Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

PLISKA
STUDIA MATHEMATICA BULGARICA
ПЛИСКА
БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

The attached copy is furnished for non-commercial research and education use only.
Authors are permitted to post this version of the article to their personal websites or
institutional repositories and to share with other researchers in the form of electronic reprints.
Other uses, including reproduction and distribution, or selling or
licensing copies, or posting to third party websites are prohibited.

For further information on
Pliska Studia Mathematica Bulgarica
visit the website of the journal http://www.math.bas.bg/~pliska/
or contact: Editorial Office
Pliska Studia Mathematica Bulgarica
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: pliska@math.bas.bg

Pliska Stud. Math. Bulgar. **17** (2005), 229–239

PLISKA
STUDIA MATHEMATICA
BULGARICA

APPLICATION OF STATISTICAL EXPERIMENTAL DESIGN IN FOOD SCIENCES

Ute Römisch

The development of new, health supporting food of high quality and the optimization of food technological processes today require the application of statistical methods of experimental design. The principles and steps of statistical planning and evaluation of experiments will be explained. By example of the development of a gluten-free rusk (zwieback), which is enriched by roughage compounds the application of a simplex-centroid mixture design will be shown. The results will be illustrated by different graphics.

1. Introduction

The production of health supporting food of high quality on modern and effective technological and economical conditions, which have to consider also demands of ecology and food security, is the main task of food industry.

Foodstuffs and their row materials are biological products of high natural variability and build complex multicomponent systems with sensory properties. Their composition and structure depend on the respective current state of happening biochemical, chemical, microbiological and physicochemical processes. Foodstuffs have only limited durability.

The cooperation of statisticians and food engineers often is characterized by following problems:

- The not sufficient technological and natural scientific knowledge of statisticians, including experiences, can result in not correct suppositions about the process and its conditions and a following not adapted planning of experiments.

- Very often statisticians get not enough information about the whole problem, they are integrated only into the solution of partial tasks and are not able then to plan and evaluate the experiments in complex.
- During realization of experiments often the conditions of the process are changed, the experimenters already empirically look to better conditions and sometimes the used experimental material comes from different groups.
- Often repetitions of experiments are planned, which are no really independent repetitions, but results of double or multiple measured experiments. On the other hand sometimes the experimental results later prove to be means of different series of experiments. The statistician must try to solve this problems at the beginning of his cooperation.

2. Methods of statistical experimental design

The quality and the informational content of experimental results depends on the experimental design. Its aim is to get results with sufficient precision and reliability on the base of a minimal number of aim oriented experiments respectively to get as much convincing results as possible on the base of a limited number of experiments.

The following well known principles of experimental design ([1],[2],[4],[8]) should be considered:

- Repetition and multiple realisations of experimental points
- Randomization
- Block building
- Symmetry
- Confounding
- Sequential planning of experiments

On principle the methods of statistical experimental design contain the following steps, including feedbacks between them:

1. Precise formulation of the complete problem and the main aim Division into partial tasks and thinking about (conditional?) relations between the partial tasks
2. Formulation of statistical models Derivation of experimental designs Calculation of sample sizes
3. Carrying-out of experiments
4. Statistical evaluation of experiments Interpretation (technological, chemical,. . .) of the results
5. Conclusions

In the Faculty of Process Sciences and Engineering at TU Berlin the following experimental designs (ED) are applied:

- Plackett-Burman experimental designs
- Complete and fractional experimental designs of first order of type 2^k und 2^{k-p}
- Box-Behnken experimental designs
- Complete and fractional experimental designs of second order of type 3^k und 3^{k-p}
- Combined experimental designs of type 2^k und $3^{k'}$ resp. 2^{k-p} and $3^{k'-p'}$
- Central composite experimental designs of type 5^{k-p} (rotatable and /or orthogonal ED)

Therefore statistical programs as STATISTICA [7], STAVEX and STATGRAPHICS are used. The practical way of handling these methods will be demonstrated on an example of Bakery industry.

3. Investigation of roughages in gluten-free baker's goods

For persons who were fallen ill of "Zöliakie" it is not allowed to eat bakery products which contain gluten. But the supply of gluten-free products is not sufficiently. In our case a gluten-free rusk (zwieback), which is enriched by roughages should be developed.
Three partial tasks were to be solved:

3.1. Development of a rusk, which is free of Gluten

Linear model with interactions, ED of type 2^{5-1}, optimal levels of explanatory variables

3.2. Investigation of different roughage compounds concerning a possible use in rusk

Comparison of five roughage compounds

As a results of the first two tasks, which will not be explained here in detail, a dough without Gluten, but Gluten simulating sensory properties was prepared and used for the further experiments with roughages. Four roughage compounds were chosen as suitable for a use in rusk: BS1 (Sugar beet), BS2 (Pea), BS3 (Soya-bran) and BS4 (Soya-bean) [5].

3.3. Finding out an optimal combination of roughage compounds

The solution of this third partial task will be discussed in greater detail now.

The aim was to find out that optimal combination of the roughage compounds, at which the volume yield of rusk becomes a maximum on condition that sensory parameters will fulfil quality standards and the total share of roughage compounds will not change.

A quadratic model with four explanatory variables at five levels, which could be extended to an incomplete cubic model was taken as a basis for a simplex-centroid design (Table 1, Figure 1) of the mixture experiments ([3],[6],[9]). As most important response variable the volume yield of the rusk [ml/100g] was considered. The results are means of two series. The grey marked experiments have shown sufficient values for the volume yield as well as for quality parameters of the dough and the rusk. The first experiment (No. 0) without roughage compounds was used for comparisons. The following models [3] were assumed:

Quadratic model:

$$\eta = \text{ß}_0 + \sum_{1 \leq i \leq k} \beta_i x_i + \sum_{1 \leq i \leq j \leq k} \beta_{ij} x_i x_j \text{ with } \sum_{i=1}^{k} x_i = 1; \text{ (here: k=4)}$$

$$\rightarrow \eta = \sum_{1 \leq i \leq k} \beta_i * x_i + \sum_{1 \leq i < j \leq k} \beta_{ij} * x_i x_j$$

Cubic model:

$$\eta = \text{ß}_0 + \sum_{1\leq i\leq k} \beta_i x_i + \sum_{1\leq i\leq j\leq k} \beta_{ij} x_i x_j + \sum_{1\leq i\leq j\leq h\leq k} \beta_{ijh} x_i x_j x_h \text{ with } \sum_{i=1}^{k} x_i{=}1$$

$$\rightarrow \eta = \sum_{1\leq i\leq k} \beta_i * x_i + \sum_{1\leq i<j\leq k} \beta_{ij} * x_i x_j + \sum_{1\leq i<j\leq k} \gamma_{ij} * x_i x_j (x_i - x_j) + \sum_{1\leq i<j<h\leq k} \beta_{ijh} * x_i x_j x_h$$

Special (incomplete) cubic model:

$$\eta = \sum_{1\leq i\leq k} \beta_i * x_i + \sum_{1\leq i<j\leq k} \beta_{ij} * x_i x_j + \sum_{1\leq i<j<h\leq k} \beta_{ijh} * x_i x_j x_h \ .$$

Exp. No. i	BS1 X1	BS2 X2	BS3 X3	BS4 X4	Volume yield Y1
0	0	0	0	0	450,00
1	1	0	0	0	367,50
2	0	1/2	0	1/2	423,00
3	1/2	0	1/2	0	372,00
4	0	1	0	0	423,00
5	1/3	1/3	1/3	0	408,00
6	0	0	1/2	1/2	416,50
7	1/2	1/2	0	0	360,50
8	0	1/3	1/3	1/3	386,00
9	0	0	1	0	374,00
10	0	1/2	1/2	0	398,50
11	1/2	0	0	1/2	386,00
12	1/3	0	1/3	1/3	375,50
13	1/4	1/4	1/4	1/4	363,00
14	1/3	1/3	0	1/3	358,00
15	0	0	0	1	364,00

Table 1: Simplex-centroid design (mean series)

The regression and variance analytical evaluation of the results of the mixture experiments for the volume yield can be found in table 2. Models of increasing complexity will be compared concerning their goodness of fit to the observed data.

The special cubic model, which contains as well the two-component as the three-component mixtures has shown the best model fit. Its $R^2_{adj.}$(88,15) is higher then in the case of linear and quadratic models, but the simultaneous tests for

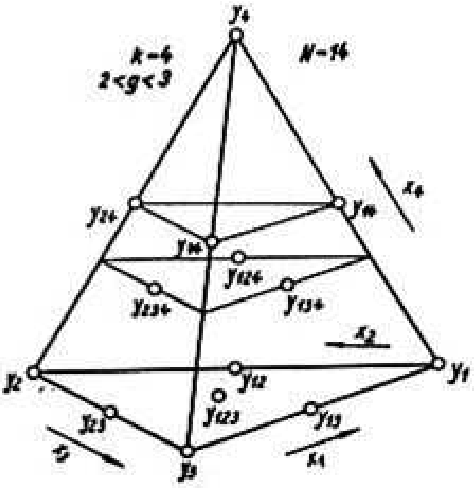

Figure 1:

all parameters of these three models were not statistically significant ($p > 0,5$). The fit of the surface only could be improved by reducing the model by the not significant two-component mixtures X_1X_3 and X_2X_3 . All parameters of this reduced model were statistically significant ($p < 0,05$) and $R^2_{adj.} = 95,86$ (Table 2).

The estimated special cubic model has the form:

Y1 = 368,078 X1 + 423,078 X2 + 374,75 X3 + 363,905 X4 - 137,274 X1X2 + 83,071 X1X4 + 121,071 X2X4 + 191,726 X3X4 + 865,494 X1X2X3 - 999,325 X1X2X4 - 715,666 X1X3X4 - 1041,17 X2X3X4

Source	SS(effect)	DF	MS	F-Ratio	P-value	St. err.	R^2	$R^2_{adj.}$
Linear model	2573,56	3	857,85	1,90	0,1877	21,2361	34,16	16,20
Quadratic model	1704,37	6	284,06	0,44	0,8293	25,5198	56,78	0,00
Special cubic model	3192,53	4	798,13	12,51	0,2057	7,98621	99,15	**88,15**
Error	63,78	1	63,78					
Total	2,23129E6	15						
Special cubic model (without X1X3, X2X3)	7467,36	11	678,85	30,45	**0,0084**		9,11	**5,86**
Error	66,8734	3						
Total (corr.)	7534,23	14						

Table 2: Sequential fit of models of increasing complexity

A comparison of the observed and predicted values of the model (Figure 2) shows a good fit of the model to the data.

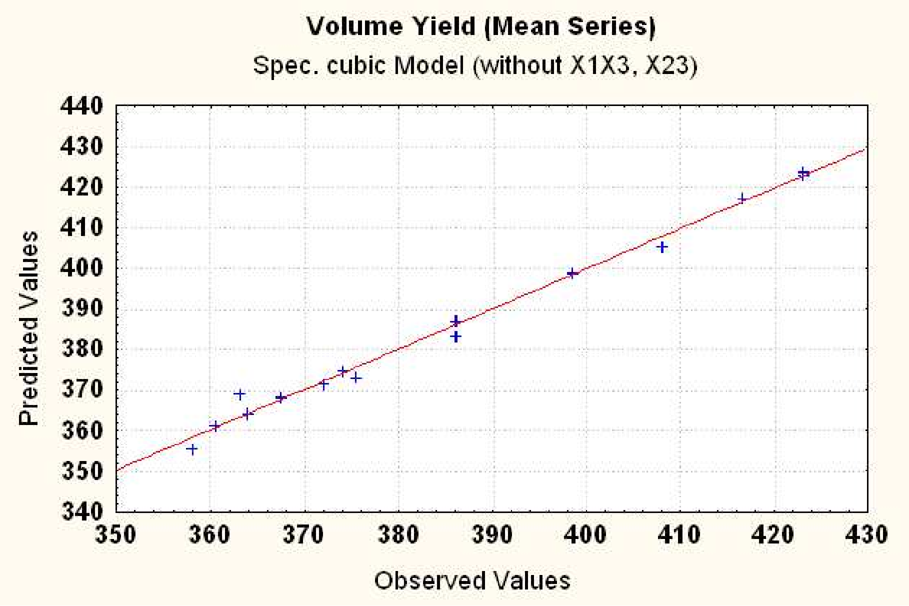

Figure 2:

The following graphics refer to the special cubic model without the two-component blends X1X3 and X2X3. The Pareto-Chart in Figure 3 shows the standardized coefficients of the so called (scale-independent, comparable) Pseudo-Components, sorted by their absolute magnitudes and the results of their significance tests.

The fitted model can be visualized by contour and surface plots (Figures 4 and 5). These plots help to locate the optimal region of the response variable. Here the three components X2 (BS2), X3(BS3) and X4(BS4) were chosen, because the costs of producing BS1 were higher than for the other three roughages and the results of optimization the volume yield comparable. The contour lines represent combinations of the three components. A high volume yield can be found, if the share of X2 (BS2) in the mixture is high.

The trace plot of the estimated response variable (Figure 6) shows the influence to the volume yield, if each component will be increased or lowered, starting from a reference blend (here: X1(BS1)=0; X2(BS2)=0,7; X3(BS3)=0,1; X4(BS4)=0,2). If a selected component will be changed, the other components

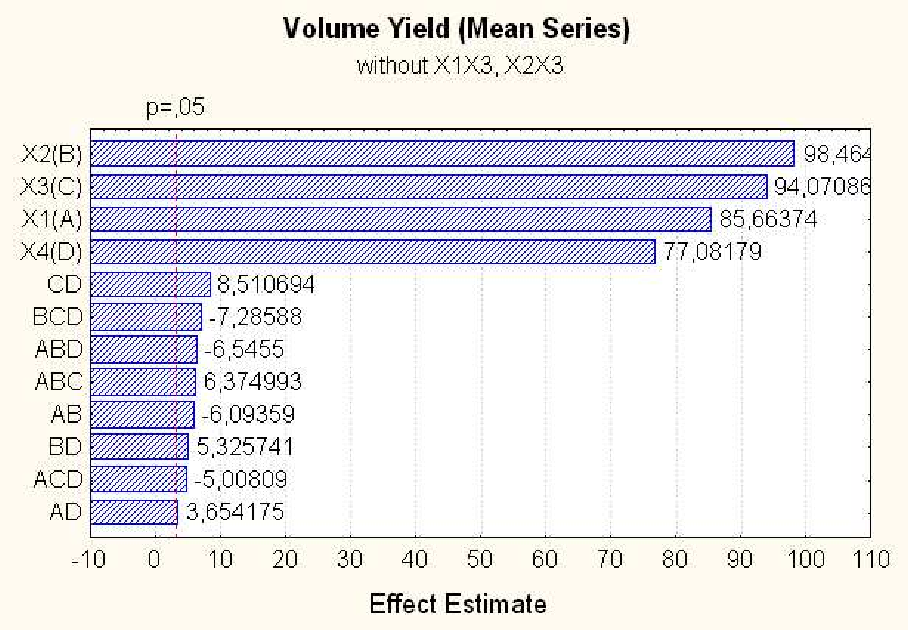

Figure 3

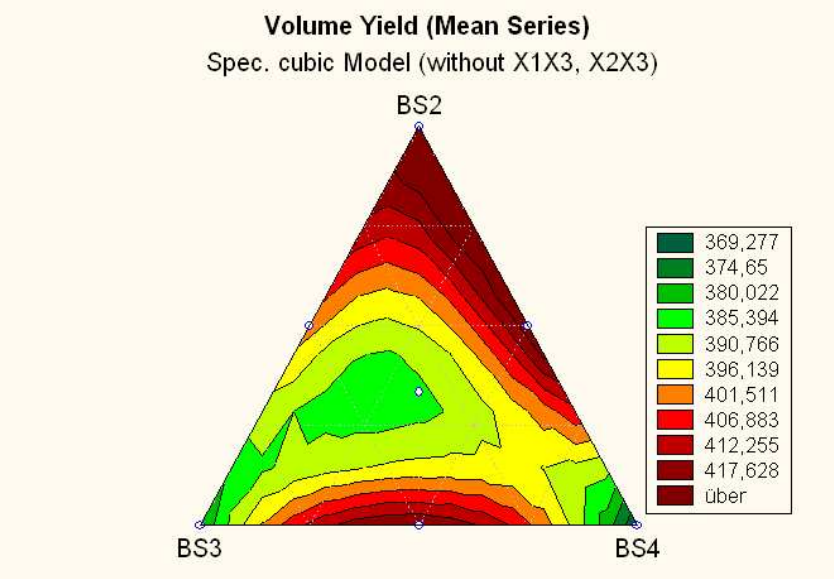

Figure 4

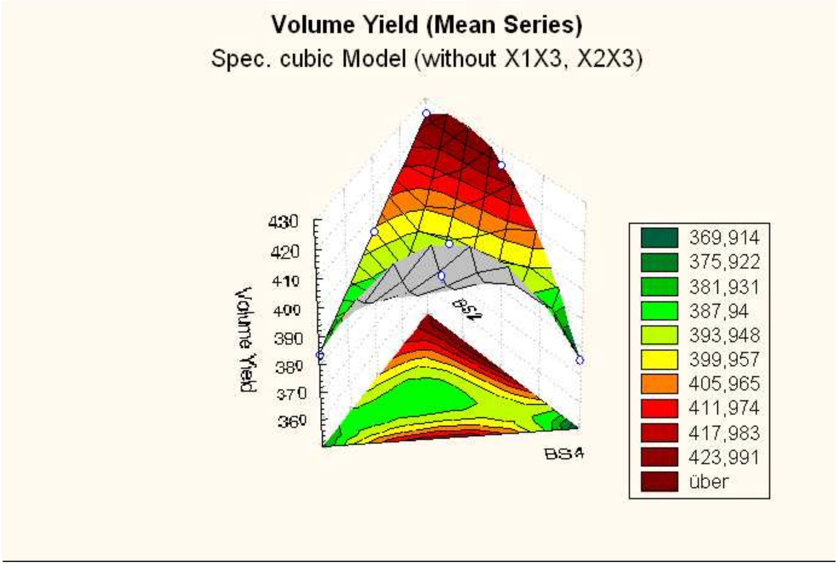

Figure 5

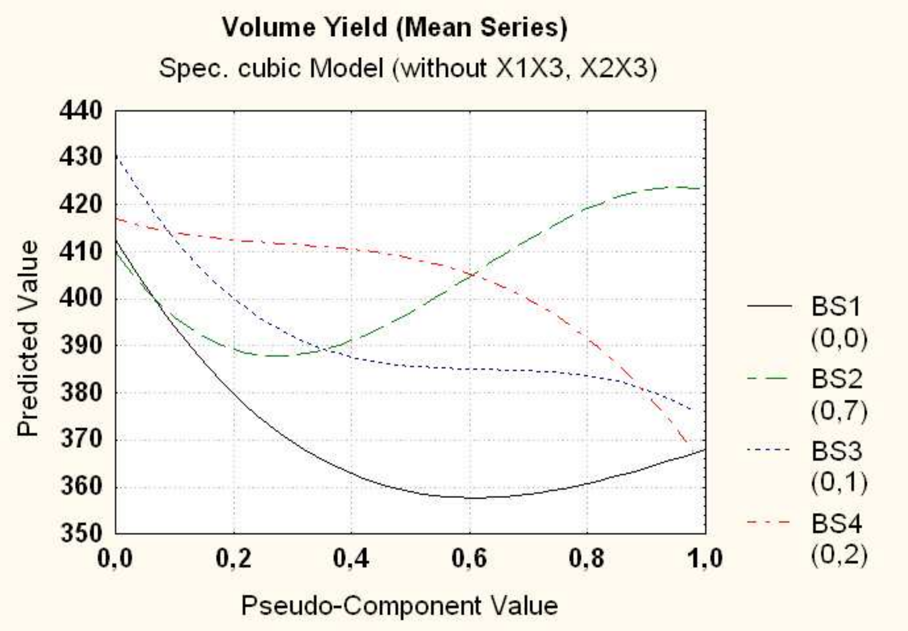

Figure 6

will be fitted proportionally. Thus a further increase of the component X2(BS2), while the other components are fitted, also result in an increase of the volume yield.

3.4. Interpretation of the results and conclusions

As expected the rusk without roughages (Experiment No. 0) has shown the highest volume yield. By adding only one roughage compound flavour changes were realized. From technological and financial point of view only the production of a gluten-free rusk which is enriched by three or two component mixtures of roughage compounds is useful. The costs for producing BS1 were higher than for the other roughages and the results of optimization the volume yield similar. That's why the three roughage compounds BS2, BS3 and BS4 were chosen.
An optimal volume yield (Y= 430,83 ml/100g) on condition that sensory parameters fulfil quality standards can be reached, if the mixture consists of roughage compounds in the following ratio: BS2 = 74,5%, BS3 = 0,1% and BS4 = 25,4%. Of course companies, which produce bread and cakes will prefer the two component variant with BS2 = 85% and BS4 = 25%. By control experiments the same good results could be realized.

REFERENCES

[1] Box, G.E.P., W.G. Hunter, J.S. Hunter Statistics for experimenters. J. Wiley & Sons, New York, 1978.

[2] Coleman, D.E., D.C. Montgomery A Systematic Approach to Planning for a Designed Industrial Experiment. *Technometrics* (1993), **35**, 7.

[3] Cornell, J.A. Experiments with mixtures: designs, models, and the analysis of mixture data. 2nd ed., J. Wiley & Sons, New York 1990.

[4] Montgomery, D.C. Design and analysis of experiments. 2nd ed., J. Wiley & Sons, New York 1991.

[5] Römisch, U. and M. Valentin Untersuchung von Ballaststoffpräparaten in glutenfreien Gebäcken durch die Anwendung eines Versuchsplanes fr Mixturen. *Getreide Mehl und Brot* (2000), **54**, 1.

[6] Scheffé, H. The simplex-centroid design for experiments with mixtures. *J. of Royal Stat. Soc.* (1963), **B25**, 235–263.

[7] StatSoft, Inc. (1998). STATISTICA for Windows [Computer programme manual]. StatSoft, Inc., 2300 East 14th Street, Tulsa, OK 74104: Tulsa, OK.

[8] Scheffler, E. Statistische Versuchsplanung und -auswertung. 3. Auflage, Dt. Verlag für Grundstoffindustrie, Stuttgart 1997.

[9] Vuchkov, I.N. and L. Boyadjieva Quality improvement with design of experiments. Kluwer Acad. Publ., Dordrecht 2001.

Ute Römisch
TU Berlin
Fac. of Process Sciences and Engineering
Dept. of Informatics
Gustav-Meyer-Allee 25
13355 Berlin, Germany
e-mail: `ute.roemisch@tu-berlin.de`