

# 18th Workshop on Quality Improvement Methods

July 16th and 17th, 2021

## Applied statistics in the food industry

Fr 15:10

Anja Schleppe

*Mondelēz International*

Mondelēz International is a company on an IL6S journey, with some aspects of DFSS being implemented at the same time. Starting to move from detection and correction to prevention is a major step for a large organization to take. This talk shares experiences from this journey and the key importance of data based decisions.

---

## Growth curve model with block random effects

Fr 15:55

### - sufficiency and estimation

Katarzyna Filipiak

*Poznań University of Technology*

We consider an extended growth curve model with random block effects. We show that under the mixed effects model the sufficient statistics for a fixed effects extended growth curve model remain the same. We also determine the maximum likelihood estimators of unknown parameters and we will look for the conditions, under which the likelihood equations simplify. We then apply the results for a real data example.

## Design selection for 2-level supersaturated designs

Fr 17:00

Rakhi Singh

*University of North Carolina at Greensboro*

The commonly used design optimality criteria are inadequate for selecting supersaturated designs. As a result, there is an extensive literature on alternative optimality criteria within this context. Most of these criteria are rather ad hoc and are not directly related to the primary goal of experiments that use supersaturated designs, which is factor screening. Especially, unlike almost any other optimal design problem, the criteria are not directly related to the method of analysis.

An assumption needed for the analysis of supersaturated designs is the assumption of effect sparsity. Under this assumption, a popular method of analysis for 2-level supersaturated designs is the Gauss-Dantzig Selector (GDS), which shrinks many estimates to 0. We develop new design selection criteria inspired by the GDS and establish that designs that are better under these criteria tend to perform better as screening designs than designs obtained using existing criteria.

This presentation is based on joint work with John Stufken, University of North Carolina at Greensboro.

---

## Factor selection in screening experiments

Fr 17:45

John Stufken

*University of North Carolina at Greensboro*

Screening designs are used in design of experiments when, with limited resources, important factors are to be identified from a large pool of factors. Typically, a screening experiment will be followed by a second experiment to study the effect of the identified factors in more detail. As a result, the screening experiment should ideally screen out a large number of factors to make the follow-up experiment manageable, without screening out important factors. The Gauss-Dantzig Selector (GDS) is often the preferred analysis method for screening designs. While there is ample empirical evidence that fitting a main-effects model can lead to incorrect conclusions about the factors if there are interactions, including two-factor interactions in the model increases

the number of model terms dramatically and challenges the GDS analysis. We discuss a new analysis method, called Gauss Dantzig Selector Aggregation over Random Models (GDS-ARM), which aggregates the effects from different iterations of the GDS analysis using different randomly selected interactions columns each time. This is joint work with Rakhi Singh, University of North Carolina at Greensboro.

---

**Reducing the number of doses in dose-finding trials**  
-  
**impact on power and estimation**

Sat 11:00

Christoph Neumann  
*Bayer AG, Wuppertal*

In clinical trials in the phase 2b the usage of MCP-Mod is quite common. The aim of these studies is to show a dose-response effect and the model the dose-response relationship thereafter. Usually these studies are powered to show the dose-response in the MCP step. When it comes to the selection of concrete doses often the idea is to save a dose for costs, ethical reasons or feasibility of the study. For this, we analyzed the influence of reducing the number of patients at one dose to a minimum or even leaving one dose out of the study on the power, the estimation of a minimum effective dose and the estimation of a response on a specific dose with comprehensive simulations. We found that in general it is not recommended to reduce the number of doses but that there is no big difference between leaving a dose completely or just reducing the number of patients. Especially we show that a powered MCP-Mod study does not guarantee good estimation of the dose-response relationship.

# An algorithm for blocking regular fractional 2-level designs with clear two-factor interactions

Sat 11:45

Ulrike Grömping

*Beuth University of Applied Sciences Berlin*

Regular fractional factorial designs with 2-level factors are among the most frequently used experimental plans. In many cases, designs should be blocked for dealing with inhomogeneity of experimental units. At the same time, the research question at hand may imply a focus on specified sets of two-factor interactions, while it is not justified to assume negligibility of other low order effects.

The talk presents an algorithm for blocking a regular fraction into – possibly small – blocks while keeping specified two-factor interactions clear from confounding with main effects or other two-factor interactions (Grömping 2021). The proposed algorithm is implemented in the R package FrF2 (Grömping 2007-2020) and combines a graph-based estimability algorithm by the author (Grömping 2012) with an automated implementation of a recent proposal by Janet Godolphin for blocking fractions by hand (Godolphin 2021).

## References

- Godolphin, J. (2021) Construction of blocked factorial designs to estimate main effects and selected two-factor interactions. *Journal of the Royal Statistical Society B* 83, 5–29. <https://doi.org/10.1111/rssb.12397>.
- Grömping, U. (2007-2020) *The FrF2 Package (Fractional Factorial designs with 2-level factors)*. R package version 2.2-2.
- Grömping, U. (2012) Creating clear designs: a graph-based algorithm and a catalog of clear compromise plans. *IIE Transactions* 44, 988–1001. <https://doi.org/10.1080/0740817X.2012.654848>.
- Grömping, U. (2021) An algorithm for blocking regular fractional factorial 2-level designs with clear two-factor interactions. *Computational Statistics and Data Analysis* 153, 1–18. <https://doi.org/10.1016/j.csda.2020.107059>.

## Predictive asset maintenance at scale

Sat 13:30

### the challenge of large numbers

Winfried Theis

*Shell*

In Shell Predictive Asset Maintenance was from the start one of the major challenges to be tackled by Digitalisation and Data Science. The opportunity in this area is clear, when you realise that any unplanned outage of equipment can quickly lead to deferrals of several thousand if not millions of dollars worth. The challenges range from dealing with the sheer amounts of data needed for training of the models, to the number of equipment in need of monitoring – in the first project on valves there were immediately about 600 in scope – and finally central equipment being tripped by problems in connected systems. Here some of the answers to these challenges are presented at a high level, mainly to increase the awareness that also our data science processes share the same challenges in scaling up as production processes generally do.

---

## Mixed models in sensometrics, and many other places

Sat 14:15

Per Bruun Brockhoff

*DTU Compute, Technical University, Denmark*

I have worked with statistical methods, applications and developments in general and for sensory and consumer data since 1991 and until 2017 where I started a pause in my own research and innovation endeavors to take a management role. A recurrent topic has been the design of experiments and analysis of experimental and complex structured data, with more focus on the analysis part than on the design part. The design part other more skilled researchers have contributed much more and better to, like Prof. Dr. Joachim Kunert. A recurrent analysis method has been *mixed models* in one version or the other. I have been teaching applied linear mixed models in The Royal Veterinary and Agricultural University (now Copenhagen University) and since 2004 at The Technical University of Denmark, and created online teaching material including a complete textbook and lecture videos (<https://02429.compute.dtu.dk>).

In sensometrics, mixed models are often relevant to consider due to the very basic issue of replicated and/or hierarchical data of any kind. The mixed models come in as well linear normal versions as non-normal, non-linear versions. At the DTU Compute Sensometrics Group we developed new and extended mixed model methodology and new R-packages for both linear normal and certain non-normal/non-linear situations highly relevant in sensometrics. Some of them are generic statistical R-packages (*lmerTest*, *ordinal* and *mumm*) and two of them are dedicated sensometrics R-packages: *sensR* and *SensMixed*.

I will during this talk touch on some of what these packages can do for us and what impact they have had and still have. With more than 1000 daily CRAN downloads, and a total Google Scholar citation number above 10.000 the *lmerTest* package and the corresponding 2017 Journal of Statistical Software paper has had quite some impact, and still growing.

Given the type of arrangement, I plan to also allow for a little time to include an anecdote or two with some links back to the scientific history of the Copenhagen area including the Bohr-family.

#### **Some software and book references:**

1. Poedenphant, S. and Brockhoff, P.B. (2018). *mumm*: Multiplicative Mixed Models using Template Model Builder. R package version 0.2.0.  
<https://CRAN.R-project.org/package=mumm>
2. Christensen, R.H.B. and Brockhoff, P.B. (2015). *sensR*: An R-package for sensory discrimination. R package version 1.4-5. <https://cran.r-project.org/web/packages/sensR/>
3. Kuznetsova, A., Brockhoff, P.B. and Christensen R.H.B. (2017). *lmerTest* Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13), 1–26.  
(<https://doi.org/10.18637/jss.v082.i13>)
4. Kuznetsova, A., Brockhoff, P.B. and Christensen, R.H.B. (2014). *SensMixed*: Mixed effects modelling for sensory and consumer data. R package version 2.0-8.  
<http://cran.at.r-project.org/web/packages/SensMixed/>
5. *PanelCheck*, Open source software for sensory profile data, [www.panelcheck.com](http://www.panelcheck.com)
6. *ConsumerCheck*, Open source software for sensory and consumer data,  
<http://consumercheck.co/>
7. Næs, T., Brockhoff, P.B. and Tomic, O. (2010). *Statistics for Sensory and Consumer Science*, John Wiley & Sons.
8. Brockhoff, P.B. (2015) *Analysis of correlated Data: Mixed Linear Models* - eBook, 400pp. Available from: <https://02429.compute.dtu.dk/eBook>

# Farewell celebration

Sat 16:00

## Statistics, geometry and pharmacokinetics

Holger Dette

*Ruhr-Universität Bochum*

---

## Multiblock data analysis applied to the sensometrics field

El Mostafa Qannari

*Oniris Nantes*

Some thirty years ago (how time flies!), I was lucky enough to make the acquaintance of Joachim Kunert within a European program devoted to sensory analysis. Sensometrics was at a fledgling stage. An example of questions that puzzled the practitioners faced with data that are structured into around twelve to twenty blocks of variables (sensory profiling data) was whether they should use Generalized Procrustes Analysis or STATIS method. Joachim's (and Michael Meyner's) answer to the controversy that opposed advocates of both methods (mainly French against English believers) was to set up a common model to both methods and perform a simulation study. Not a common approach at that time! And the winner is... Well, certainly Sensometrics because the lesson we learnt was to put the arguments on the rails of rigorous thinking. Sensory analysis has very much evolved and so has sensometrics. In the meantime, the interest for multiblock data analysis has grown in numerous domains of science and application. We discuss some achievements in this topic within the field of sensory analysis.

---

## Mishaps, brainstorm and strokes of genius

—  
some recollection

Michael Meyners

*Procter & Gamble Service GmbH*