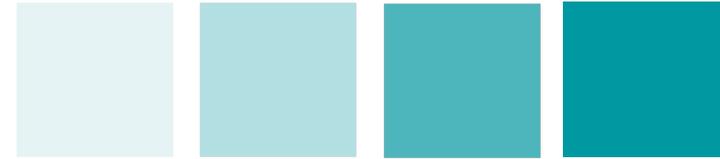


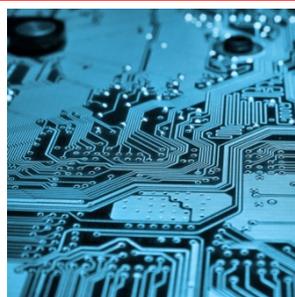
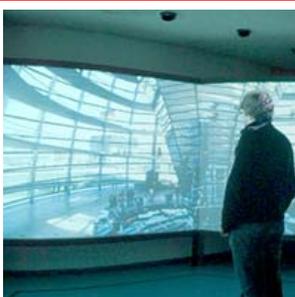


BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences



Design of Experiments in R

*Prof. Ulrike Grömping
BHT Berlin*





Design of Experiments (DoE) in R

- High-level goals
- Structure / Output objects
- Scope
- Some usability aspects
- Call for contributions





Mission: Support application of (Industrial) DoE in R

Target users:

- Inexperienced / insecure users, who need to be presented with a minimal set of preselected choices
- Expert users, who need state-of-the-art methods and the flexibility for making the most of their expertise

Make R competitive in the market for DoE software

- Provide full base functionality for DoE in R
- Implement some advanced methods
- Well-structured GUI (comfort cannot be fully competitive)

Hope: Laying the foundation

→ others will also implement advanced DoE functionality in R



Package **DoE.base** (early version on CRAN)

for full factorials, orthogonal arrays,
and base utilities for the other packages → avoid bundling



Package **FrF2** („medium“ version on CRAN)

for (regular and non-regular) 2-level fractional factorials



Package **DoE.wrapper** (not yet on CRAN)

- wrapper for existing functionality to unify syntax and output structure (class `design`) and add comfort where necessary
- **Challenge:** help users choose a design cross package



GUI interface as an **R-commander plugin**:

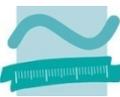
Package **RcmdrPlugin.DoE** (not yet on CRAN)

BsMD
igraph
scatterplot3d
sfsmisc

lhs
AlgDesign
rsm

Rcmdr





The same output structure for all types of design:
object of S3 class `design`

- is a data frame with attributes
 - has been inspired DoE-functions from the White Book (Statistical models in S)

- the data frame itself:
the design as factors or uncoded data

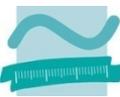
- the attributes
 - `desnum`: numeric or coded version of the design
 - `run.order`: data frame with run order information for looking at standard order and returning to randomized order
 - `design.info`: list with design type-dependent information



- Full factorials, orthogonal arrays for main effects designs (package **DoE.base**)
- Orthogonal plans for 2-level factors (package **FrF2**)
 - Regular fractional factorial designs (function **FrF2**)
 - based on catalogues of non-isomorphic designs
 - blocking, split-plot, hard-to-change factor levels
 - estimable 2-factor interactions
 - not yet: augmentation by foldover or star points intended
 - not yet: designs with 2- and 4-level factors
 - Non-regular designs (function **pb**)
 - Plackett-Burman, some exceptions (16, 32, 64 runs), where better for screening
 - not yet: blocking
- Latin hypercube samples, response surface designs for quantitative variables (package **DoE.wrapper**)
- D-optimal plans, perhaps mixture designs (package **DoE.wrapper**)



- Full factorials, **orthogonal arrays** for main effects designs (package **DoE.base**)
 - a few standard special arrays (like Taguchi's L18(2^13^7))
 - All of Warren Kuhfeld's parent arrays are available (http://support.sas.com/techsup/technote/ts723_Designs.txt),
 - soon: child arrays (by expansive replacement method)
 - vision:
 - more intricate SAS-like ways of combining these, **a lot of effort!**
- Orthogonal plans for 2-level factors (package **FrF2**)
- Latin hypercube samples, response surface designs for quantitative variables (package **DoE.wrapper**)
- D-optimal plans, perhaps mixture designs (package **DoE.wrapper**)
- Not: various special types of design available in  as described in the CRAN Task View „ExperimentalDesign“



Near Future

- Make existing analysis capabilities accessible through **RcmdrPlugin.DoE** package:
 - linear model functions in general (are in R-commander already)
 - simple plotting facilities for orthogonal 2-level experiments from package **FrF2**
 - analysis facilities for response surface designs from package **rsm**

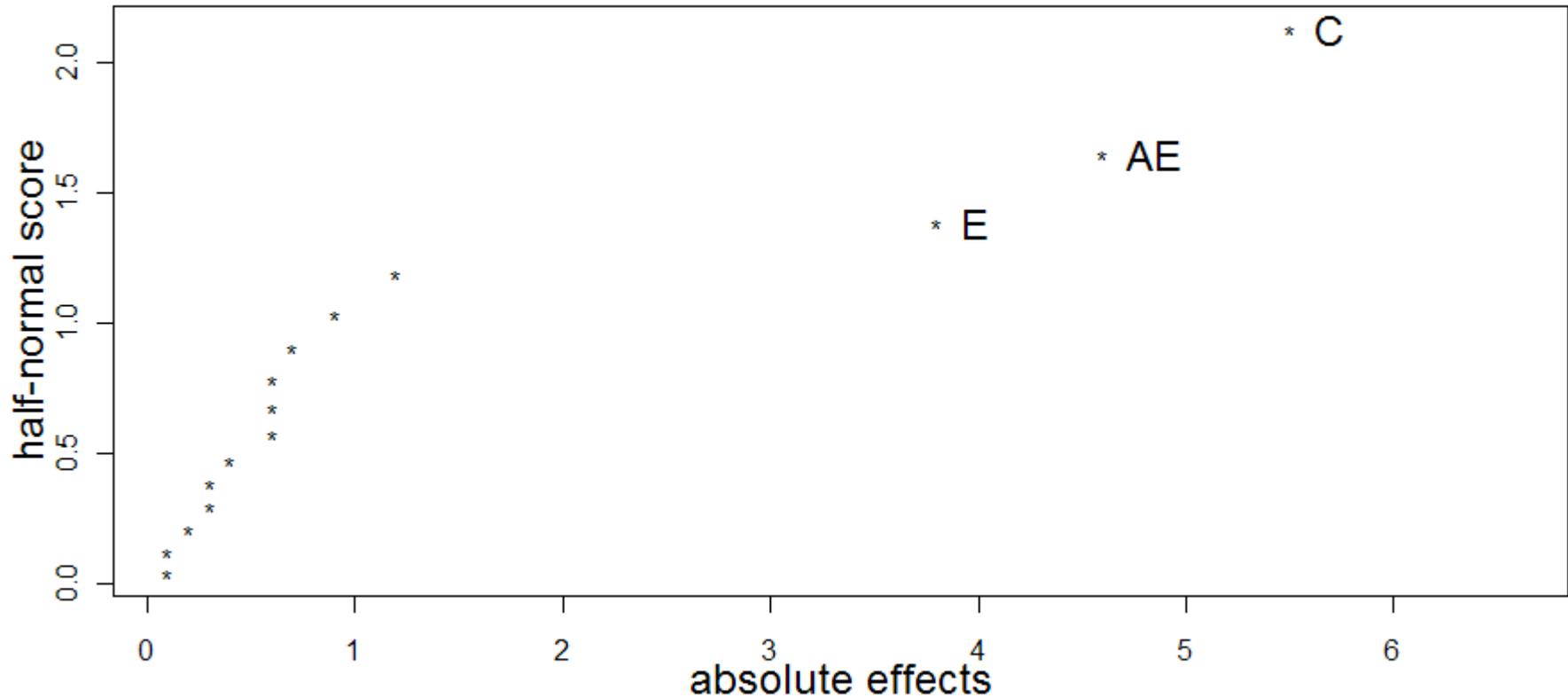
Later

- Special analysis functions (command line use) that make use of the info in class **design** objects for providing reasonable default analyses





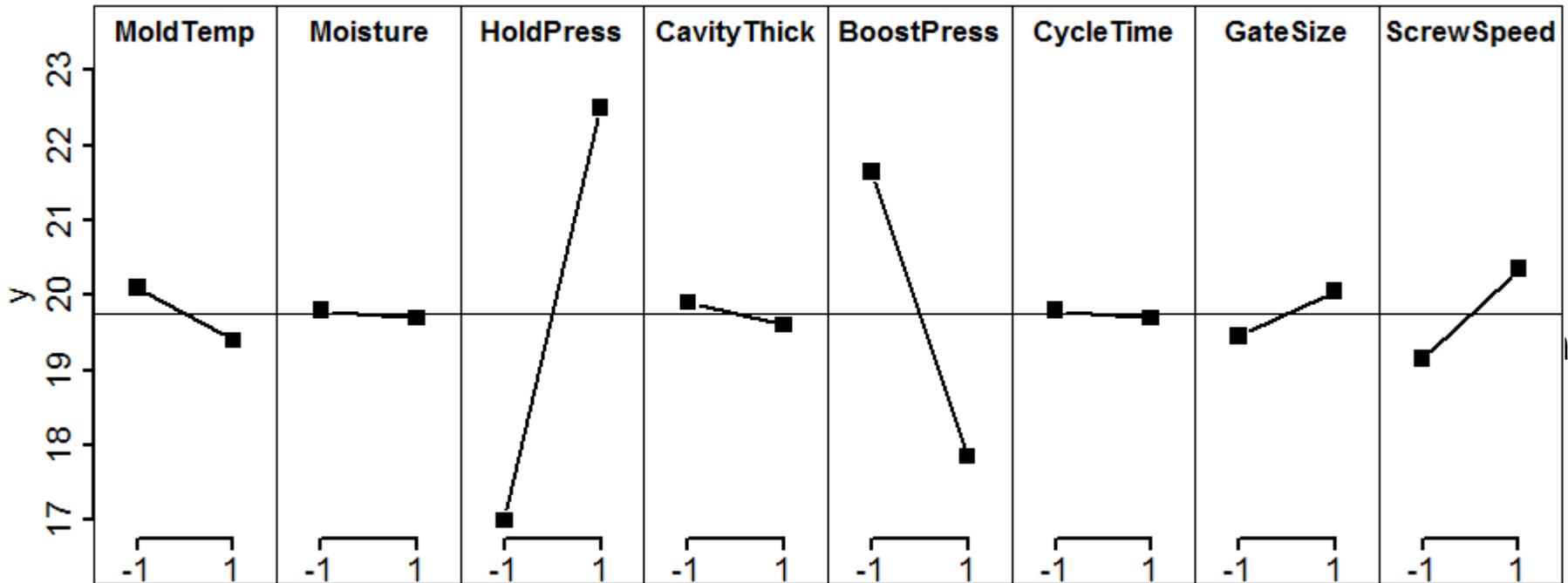
Half Normal Plot for y, alpha=0.05



A = MoldTemp , B = Moisture , C = HoldPress , D = CavityThick , E = BoostPress , F = CycleTime , G = GateSize , H = ScrewSpe



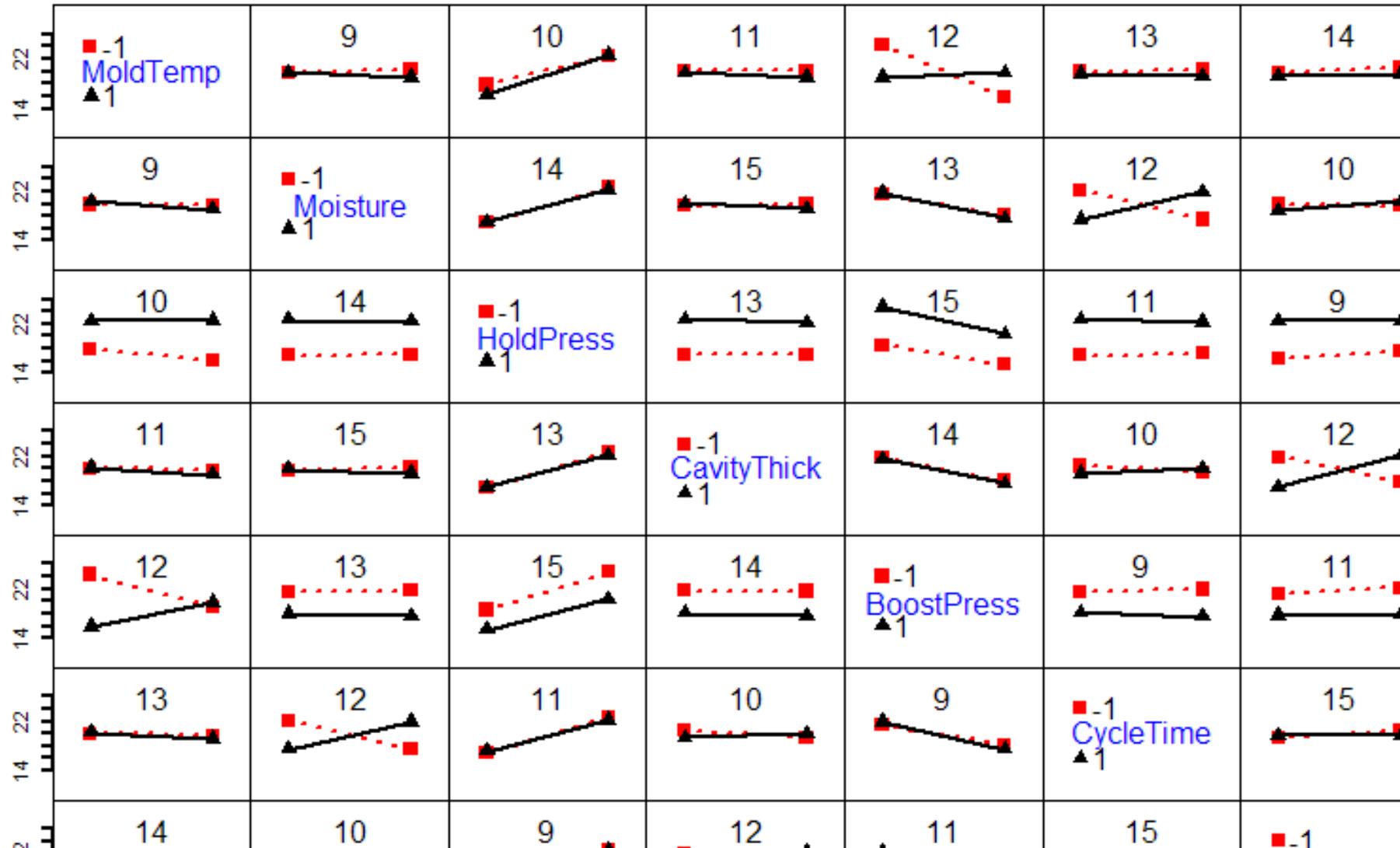
Main effects plot for y

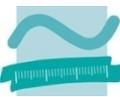


Scope: Analysis



Interaction plot matrix for y





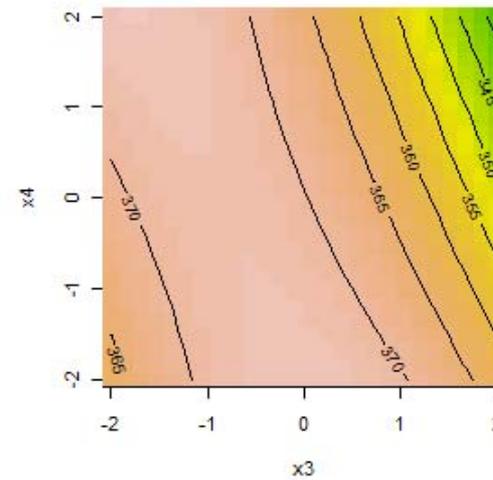
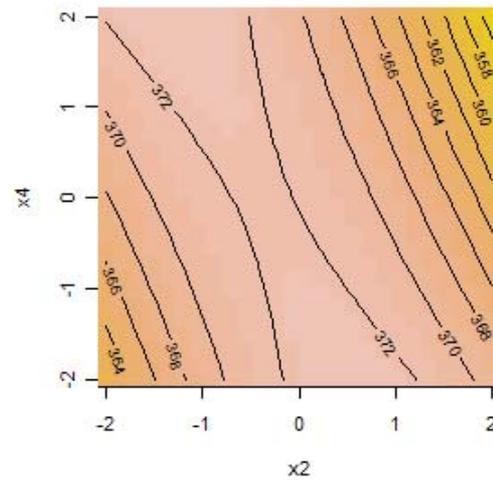
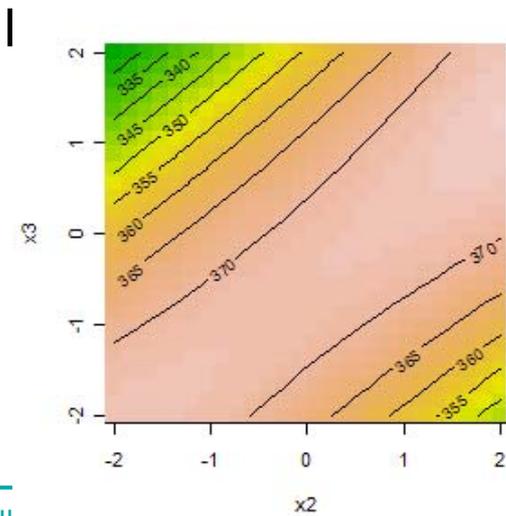
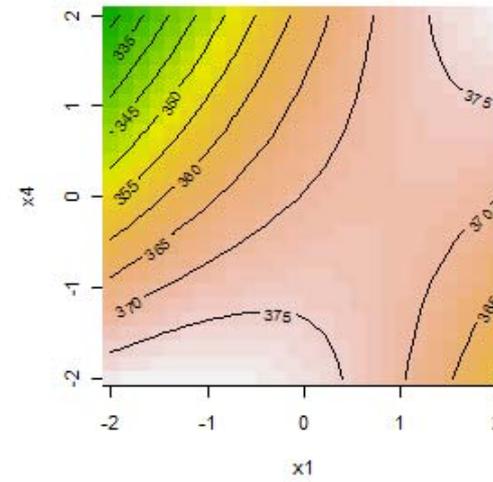
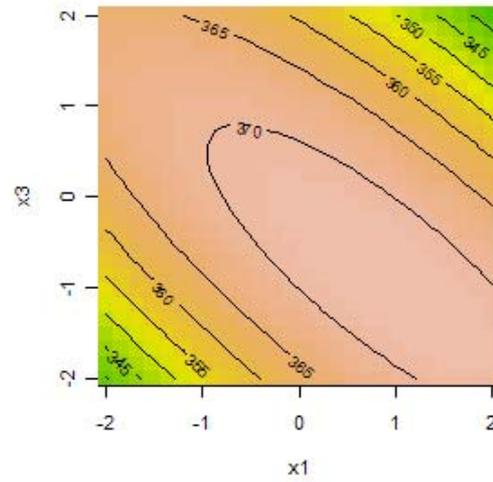
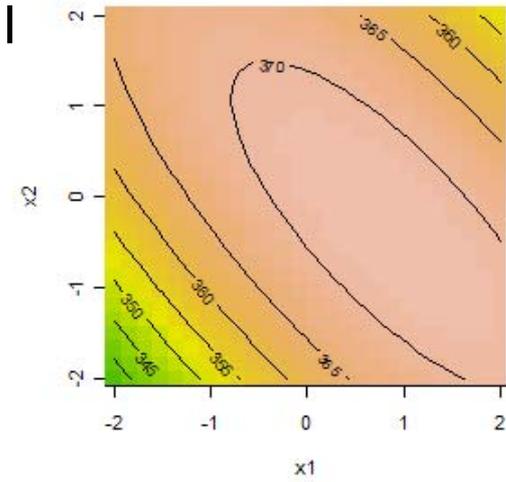
Near Future

- Make existing analysis capabilities accessible through **RcmdrPlugin.DoE** package:
 - linear model functions in general (are in R-commander already)
 - simple plotting facilities for orthogonal 2-level experiments from package **FrF2**
 - analysis facilities for response surface designs from package **rsm**

Later

- Special analysis functions (command line use) that make use of the info in class **design** objects for providing reasonable default analyses





ady)

je rsm





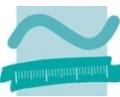
Near Future

- Make existing analysis capabilities accessible through **RcmdrPlugin.DoE** package:
 - linear model functions in general (are in R-commander already)
 - simple plotting facilities for orthogonal 2-level experiments from package **FrF2**
 - analysis facilities for response surface designs from package **rsm**

Later

- Special analysis functions (command line use) that make use of the info in class **design** objects for providing reasonable default analyses





Usability is very important for the intended user group!

- Work directly with standard R installation
 - direct exporting to xls not possible, produce formatted Excel sheet via html

GUI aspects

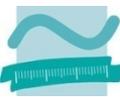
- Support both experts and DAUs
 - Simple interface that can be extended to an expert level
- Good help facilities, both on content and interface
- Store inputs, so that
 - interruption of tedious input work is safe
 - modifications of inputs are comfortably possible during the planning phase of an experiment



- The project is **progressing** well
 - Roughly on time, useful result available by 30/09/2009
 - Still quite a way to go after September 2009

- **Potential contributions** (more ideas welcome):
 - Bug reports, suggestions for improvement, wishes, Contributions of orthogonal arrays for **DoE.base**
 - Beta-testing for **RcmdrPlugin.DoE** (not quite yet)
 - Support on internationalization (not quite yet)
 - implementation of special functionality into **DoE.wrapper** or **RcmdrPlugin.DoE**
 - separate packages that fit into the project input and output structure
 - SAS macro-like functionality (MktEx) for intricate (market research) designs based on orthogonal arrays

- **Bob Wheeler is looking for an „heir“ for AlgDesign (optimal DoE)**



■ R-packages

- **AlgDesign**: Algorithmic experimental designs. *Bob Wheeler*
- **BsMD**: Bayes Screening and Model Discrimination. *Ernesto Barrios*
- **igraph**: Routines for simple graphs, network analysis. *Gabor Csardi*
- **lhs**: Latin Hypercube Samples. *Rob Carnell*
- **scatterplot3d**: 3D Scatter Plot. *Uwe Ligges*
- **sfsmisc**: Utilities from Seminar für Statistik ETH Zürich.
Martin Mächler and many others
- **Rcmdr**: R commander. *John Fox* with contributions from many others
- **rsm**: Response-surface analysis. *Russ Lenth*

■ Other sources

- Chambers, J.M. and Hastie, T.J. (1991, eds.). *Statistical models in S*. Chapman and Hall, London. (*The White Book*)
- Grömping, U. (2008-2009). *CRAN Task View on Design of Experiments*.
<http://<your CRAN mirror>/web/views/ExperimentalDesign.html>.
- Warren Kuhfeld (2009). Orthogonal arrays.
<http://support.sas.com/techsup/technote/ts723.html>.

