, though more computationally demanding, are better suited to the complexities of sound rendering. In this talk, an overview of physical modelling techniques will be presented, followed by an a variety of applications, including Matlab-generated sound and video, and a discussion of numerical stability issues.
Stefan Bilbao is currently a Senior Lecturer in the Acoustics and Fluid Dynamics group/Music subject area at the University of Edinburgh. He was previously a lecturer at the Queen's University Belfast, and a post-doctoral research associate at the Stanford Space Telecommunications and Radioscience Laboratory, after completing his PhD in Electrical Engineering at Stanford University, in 2001.

## Process Control in practice at TNT Post

*Thijs Vermaat (TNT Post, Netherlands)*

Process control is a core management task at Soring Centres at TNT Post, such as the Sorting Centre in the Rotterdam Area, where I am deputy manager. Process control can be done both beforehand (such as feed forward control) as well as afterwards (SPC and feedback control). For both instruments different theories and a number of tools are available, such as the control chart. I want to share a number of problems/challenges with you which arise applying these theorems in practice:
- violation of the iid-assumption;
- application in a system in which people act continuously (a non-stable system);
- monitoring summations of processes (where the summations are of non iid-variables).

Theories are either too complex to understand and implement or just do not exists. I will explain the solutions we came up with at TNT Post and the impact the applied theories have on the day-to-day business.

## Claim Development and Verification

*Anja Schleppe (Kraft Foods, Germany)*

This important topic is shared with the audience in the format of an eLearning module. The presenter explains the background when a Kraft Foods employee would be asked to run this eLearning module to get a basic understanding of Claim Development and Verification. The audience will then be put into the shoes of the Kraft Foods employee: One audience member gets the mouse and the audience as a team walks through the pages and tries to answer the interactive quizzes. The wrap-up session will be moderated by the presenter.