Reconciling mixture designs and factorial designs in order to identify best recipes in a holistic way

Andreas Rytz[1], Mireille Moser, Mélissa Lepage, Cassandra Mokdad, Marie Perrot, Nicolas Antille, Nicolas Pineau

[1] *Nestlé Research Center, Lausanne, Switzerland*
E-mail : *andreas.rytz@rdls.nestle.com*

Abstract
All food recipes are composed of three elements: ingredients, their respective amounts, and the way to process them. The art of a chef consists in mastering these three elements in a holistic way in his kitchen in order to deliver most pleasurable food. The challenge of the food industry is to do the same, but on a scale where intermediate tasting and adjusting is almost impossible. As a consequence, Design of Experiments (DoE) are commonly used to model the impact of mixture properties (e.g. mixture designs) or of process factors (e.g. orthogonal arrays) on the nutritional and sensorial characteristics of the foods. Only few approaches combine mixture and process in an efficient way (e.g. coverage, optimal) but they require defining global multivariate spaces that somehow artificially combine mixture and process parameters. This paper proposes a combined approach - based on orthogonal arrays - that projects ingredients on the mixture hyperspace in a very intuitive way. It is illustrated through an example that allowed significantly improving the nutritional value of a cereal product (i.e. 50% sugar reduction), while increasing consumer liking.

Keywords: Design of Experiment (DoE), orthogonal arrays, mixture designs, cereals, nutrition, sugar reduction.

1. Introduction

Very schematically, a basic all-family cereal recipe is made of 80% wheat flour and 20% sucrose that is mixed, soaked, cooked and dried. In order to improve the nutritional value of the product, it was asked to reduce its final sugar content by 50%, without using any artificial sweetener or flavor, while maintaining highest consumer liking.

This presentation describes the two-step approach that was used to achieve this goal: 1) test if natural generation of biscuit and caramel notes during the process could be a winning strategy thanks to their congruency with perceived sweetness and 2) optimize nutritional offer and achieve taste superiority.

These two steps lead to a final product containing 75% wheat flour, 15% whole grains and 10% sugars.

This example is used to illustrate the efficiency of our very intuitive way of reconciling mixture and factorial parameters in one experimental design.

2. Method to reconcile mixture and factorial parameters

2.1 Projecting three mixture parameters of a 2^{7-4}

2.1.1 Constructing the design, including projection

Biscuit and caramel flavors can be generated during cooking and drying if the soaked mix contains the necessary precursors (i.e. free reducing sugars, amino acids) and an appropriate humectant.
The commonly used humectant is water, but it is known that the underlying chemical reaction could be triggered if adding another humectant to the water. As a consequence, humectant is a factorial parameter with two levels (water, water+).
Amino acids cannot be added as such as ingredients, but they have to come from natural sources (i.e. 3 grains or pulses with specific amino acid profiles are tested). Since it is not known which amino acid cocktail leads to best flavor generation, it is proposed to test various combinations of wheat replacement. Independent replacement of wheat flour by respectively 2% of each of these three sources is handled as three factorial parameters.
The three reducing sugars (A, B, C) of interest for the reaction are all available as natural ingredients and they can simply replace part of the sucrose in the recipe. It is known that 3% of reducing sugars is required to generate biscuit and caramel notes (Illmann. et.al, 2009). The three reducing sugars are therefore considered as a mixture summing to a constant of 3% in final recipe, and following bounds were established for cost and nutritional constraints (expressed as % reducing sugars): A=0-58%, B=34-85% and C=12-47%. Considering that B and C sum to a minimum of 46% (=34+12), ingredient A cannot exceed 54%. As a consequence, the experimental region is defined by the new bounds A=0-54%, B=34-85% and C=12-47%.

This experiment therefore features 1 humectant (factorial), 3 potential flour replacements (factorial) and 3 reducing sugars (mixture). These 7 parameters were investigated using a 2^{7-4} saturated design (Hedayat et.al, 1999), as shown in table 1.

	D_{Init}							$D_{Intermediate}$				D_{Final}						
	A	B	C	AB	AC	BC	ABC	A	B	C	vs. *c*	A Sugar A	B Sugar B	C Sugar C	AB Grain 1	C Grain 2	BC Grain 3	ABC humect.
C1	0	0	0	0	0	0	0	0	34	12	-54	21	54	25	0	0	0	W
C2	1	0	0	1	1	0	1	54	34	12	0	54	34	12	2	2	0	W+
C3	0	1	0	1	0	1	1	0	85	12	-3	2	85	13	2	0	2	W+
C4	1	1	0	0	1	1	0	54	85	12	51	28	60	12	0	2	2	W
C5	0	0	1	0	1	1	1	0	34	47	-19	10	43	47	0	2	2	W+
C6	1	0	1	1	0	1	0	54	34	47	35	33	34	33	2	0	2	W
C7	0	1	1	1	1	0	0	0	85	47	32	0	66	34	2	2	0	W
C8	1	1	1	0	0	0	1	54	85	47	86	21	54	25	0	0	0	W+

Table 1: Design 2^{7-4} with 3 projected mixture parameters

The projection of the mixture part is a pragmatic simplification of the approach by Box and Hau (2001) and is described below:

- Start with the initial orthogonal array D_{Init}. For each of the n experiments and each of the q ingredients, replace the low level by the lower bound a_i and the high level by the upper bound b_i, leading to an intermediate design $D_{Intermediate}$ with elements p_{ti} (t=1...n, i=1...q). In $D_{Intermediate}$, none or almost none of the mixtures sum to the constant total amount c.
- Transform $D_{Intermediate}$ into D_{Final} by adjusting the mixtures according to their excess or lack vs. total amount c:

- In case of a mixtures summing to less than c, let the lack of total amount be w-. Focus on the low level setting components p_{ti} of this mixture. One has to increase by w- the sum of the selected levels. In order to accomplish it, allocate a portion of w- to each of these levels proportionally to the range of variation of their respective components.
- In case of a mixture t in excess of amount w+, the principle remains the same. In this situation, one has simply to select the high levels p_{ti} of the mixture. And then decrease the selected high levels proportionally to the range of variation of their respective components.

2.1.2 Comparing the projected mixture part with other approaches

The proposed design is fully saturated and it allows to estimate a first order model for the three mixture parameters and main effects of factorial parameters. When considering the coverage of the mixture, the proposed design covers the experimental region relatively homogeneously. It is therefore closer to a space-filling logic than a D-Optimal logic (i.e. a D-optimal design would select the 3-4 most extreme points for a first order model).
In order to better compare our approach to four common approaches that cope with constraints in mixtures (Cornell, 2002), let us consider the projected mixture part only and let us consider a 2nd order model (for which our projection would be the same). The resulting designs are visualized in figure 1.

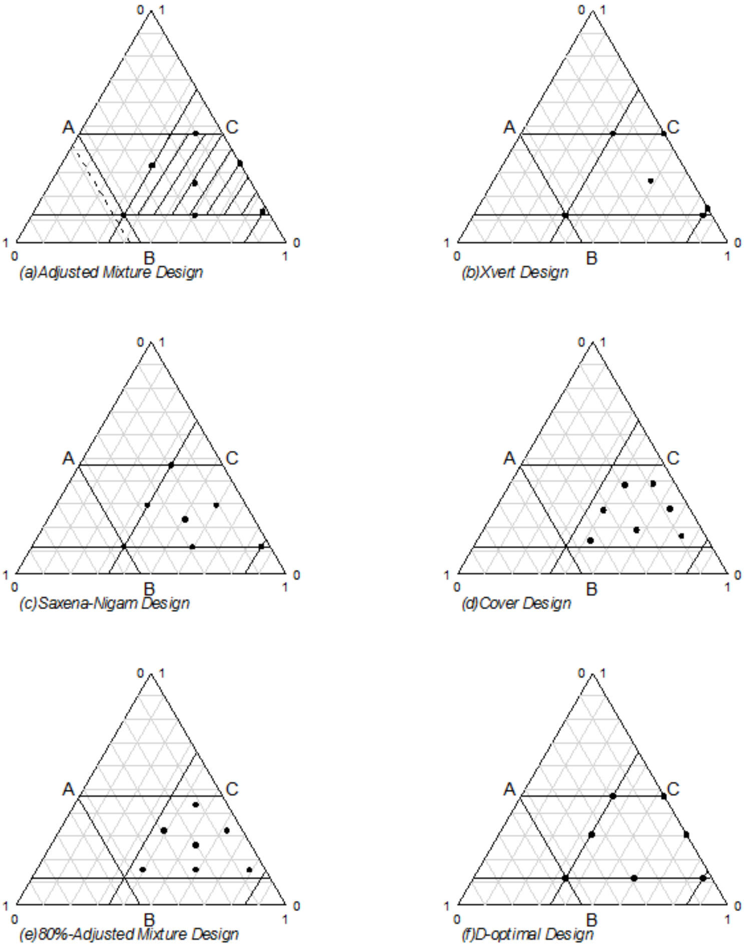

Figure 1: Comparing 6 mixture designs for a 2nd order model

The proposed design leads to a very good compromise between D-efficiency (Atkinson, 1992) and coverage (Johnson et.al., 1990), as shown in table 2: it is the only technique yielding values higher than 0.5 for both indexes.

	D-efficiency (2nd order)	Coverage Index
(a) Adjusted Mixture	0.513	0.579
(b) Extreme Vertices	0.070	0.300
(c) Saxena-Nigam	0.215	0.588
(d) Space-Filling	0.001	1.000
(e) Adjusted Mixture (80%)	0.008	0.944
(f) D-optimal (2nd order)	1.000	0.435

Table 2: Comparing design strategies using D-efficiency and Coverage Index

2.2 Projecting five mixture parameters of a 2^{9-5}

The first design allowed to show that most interesting biscuit and caramel notes (i.e. most congruent with sweetness perception) were generated using 3% of the second tested sugar, that all three additional grains were useful (i.e. all amino acids were required for the best reaction), but no clear findings could be made for the humectant part. The necessary addition of three flours was a good opportunity to move from a pure wheat-based product to a multi-cereal product with interesting levels of whole grain.
The second design therefore was about optimizing nutritional offer (i.e. maximizing whole grain content at a 50% reduced final sugar content) and achieve taste superiority.
Five whole grains were tested summing to a total of 15% of the final recipe, but with different levels depending on nutritional and known taste impact (Grain1=3-10%, Grain2=2-10%, Grain3=2-5%, Grain4=0-2% and Grain5=0-8%).
In parallel, 4 process parameters (i.e. humectant, soaking duration, cooking condition, and drying condition) were investigated for main effects and 2x2 interactions.
The initial design was a 1/32 fraction of a 2^4x2^5 factorial design, which is a compromise plan of class one given by Addelman (1962). This saturated design covers the experimental region very well with as few as 16 different mixtures and allows investigating main effects and 2x2 interaction for the process part (A, B, C, D of the underlying Yates table) as well as a first order model for the mixture part (ABC, ABD, ACD, BCD, ABCD of the underlying Yates table).
The projection of the mixture part follows the same approach and the final design is shown in table 3.

	Humect.	Soaking duration	Cooking condition	Drying condition	Grain1	Grain2	Grain3	Grain4	Grain5	∑Grains
P01	W	short	mild	mild	2.0	2.0	3.0	0.0	8.0	15.0
P02	W	short	mild	strong	2.0	5.8	6.3	0.9	0.0	15.0
P03	W	short	strong	mild	4.0	2.0	7.7	1.3	0.0	15.0
P04	W	short	strong	strong	3.3	5.4	3.0	0.0	3.4	15.0
P05	W	long	mild	mild	3.8	6.9	3.0	1.2	0.0	15.0
P06	W	long	mild	strong	3.3	2.0	6.1	0.0	3.6	15.0
P07	W	long	strong	mild	2.0	4.8	5.4	0.0	2.8	15.0
P08	W	long	strong	strong	2.7	3.8	4.6	2.0	1.8	15.0
P09	W+	short	mild	mild	3.3	5.6	6.1	0.0	0.0	15.0
P10	W+	short	mild	strong	3.8	2.0	3.0	1.2	4.9	15.0
P11	W+	short	strong	mild	2.0	5.6	3.0	0.9	3.6	15.0
P12	W+	short	strong	strong	2.1	2.4	10.0	0.1	0.4	15.0
P13	W+	long	mild	mild	2.0	2.0	6.3	0.9	3.8	15.0
P14	W+	long	mild	strong	2.0	10.0	3.0	0.0	0.0	15.0
P15	W+	long	strong	mild	5.0	3.6	4.4	0.4	1.6	15.0
P16	W+	long	strong	strong	2.9	4.3	5.0	0.6	2.3	15.0

Table 3: Design 2^{9-5} with 5 projected mixture parameters

3. Discussion and conclusion

Although most foods can be described as a processed mix of ingredients, few design of experiment techniques efficiently combine mixture and process parameters. We propose a very simple, intuitive approach basing on the partial projection of orthogonal arrays on a mixture hyperspace, following the ideas of Box and Hau.

This approach appears to be extremely easy to implement even for very complex cases such as products with multiple mixtures (e.g. milk chocolate is a mixture of fat, protein and sugar, but fat is a mixture of various sources of fats, protein a mixture of various sources of proteins and sugar a mixture of various sugars).

It is easy to show that for non-constrained mixtures, the approach leads to simplex-lattice designs. In case of constrained mixture, it is shown that this approach generally leads to best consensus between coverage and D-efficiency.

The approach is illustrated through an example that allowed, in 24 trials (8+16), to significantly improve the nutritional value of a cereal product, moving from a basic product containing 80% wheat and 20% sucrose to a nutritional proposition, that achieves higher liking, and that features 75% wheat, 15% whole grain and 10% sugars.

The same design approach has been successfully applied to other food categories including confectionary, dairy, ice-cream, culinary, beverages and pet-food.

References

Addelman, S. (1962) Symmetrical and asymmetrical fractional factorial plans. Technometrics, 4(1), 47-58.

Atkinson, A.C., Donev, A.N. (1992) Optimum Experimental Designs. Oxford Statistical Science Series, vol. 8 Oxford: Clarendon Press.

Box, G.E.P., & Hau, I. (2001) Experimental designs when there are one or more factor constraints. Journal of Applied Statistics, 28(8), 973-989.

Cornell, J.A. (2002) Experiments with Mixtures. 3rd edition. New York: Wiley.

Hedayat, A.S., Sloane, N.J.A., Stufken J. (1999) Orthogonal Arrays, Theory and Applications. Springer, New York.

Illmann, S, Davidek, T., Gouezec, E., Rytz, A, Schuchmann, H., Blank,I. (2009) Generation of 4-Hydroxy-2,5-Dimethyl-3(2H)-Furanone from rhamnose as affected by reaction parameters: Experimental Design Approach, J. Agric. Food Chem. 57, 2889–2895

Johnson, M.E., Moore, L.M., & Ylvisaker, D. (1990) Minimax and maximin distance designs. Journal of statistical Planning and Inference, 26, 131-148.