Modeling the inhibition of Salmonella typhimurium growth by combination of food antimicrobials

Andre Kobilinsky, A. I. Nazer, Florence Dubois-Brissonnet

Available online at www.sciencedirect.com
ScienceDirect
International Journal of Food Microbiology 115 (2007) 95–109
INTERNATIONAL JOURNAL OF Food Microbiology
www.elsevier.com/locate/ijfoodmicro

Modeling the inhibition of *Salmonella typhimurium* growth by combination of food antimicrobials

A. Kobilinsky [a,*], A.I. Nazer [b], F. Dubois-Brissonnet [b,1]

[a] *INRA, Biométrie (MIA), Domaine de Vilvert., F-78352 Jouy-en Josas Cedex, France*
[b] *ENSIA, UER Qualité et Sécurité Microbiologiques des Aliments et Procédés, 1 av. des Olympiades, F-91744 Massy Cedex, France*

Received 23 November 2005; received in revised form 31 August 2006; accepted 8 October 2006

Abstract

Through the use of a sequence of fractional factorial designs, the growth inhibition of *Salmonella typhimurium* by many natural antimicrobial compounds is studied and modeled. Two very important predictive variables are an appropriately weighted total of organic acid concentrations on the one hand, and of aromatic compound concentrations on the other.

Keywords: Synergy; Growth inhibition; Acidic compounds; Aromatic compounds; Experimental design; Smoothing; *Salmonella Typhimurium*

1. Introduction

Nowadays, consumers request additive-free, fresher and more natural tasting food products, while maintaining microbiological safety (Gould, 1996). The use of natural antimicrobial compounds, which belong to the general framework of preservation processes, is thus of utmost interest in the food industry. Organic acids and aromatic compounds belong to this type of additives, as well as some salts.

The main drawback to the use of such a compound when it is used alone as food preservative is the high effective concentration needed for a lethal effect on the spoilage microbiota: it often exceeds the threshold acceptable to consumers. In that context, associating several of these compounds can be interesting. Combinations of preservative effects including antimicrobials addition have been already described as the *Hurdle Technology* concept in food safety by Leistner (1985) (Gould et al., 1995). This process for food stabilisation allows the same microbiological security in food by using a much lower amount of each preservative compound or a lower intensity of the physical treatment involved.

This paper describes an approach using fractional experimental designs to find which combinations of compounds used at moderate levels are the most effective in preventing the growth of *Salmonella typhimurium* which is of major concern to public health and represents one of the most important serovars in *Salmonella* gastroenteritis in almost all countries (Ray, 2001; Moll and Moll, 2002). We here mainly consider the statistical aspect of the study. See Nazer et al. (2005) for other aspects.

2. Materials and methods

2.1. Antimicrobial agents

The compounds used in the study, the abbreviations used to denote them here and their origin are listed in Table 1. Stock solutions of organic acids and phosphates were made in distilled water. Stock solutions of aromatic compounds were prepared by dissolution in absolute alcohol.

2.2. Strain and growth conditions

The strain used in this study was *Salmonella enterica* subsp. *enterica* serovar Typhimurium ATCC 13311. The bacteria were

* Corresponding author. Tel.: +33 1 34 65 22 30.
E-mail addresses: Andre.Kobilinsky@jouy.inra.fr (A. Kobilinsky), dubois@ensia.fr (F. Dubois-Brissonnet).
[1] Tel.: +33 1 69 53 64 72.

0168-1605/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2006.10.039

Table 1
Antimicrobial compounds

Compound	Abbreviation	Provider
Acetic acid	acetic	Labosi (France)
Citric acid	citric	Sigma-Aldrich (USA)
Lactic acid	lactic	Merck (Germany)
Benzoic acid	benzoic	Sigma-Aldrich (USA)
Paraminobenzoic acid	paba	Accros Organic (France)
Sodium acetate	asod	Labosi (France)
Potassium acetate	apot	Labosi (France)
Sodium lactate	lacso	Sigma-Aldrich (USA)
Sodium nitrite	nitsod	Prolabo (France)
Pyropolyphosphoric acid	aphos	Sigma-Aldrich (USA)
Sodium hexametaphosphate	shmp	Sigma-Aldrich (USA)
Sodium tripolyphosphate	stpp	Sigma-Aldrich (USA)
Geraniol	geran	Sigma-Aldrich (USA)
Carvacrol	carvac	Sigma-Aldrich (USA)
Eugenol	eug	Sigma-Aldrich (USA)
Thymol	thymol	Sigma-Aldrich (USA)
Citral	citral	Sigma-Aldrich (USA)
Menthol	menth	Sigma-Aldrich (USA)
trans-cinnamaldehyde	cinnam	Sigma-Aldrich (USA)
a-terpineol	a-terp	Sigma-Aldrich (USA)

grown in BHI broth (Brain–Heart Infusion, Oxoid, UK) in the absence or presence of antimicrobials. The medium was inoculated at 1% v/v (approximately 10^6 cells/mL) with a standardized inoculum.

Growth was monitored with a Bioscreen Microbiological Growth Analyser (Bioscreen C, Labsystems, Finland), an apparatus using 100-well microplates which can thus follow up to 100 growth curves simultaneously by measuring the optical density (OD). In order to analyse the growth, the growth percentage at 12 h of culture was measured (Nazer et al., 2005). It is defined as:

$$G = \frac{(\mathrm{OD}_t - \mathrm{OD}_{t_0})_{\mathrm{test}}}{(\mathrm{OD}_t - \mathrm{OD}_{t_0})_{\mathrm{control}}} \quad (1)$$

where OD is the optical density at 600 nm, t=12 h, t_0=0 h, test makes reference to the culture grown with antimicrobials, and control makes reference to the culture grown without antimicrobial. This variable G indicates how much the growth is reduced in the presence of antimicrobials. The time 12 h has been chosen for the best discrimination of growth curves.

2.3. Experimental designs and statistical analysis

For this problem, an usual experimental approach starting with a low resolution fractional design to screen the factors followed by designs of higher resolution to model the influence of the more important factors was used. The analysis of each design led to the next one. Through a final global analysis, the action of the compounds under investigation, which at the beginning of the study were the 20 listed in Table 1, was modeled. This statistical approach is described in Sections 4 and 5.

3. Determination of antimicrobial activity of individual compounds

To define the levels used in the experimental designs, a first study was made on each compound separately. The growth percentage G at 12 h was represented as a function of the concentration and this graph was used to determine first visually the 5% inhibitory concentration, denoted by IC5, which is the concentration reducing the growth by 5% with respect to the control. This visually obtained IC5 was then used as a reference to determine the levels used in each design. For instance the levels used in the first screening design were IC5/8 and IC5/4 (see Section 4.1).

As a particular case, graphs of G as a function of the concentration for 5 aromatic compounds on one side, 4 organic acids on the other are shown in Figs 1,a and 2,a. The curves in Fig. 1,a seem to have similar shapes, up to a change of scale on the concentration prompting a search for suitable changes of scale to make the inhibition curves for the aromatic compounds coincide as much as possible. The following procedure was used to determine these changes of scale.

First for each compound, the curve giving the growth percentage as a function of the concentration was smoothed by local polynomial approximation (Fan and Gijbels, 1996), as

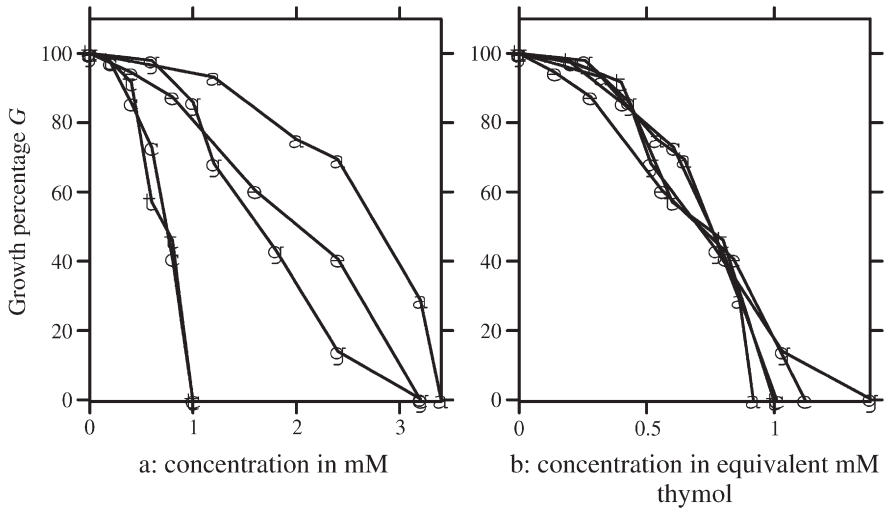


Fig. 1. Growth percentage G at 12 h for each aromatic compound e: eugenol, c: carvacrol, g: geraniol, t: thymol, a: citral.

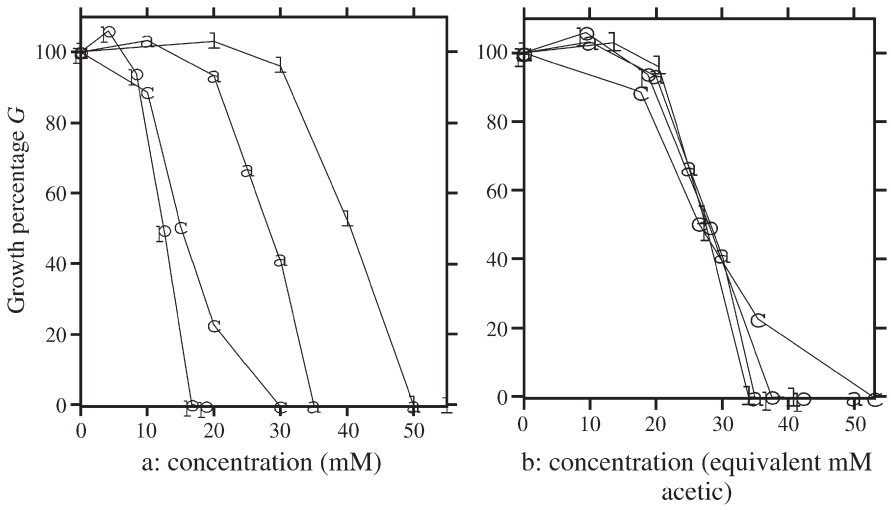


Fig. 2. Growth percentage G at 12 h for each acidic compound a: acetic, c: citric, l: lactic, p: pyropolyphosphoric.

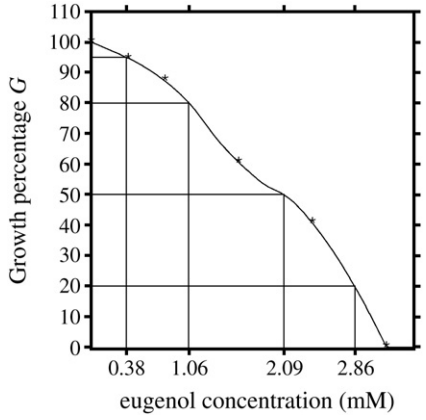
summarised in Section 3.1. The smoothed curve was then used to estimate the inhibitory concentrations at different levels (IC5, IC20, IC50, IC80) and the scales were then chosen so as to make these inhibitory concentrations coincide as much as possible as described in Section 3.2. The equivalences thus found between aromatic compounds (given in Section 3.3) suggest introduction, as a predictive variable in the model, of their total after the changes of scale.

The strong analogy between the curves in Fig. 2 similarly leads to the introduction of a total organic acid concentration, after suitable changes of scale. Section 3.3 explicitly gives these two new global explaining variables, the total aromatic and the total acid concentrations, the use of which then allowed development of much simpler models.

3.1. Detail of the smoothing

The local polynomial fitting (Fan and Gijbels, 1996) leads to a smoothed curve such as the one on the left of Fig. 3. If (x_i, y_i) are the pairs of observations, where x stands for the concentration and y for the growth percentage, the adjusted value at x is obtained by a weighted polynomial regression of the y_i on the x_i. The weights decrease with the distance between x_i and x. They are of the form $K((x_i-x)/h)$ where K, the kernel function, is maximum at $x_i-x=0$ and decreases symmetrically as the argument increases in absolute value. So for each x, the method finds a polynomial $P_x(z)$ locally adjusting the observed points and the value $P_x(x)$ of this polynomial at x is precisely the smooth ordinate.

A thorough discussion on the choice of the kernel function K, of the parameter h, of the degree of the polynomial regression can be found in Fan and Gijbels (1996). In our case, the kernel was chosen as Gaussian, and the choice of the degree and of h was made so that the smoothed curve nearly goes through the few observed points. The degree 2 together with h equals to 1/8 of the concentration which completely inhibits the growth that was found adequate for that purpose. As an example, let us consider the case of eugenol for which the tested concentrations and corresponding growth percentages are given in Table 2. In that case, h is taken as equal to $0.4=3.2/8$ and thus the weight associated to the point of the abscissa x_i (where x_i takes the values 0, 0.4, 0.8, 1.6, 2.4, 3.2) is: $\exp -0.5[(x_i-x)/0.4]^2$. For instance if $x=1$, the corresponding weights are 0.044, 0.325, 0.882, 0.325, 0.002, 0. The local adjustment mainly takes



	IC5	IC20	IC50	IC80
acetic	18.8	22.8	28.2	32.4
citric	8.5	11.2	15.1	20.6
lactic	29.8	34.2	40.3	45.9
aphos	8.1	9.9	12.4	14.8
eug	0.4	1.1	2.1	2.9
carvac	0.25	0.51	0.75	0.91
geran	0.7	1	1.6	2.3
thymol	0.35	0.47	0.75	0.93
citral	1.1	1.8	2.9	3.3
asod	54.7	170.5	357.4	526.4
apot	49	156.8	365.3	572.9
lacso	130.4	424.8	620	754.1
shmp	18.6	32.7	32.7	32.7

Fig. 3. 5%, 20%, 50%, 80% Inhibitory Concentrations of the Growth percentage at 12 h. left: smoothed curve and ICs in the case of eugenol, right: table of the ICs for the main compounds.

Table 2
Growth inhibition by eugenol

Concentrations (mM)	0	0.4	0.8	1.6	2.4	3.2
Growth percentage	100	94.5	87.5	60.5	40.7	0

into account the three points of coordinates 0.4, 0.8, 1.6 and since it is a polynomial of degree 2 which is adjusted, it nearly goes through these three points. This explains why, with such parameters, the smoothed curve nearly goes through the points.

3.2. Changes of scale

A simple dichotomic process was then used to find on the smooth curve the concentrations IC5, IC20, IC50, IC80 leading to the desired growth inhibition. As an example, we give on the left of Fig. 3 the fitted curve for the eugenol, with the corresponding ICs reported in the abscissa. On the right, we give the ICs thus found for the more important compounds.

Change of scale coefficients δ were determined so as to make the ICs in the new scale as near as possible, in the sense of minimizing the sum of square differences, to the corresponding ones for the reference compound, thymol for aromatics, acetic acid for acids. In this way, for example, 1 mM eugenol was found to be equivalent to $\delta = 0.35$ mM thymol. Multiplying by this coefficient the eugenol ICs 0.4, 1.1, 2.1, 2.9, we find equivalent thymol concentrations 0.14, 0.385, 0.735, 1.015 which are similar to the corresponding ICs 0.35, 0.47, 0.75, 0.93 obtained for thymol. Indeed, δ is the coefficient minimizing the sum of squares

$$S = (0.35-\delta\ 0.4)^2 + (0.47-\delta\ 1.1)^2 + (0.75-\delta\ 2.1)^2 + (0.93-\delta\ 2.9)^2.$$

A similar procedure was applied for acidic compounds.

Initially, we tried to find equivalence in the same way for all the compounds found active in the first screening design. They were thus all equivalenced to an acetic acid concentration with the hope that the total concentration of the compounds in this common unit would be a good predictor of the inhibition. But that global equivalencing failed, as well as less global ones trying to assimilate aromatics and acids, or acids and salts. Similarly we found that the salts cannot be all efficiently equivalenced, but that sodium acetate and potassium acetate, which are quite active and have very similar growth percentage curves, can be equivalenced.

3.3. Equivalence between compounds

The equivalences with thymol concentrations found by the method just described are

1 mM carvacrol ~ 1.01 mM thymol, 1 mM geraniol ~ 0.43 mM thymol, 1 mM citral ~ 0.27 mM thymol, 1 mM eugenol ~ 0.35 mM thymol.

The transformation of the Fig. 1,a curves after the change of scale, that is when all concentrations are expressed in equivalent mM of thymol, is shown in Fig. 1,b.

As already indicated, the strong similarity between the aromatic compounds suggests introduction of their total after the scale change as a predictive variable for the inhibition in the presence of several compounds. If thymol, carvacrol, etc denote the concentrations in mM, this total denoted by AR, is thus defined as

$$\mathrm{AR} = \mathrm{thymol} + 1.01\ \mathrm{carvacrol} + 0.43\ \mathrm{geraniol} + 0.27\ \mathrm{citral} + 0.35\ \mathrm{eugenol}.$$

A similar approach works for organic acids. Again the inhibition curves are very similar after all the concentrations have been transformed in the same unit *equivalent acetic mM* (Fig. 2,b). This leads to introduce the total

$$\mathrm{AC} = \mathrm{acetic} + 1.77\ \mathrm{citric} + 0.68\ \mathrm{lactic} + 2.25\ \mathrm{aphos}$$

as a predictive variable.

4. Experimental designs

4.1. The first screening design of resolution 4

It was decided to study the 20 compounds in 64 wells, each at two levels which after a first trial were established as IC5/8 and IC5/4 to give, when the compounds are combined together, results varying between 0 and 100% of growth percentage G at 12 h (IC5 was in that case determined as explained in the beginning of Section 3). It is well known that it is possible to find in this context a resolution 4 design, authorizing the optimal unbiased estimation of all main effects even when there are two-factor interactions in the model. Such designs are known to be far more robust than resolution 3 designs, currently known as Plackett and Burman designs, which are deduced from Hadamard matrices (Plackett and Burman, 1946; Diamond, 1981).

To choose an adequate fraction, it would have been useful to get all resolution 4 regular 2^{20-14} possible fractions, and to compare their properties in term of aliasing of the two-factor interactions. However even with the algorithm of Draper and Mitchell (1967), described in some details by Kobilinsky (1997), this cannot be achieved in a reasonable time. So the investigation was limited to an assortment of such resolution 4 designs, obtained by the software PLANOR (Kobilinsky, 1994; Kobilinsky, 1997) which makes possible backtrack searches of regular fractions in a random order, and, thus, to find several solutions.

However the direct search was very long and stopped after obtaining the first solution. It was then decided to look for resolution 3 regular fractions, and to double them by their opposite to get the searched resolution 4 fractions (Kobilinsky, 1997; Montgomery, 1997; Ankenman, 1999). Four solutions were thus derived from four different resolution 3 fractions 2^{20-15} and four others from resolution 3 fractions 2^{19-14} to which was added a supplementary factor equal to 1 on the first half and −1 on the opposite part.

In each of the 9 fractions thus obtained, there are 31 groups of aliased two-factor interactions. Recall that in each group, only one linear combination of the two-factor interactions is

estimable. This function is the associated *canonical estimable function*, abbreviated CEF (Kobilinsky and Monod, 1995).

Since the factors were a priori considered as equally important, it was decided to select a fraction in which the groups of aliased effects are approximately of the same size. One avoids in this way too big groups that make the analysis more difficult if the associated CEFs are significantly different from 0. Table 3 gives, for three of the compared fractions, the number of groups of aliased two-factor interactions (2fi) of each size. The one selected is the third where 24 of the 31 groups include 6 two-factor interactions. The corresponding effects deduced from the results of the design are reported just below, sorted by order of decreasing absolute values. They are normalized as explained in Appendix A (see also Kobilinsky, 1997) and may thus be compared.

Moreover, since the fraction is regular, they all have the same standard deviation so that the 95% (or 99%, or 99.9%) confidence intervals all have the same width, the half of which is given at the right top of Table 3. By adding and subtracting these half widths, the desired confidence interval is found and it is thus immediately seen which effects are significant at the 5% (or 1%, or 0.1%) level. For instance the effect "asod" (sodium acetate) has a 99.9% confidence interval -2.49 ± 2.24, that is $[-4.73, -0.25]$, which does not includes 0. Thus it is significantly different from 0 at the 0.1% level, while "geran" which has a 99.9% confidence interval -1.87 ± 2.24 and a 99% confidence interval -1.87 ± 1.57 is only significantly different from 0 at the 1% significance level. Effects smaller in absolute value than 1.11 are not significantly different from 0 at the 5% level. Only the first of them "acetic.a-terp+⋯" is reported in Table 3 to show the limit between the effects which are significantly different from 0 at the 5% level and those which are not.

It is to be noted that the sign of an effect depends on the way the levels are coded in the analysis of variance. In that case the low level IC5/8 was systematically coded −1, while the high level IC 5/4 was coded 1. A main effect in Table 3 has, therefore, to be added if the corresponding compound is at its high level, subtracted otherwise. All the main effects except stpp (sodium tripolyphosphate) have a negative effect. Adding more of the corresponding compound therefore increases the inhibition (it adds a negative effect hence it lowers the growth percentage at 12 h). The 1% significant positive effect of stpp may indicate that this compound, which was found to be inhibiting when alone, can on the contrary lower the global inhibition in the presence of several other inhibiting compounds.

To evaluate the variation between microplates, the 64 wells were distributed among two microplates, with 32 wells in each. The division was selected so that the corresponding block effect (denoted by Bl in Table 3) can be estimated in the model including all two-factors interactions.

The analysis of variance of such a design is straightforward. Once the constant, the 20 main effects, the block effect and finally the 31 CEFs are estimated, there remains $11 = 64 - (1 + 20 + 1 + 31)$ degrees of freedom to estimate the residual variance. Note that the degrees of freedom used to compute the residual variance correspond in that case to interactions of strictly more than two factors. It is expected that even if some of them are not negligible, the corresponding main effects are quite bigger and therefore appear as significant when compared to this residual variance.

The results in Table 3 show no evidence of any strong two-factor interaction. The more important main effects are those of the four organic acids, particularly those of the citric and aphosphoric acids. Two salts, i.e. shmp (sodium hexametaphosphate) and sodium lactate (lacso), come after and then two

Table 3
Selection of the screening design and associated results

Comparison of fractions									Half width of confidence interval	
	Number of aliased 2fi	4	5	6	7	8	9	10	95%:	±1.11
Fraction 1	Number of groups	16	0	0	0	12	0	3	99%:	±1.57
Fraction 2	Number of groups	5	11	0	8	4	3	0	99.9%:	±2.24
Fraction 3	Number of groups	0	2	24	4	1	0	0		
…										

Effects and CEFs sorted by order of decreasing absolute values

general mean	55.43	shmp	−3.63	paba	−2.24
citric	−5.54	lacso	−3.39	stpp	1.98
aphos	−5.09	citral	−3.13	geran	−1.87
lactic	−4.3	thymol	−2.73	Bl	−1.8
acetic	−4	asod	−2.49	apot	−1.62
				carvac	−1.6
acetic · citric + benzoic · menth + apot · cinnam + lacso · citral + nitsod · a-terp + aphos · geran + eug · carvac					−1.54
a-terp					−1.49
lactic · paba + benzoic · aphos + asod · thymol + lacso · eug + shmp · a-terp + carvac · citral + geran · menth					−1.4
acetic · aphos + citric · geran + asod · a-terp + apot · carvac + shmp · thymol + eug · cinnam					−1.36
acetic · apot + citric · cinnam + lactic · thymol + benzoic · citral + paba · asod + lacso · menth + aphos · carvac + eug · geran					−1.25
eug					−1.22
benzoic					−1.15
acetic · a-terp + citric · nitsod + lactic · citral + benzoic · thymol + paba · carvac + asod · aphos					−1.01
					…

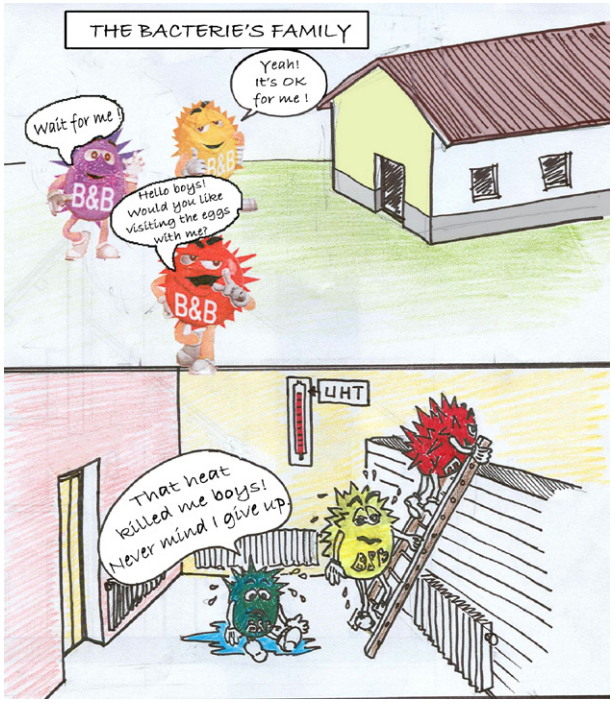


Fig. 4. Is hurdle technology a right concept? (Céline)

aromatic compounds, i.e. citral and thymol, while sodium acetate (asod) follows.

4.2. The designs of resolution 5 and more

These results prompted us to make a second experimental design including the four organic acids, two aromatic compounds, citral and thymol, and shmp and sodium acetate (abbreviated asod). The design chosen was a classical regular 2^{8-2} of resolution 5. This design gives unbiased optimal estimation of the main effects if it is assumed that there is no interaction between 4 or more factors and of the two-factor interactions if the model does not include interactions between 3 or more factors. Note that the mentioned property for main effects is valid even if there are three-factor interactions in the model making this design very robust.

The two levels were selected as IC5/6 and IC5/3, except for citral and thymol where the low level was chosen as equal to 0 in order to get mixtures including only acids and salts.

The results not shown here confirm the results previously found, and show the existence of some two-factor interactions smaller than main effects, mainly for pairs of acidic compounds. In fact, the mean inhibition obtained from the data where two such acidic compounds are at their high levels is greater than what would be predicted by simply adding the general mean and the main effects of these two acidic compounds.

However this result, well in accordance with the hurdle theory (Figs. 4 and 5), is not surprising. Fig. 2 clearly shows that the inhibition does not vary linearly with the concentration of an acidic compound. There is a threshold beyond which the inhibition strongly increases. Hence doubling an initial low concentration, for instance the IC5, gives an inhibition far more important than the inhibition 10%=5%+5% that would be explained by a linear model. If instead of putting two IC5 of a unique acidic compound, one adds two different acidic compounds, each at the IC5 level, it is therefore not surprising to get an inhibition stronger than the 10% inhibition that would result from an additive model.

This suggests introducing into the model for predicting the growth percentage at 12 h the two totals AR and AC defined in Section 3.3 (AR only involves here the two compounds citral and thymol). Recall that these totals are made only after all the units had been transformed to equivalent thymol or acetic acid concentrations by the method described in Section 3.2. A polynomial of degree 2 in these two totals, completed by the concentrations in sodium acetate and shmp gives with only 8 parameters in the model an adjustment as good as the ANOVA model including all main effects and two-factor interactions between the 8 compounds, that is 37 parameters.

Two other designs were then performed. A third one, a complete 2^5 factorial design replicated twice on two separate bioscreen plates, was done to study more specifically the aromatic compounds, whose efficiency clearly appeared in the first two designs and which are of great interest for the industry. Then, a fourth design was used which mixed the four acidic compounds with the three most efficient aromatic compounds. This latter one was a 2^{7-1} of resolution 7 thus allowing one to

Fig. 5. Is hurdle technology a right concept? (Céline)

Table 4
Models of growth percentage G as a function of aromatic compound concentrations

	Model	Nb.par	Res *df*	σ
1	$(\text{Eug}+\text{carvac}+\text{geran}+\text{thymol}+\text{citral})^2$	21	83	10.4
2	AR^2	3	101	11.4
3	$\text{AR}^2+\text{thymol}\cdot\text{carvac}+\text{citral}^2$	8	96	10.4
4	Gompertz (ln(AR))	2	102	12.4
5	Truncated Schnute (ln(AR))	3	101	11.2

The adjustment is made on the points of the two successive designs 3 for which $\text{AR} \leq 1$ (104 points).
Models 2 and 3 include all the subterms of those appearing, i.e. the constant and linear term AR and also the main effects thymol, carvacrol, citral in model 3.
Nb.par.: number of parameters in the model.
Res *df*: residual degrees of freedom = nb. of data points − nb.par.
σ: residual standard deviation.

study all three-factor interactions in a model without four-factor ones. In fact, the third design was performed twice. In the first attempt, the high level was IC5 and the low level IC5/2. The inhibition was judged globally too high and so the design was used a second time with the low level IC5/3 instead of IC5/2. Similarly the fourth one was made twice, with a low level of 0, a high level of IC5/3 the first time, IC5/2 the second time.

As in the second design, the totals AR in equivalent thymol mM and AC in equivalent acetic acid mM were computed and introduced in the analysis of variance of the third and fourth designs.

5. Statistical analysis and models

5.1. Polynomial factorial effects and orthonormal polynomials

In all the analysis with a polynomial model, orthonormal polynomials were used instead of monomials (Appendix B). The corresponding coefficients are the classical factorial polynomial effects: linear (lin), quadratic (quad), cubic (cub), lin · lin, lin · quad, etc.… They have a meaning that makes the analysis easier, and their estimates are usually not too correlated contrarily to the coefficients of the monomials in a classical polynomial regression. It is important to note that these orthonormal polynomials, and so the corresponding polynomial effects, are defined through a reference measure and not through the particular set of data point analysed (see Appendix B). This makes possible a comparison with the results of other experiments. Note that the same approach through a reference measure also provides efficient tools for the comparison of different designs (Goos et al., 2005). Moreover, when there are several factors, the reference measure is a product measure and, therefore, the definition of the polynomial effects does not depend on the order of introduction of the factors in an orthogonalisation process.

These orthonormal polynomials, which have long been used to simplify the computations in polynomial regression, were initially defined and tabulated in Fisher and Yates (1957) and Pearson and Hartley (1976). Although their interest for calculation has now disappeared, their use is still highly recommended in the general context of polynomial regression to obtain meaningful parameters whose estimates are the least correlated possible (Kobilinsky, 1988; Cliquet et al., 1994). As indicated in Appendix B, they are *normalized* so as to make their coefficients, that is the factorial polynomial effects, comparable.

5.2. Inhibition by aromatic compounds alone

The data coming from the two successive version of design 3 (with respective low levels IC5/2 and IC5/3) were merged together with the data giving the inhibition for pure aromatic compounds. Table 4 shows the residual standard deviation for some models adjusted by ordinary least squares on these data without the points for which $\text{AR} > 1$ which almost all lead to no observable growth at 12 h.

Model 1 is a polynomial of degree 2 in the 5 compound concentrations. This model has 21 parameters: the constant, 5 linear terms, 5 square terms and finally 10 products. With the only 3 parameters of the polynomial of degree 2 in AR (model 2), a slightly greater residual standard deviation is obtained. But by adding to these 3 parameters 5 other parameters (linear effects of thymol, carvacrol, citral, quadratic effect of citral and product thymol · carvacrol) one gets the same residual standard deviation with only 8 parameters. Finally, as done in Lambert and Pearson (2000), one can fit the non-linear model $\mathcal{G}(\ln(\text{AR}))$ where $\mathcal{G}$ is the classical sigmoidal function of Gompertz (model 4), or similarly the function $\mathcal{L}(\ln(\text{AR}))$ where $\mathcal{L}$ is the classical sigmoidal logistic function. However these two non-linear differentiable functions do not fit very well especially at values of AR near 1 where the growth becomes 0 (see Fig. 6). The function $\mathcal{S}(\ln(\text{AR}))$ where $\mathcal{S}$ is the Schnute function does not either fit properly unless it is modified to be exactly 0 after it comes down to 0. The functions $\mathcal{G}(t)$, $\mathcal{L}(t)$, $\mathcal{S}(t)$ are often used to described a bacterial growth as a function of time (Zwietering et al., 1990). The predictions made by models 2, 4 and 5 using only AR as predictor are displayed in Fig. 6.

The significant reduction of the residual standard deviation obtained when adding some terms in thymol, carvacrol and citral (model 3) shows that the aromatic compounds cannot be considered as strictly equivalent after rescaling. But from the practical point of view, it is sufficient to consider that there is no more growth as soon as AR reaches about 1 mM of equivalent thymol, and to approximate the growth percentage below $\text{AR} = 1$ by one of models 2 or 5, functions only of AR. Model 2 is inadequate for low level of AR, but this part of the curve is uninteresting for practical purposes. Of course, a local polynomial fitting as described in Fan and Gijbels (1996) could be used, but it does not lead to a meaningful formula.

A supplementary adjustment was made where the change of scale parameters given in Section 3.3 could also vary. But this adjustment did not consequently modify these scale parameters nor the quality of the global adjustment and is therefore not shown here.

5.3. Modeling the global inhibition

If the points of the different experimental designs are graphed with AC (the total acidic concentration) as the abscissa

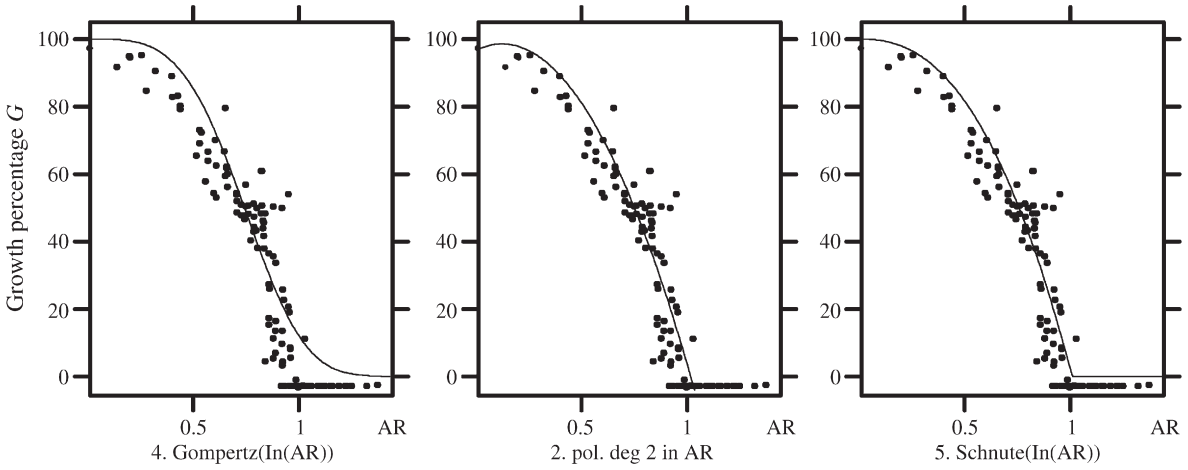


Fig. 6. Three adjustments of the growth percentage at 12 h as a function of AR.

and AR (the total aromatic concentration) as the ordinate, it clearly appears that only in design 3 where there are no acidic compounds, the total aromatic concentration exceeds 0.5 in equivalent thymol mM (Fig. 7).

It is, therefore, natural after the separate adjustment previously made for aromatic compounds, to make a global adjustment with all experimental points, including points with pure compounds, such that $AR \leq 0.5$. This adjustment was first made by fitting a polynomial. As seen in the case of aromatic compounds when $AC = 0$ (Section 5.2) this choice can be locally inadequate, but the adjustment is very easily handled even when there are many factors involved (Kobilinsky, 1988; Cliquet et al., 1994; Kobilinsky, 1997) and the procedure allows one to quickly find the main features of the data.

To avoid fitting the response in too outlying points, the points with pure compounds were added only if the concentration of this compound was of the same order as in the designs.

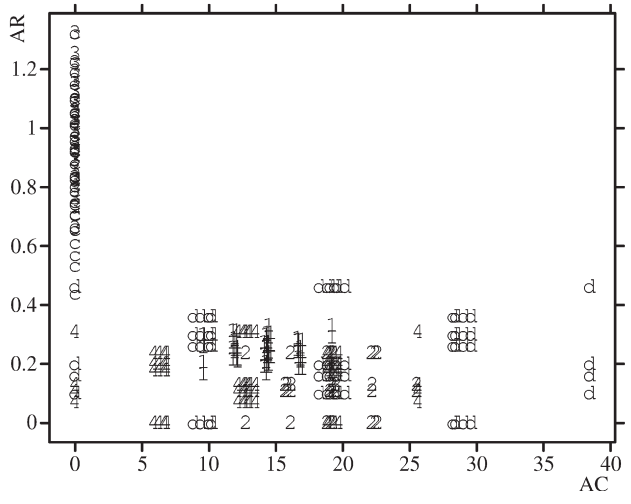


Fig. 7. Experimental points in the graph of AC versus AR. The symbol used for the representation indicates the number of the design: 1, 2, 3 and c for designs 3.1 and 3.2, 4 and d for designs 4.1 and 4.2.

Moreover, the points leading to no growth at 12 h (which generally correspond to values of AC greater 34, or to combination of values $AC \geq 28$ and $AR \geq 0.3$) were omitted too. Finally 273 points remained for the adjustment.

In the symbolic form admitted by the program, the model initially fitted on this data was:

$$AC^4 \cdot AR^3 + AS^2 + AS \cdot (AC + AR) + P + Q \cdot Q \qquad (2)$$

with

Q: acetic + citric + lactic + carvac + geran + thymol,
P: paba + asod + lacso + nitsod + shmp + stpp + cinnam + a-ter + menth.

It included all the monomials $AC^i\ AR^j$ of degree $i \leq 4$ and $j \leq 3$. Then AS, total of sodium and potassium acetate made after equivalencing these two compounds (AS = asod + 0.94 apot) and also its square AS^2 and the products $AS \cdot AC$, $AS \cdot AR$. Finally all the main effects appearing in P, Q as well as the two-factor interactions between the terms in Q (three organic acids and three aromatic compounds).

This initial model 2 has 60 parameters and was introduced as a tool to screen the factorial effects. Table 5 gives on its left the more significant terms in the analysis of variance with this model. The more important are lin AC, lin AR, quad AC, lin AS. But some salts shmp, lacso, paba, stpp and some complementary terms involving acids of aromatic compounds also have a non-negligible impact.

For stpp the estimated effect 3.2 on G is, as already noted, opposite to what is normally expected. Recall that, as explained in the end of Section 4.1, this effect has to be added if stpp is at its high level, subtracted otherwise. This salt stpp could have a tendency to limit the inhibition created by others. Terms involving isolated acid or aromatic compounds, like acetic, lactic, quad carvac, lactic · thymol also allow a better adjustment showing that though the order of magnitude of the growth percentage G essentially depends on them through the totals AC

and AR, the way these totals are obtained also has a non-negligible impact.

The residual standard deviation with the 60 parameters of model 2 is 4.778. The submodel keeping all polynomial terms significant at the 5% level and the subterms of them (for example acetic, thymol since it includes acetic · thymol) has 27 parameters and a slightly bigger residual standard deviation of 5.383 with 246 = 273 − 27 degrees of freedom. Table 5 gives it on the next to last column on its right. If only 1% significant terms are kept, the residual standard deviation increases to 6.993 though there are still 22 parameters in the model. If only 0.1% significant terms are kept, except paba which becomes non-significant in such a reduced model, the residual standard deviation increases to 9.587. The 9 parameters of this latter model are also given for information on the right of Table 5 but it is clear that keeping only them drastically increases the residual variance.

Appendix B gives the general form of orthonormal polynomials when the reference measure used to define orthogonality is symmetric. But here since the histograms of the different factors on the considered 273 points were not symmetric at all, it was decided to base the reference measures on them. Table 6 gives these reference measures and the corresponding orthonormal

Table 5
ANOVA of model 2 and submodels

	$F(1, 213)$	P (%)	S	Effect on G	+/− (95%)	+/− (99%)	+/− (99.9)	Model 95%	Model 99.9%
Res. std				4.778				5.383	9.587
Res. *d.f.*				213				246	264
Mean				45.5	4.6	6	7.7	47.98	47.82
AC	479.9	0.0	***	−29.4	2.6	3.5	4.5	−29.54	−23.59
AR	85.9	0.0	***	−11.2	2.4	3.1	4	−13.00	−6.42
AC^2	83.5	0.0	***	−10.5	2.3	3	3.8	−7.52	−7.57
AS	46.4	0.0	***	−9.3	2.7	3.5	4.5	−10.25	−6.20
lacso	32.1	0.0	***	−5.5	1.9	2.5	3.2	−5.08	−6.32
acetic	8.0	0.5	**	4.6	3.2	4.2	5.4	2.13	
shmp	36.3	0.0	***	−3.9	1.3	1.7	2.2	−4.47	−6.61
lactic	6.6	1.1	*	3.8	2.9	3.9	5	2.28	
paba	13.8	0.0	***	−3.6	1.9	2.5	3.2	−3.20	
$AC^2 \cdot AR$	8.2	0.5	**	3.5	2.4	3.2	4.1	0.93	
$carvac^2$	20.0	0.0	***	−3.3	1.4	1.9	2.4	−2.22	−3.88
stpp	10.7	0.1	**	3.2	1.9	2.5	3.3	3.60	
citric	4.3	3.8	*	−3	2.8	3.8	4.8	−2.03	
carvac	3.6	5.8		−3	3.1	4.1	5.2	−0.57	−1.15
$AC^2 \cdot AR^2$	4.4	3.7	*	2.4	2.3	3	3.8	1.21	
AR · AS	9.0	0.3	**	−2	1.3	1.7	2.2	−2.79	
AC · AR	2.8	9.8		−1.9	2.3	3	3.8	−3.24	
$AC \cdot AR^2$	5.0	2.7	*	1.9	1.6	2.2	2.8	0.76	
AC · AS	4.0	4.6	*	−1.9	1.8	2.4	3.1	−2.17	
AR^2	4.2	4.3	*	−1.8	1.8	2.3	3	−2.47	
$AC \cdot AR^3$	10.0	0.2	**	1	0.6	0.8	1	0.6	
lactic · thymol	9.9	0.2	**	0.9	0.6	0.7	0.9	0.89	
AC^3	4.6	3.4	*	0.8	0.7	1	1.3	2.04	
thymol	0.2	62.9		0.8	3.2	4.3	5.5	1.36	
lactic · carvac	7.3	0.8	**	0.7	0.5	0.7	0.9	0.62	
AR^3	0.1	79.9		−0.1	0.7	1	1.2	0.14	
$acetic^2$	4.9	2.8	*	1.4	1.2	1.6	2.1		
a-ter	6.0	1.5	*	−2.4	1.9	2.5	3.2		
AC^4	1.8	17.6		0.2	0.2	0.3	0.4		
...									

The points of design 3 for aromatic compounds such that AR > 0.5 were excluded from this analysis. The points such that either AC > 34, or AC ≥ 28 and AR ≥ 0.3 were also excluded because they lead to no growth at 12 h. Points with pure compounds were added if the compound concentration was of the same order as in the designs. Finally 273 points remained for the adjustment.

We use AC, AC^2,... to denote the linear, quadratic,... polynomial effects. Thus $AC^2 \cdot AR$ is quad AC lin AR. See Appendix B for the precise definition of such polynomial effects.

The 27 first terms are those significant at the 5% level plus those included in them or of smaller degree, for instance AR^3, thymol. The terms $acetic^2$ and a-ter become non-significant if the other last non-significant terms are pooled within the error and they were not included in the model with 5% significant terms.

$F(n_1, n_2)$: Fischer–Snedecor statistic with n_1, n_2 degrees of freedom.

P (%): probability to exceed the F value, when there is no effect.

S: associated significance (*: <5%, **: <1%, ***: <0.1%).

G: estimate of the effect on G, the growth percentage.

+/−, 95% (99, 99.9): quantity to add or subtract to get the limit of the 95% (resp. 99%, 99.9%) confidence interval. For instance at the 99% level quad AC^2 AR = 3.5 ± 3.2, that is the interval [0.3, 6.7] contains this polynomial effect with probability 99%.

Model 95%: coefficients of the orthonormal polynomials in the model obtained by selecting terms significant at the 5% level.

polynomials. Using them and their coefficient in Table 5, it is easy to get the corresponding prediction of the growth percentage. For instance, the predicted growth percentage with the 9 parameters model given on the right of Table 5 would be:

$$\begin{aligned}\hat{G} = {} & 47.82 - 23.59[0.12085(\mathrm{AC}-16.833)] \\ & -6.42[7.7339(\mathrm{AR}-0.2125)] \\ & -7.57[-0.84895 + 0.01194(\mathrm{AC}-16.833) \\ & + 0.01240(\mathrm{AC}-16.833)^2] - 6.20[0.0866(\mathrm{AS}-10)] \\ & -6.32[0.05497(\mathrm{lacso}-22.28)] - 6.61[0.29345(\mathrm{shmp}-5.3986)] \\ & -1.15 \times 27.574(\mathrm{carvac}-0.030625) \\ & -3.88[-1.3536 - 28.0860(\mathrm{carvac}-0.030625) \\ & + 1029.1570(\mathrm{carvac}-0.030625)^2].\end{aligned}$$

Such polynomial approximations have however the drawback to lead to abnormal extrapolations. To avoid these abnormalities, a non-linear model substituting a product $\mathcal{S}(\ln(AC))\mathcal{S}(\ln(AR))$ between two Schnute functions to the monomials in AC and AR and introducing only the major other terms with salts was also fitted. The maxima of the Schnute functions were taken as equal to 100% and another parameter was imposed upon the function of AR so that it decreases to 0 when AR = 1. This non-linear model leads to a residual standard deviation of 8 with 11 parameters.

The adjusted polynomial function f depends on too many factors to be represented in a two-dimensional graph. But as AC and AR are the more important of them and very weakly interact with the other factors, it is useful to represent G as a function of AC and AR for selected fixed values of other factors in the polynomial function. Fig. 8 gives some two-dimensional such representations. To add experimental points to such a graph, it is necessary to correct the observed growth percentage G. Let x_0 be the selected fixed vector of values for the factors other than AC and AR, and let x be the same vector of values for the experimental point considered. Then G is given by:

$$G = f(\mathrm{AC}, \mathrm{AR}, x) + \varepsilon.$$

The corrected value

$$\mathrm{G_c} = f(\mathrm{AC}, \mathrm{AR}, x_0) + \varepsilon = G - f(\mathrm{AC}, \mathrm{AR}, x) + f(\mathrm{AC}, \mathrm{AR}, x_0) \quad (3)$$

is obtained by adding to the observed growth percentage G the quantity

$$f(\mathrm{AC}, \mathrm{AR}, x_0) - f(AC, AR, x),$$

which is immediately deduced from the estimated function f. The points with these corrected G values are represented on Fig. 8.

The correction is made so as to bring back the observed values to those which would have been obtained with the conditions indicated below the graph (AS = 0, paba = 0,..., thymol = 0.07 mM). Graphs 2 and 3 show the variation of G as a function of AC for AR lying in a narrow range of values. Since AR was computed after the designs were achieved, it is not possible to keep it strictly constant while varying AC, but the restricted range of values selected on each graph insures that the variability around the represented curve is not too inflated. The same is true in graphs 5 and 6 where the abscissa is AR and where AC is restricted to a narrow range of values.

5.3.1. About the synergy between compounds

Synergy is generally reported when the combination of two components is more effective than each compound alone or when the observed inhibition of the combination is higher than

Table 6
Orthonormal polynomials and reference measures used to define them

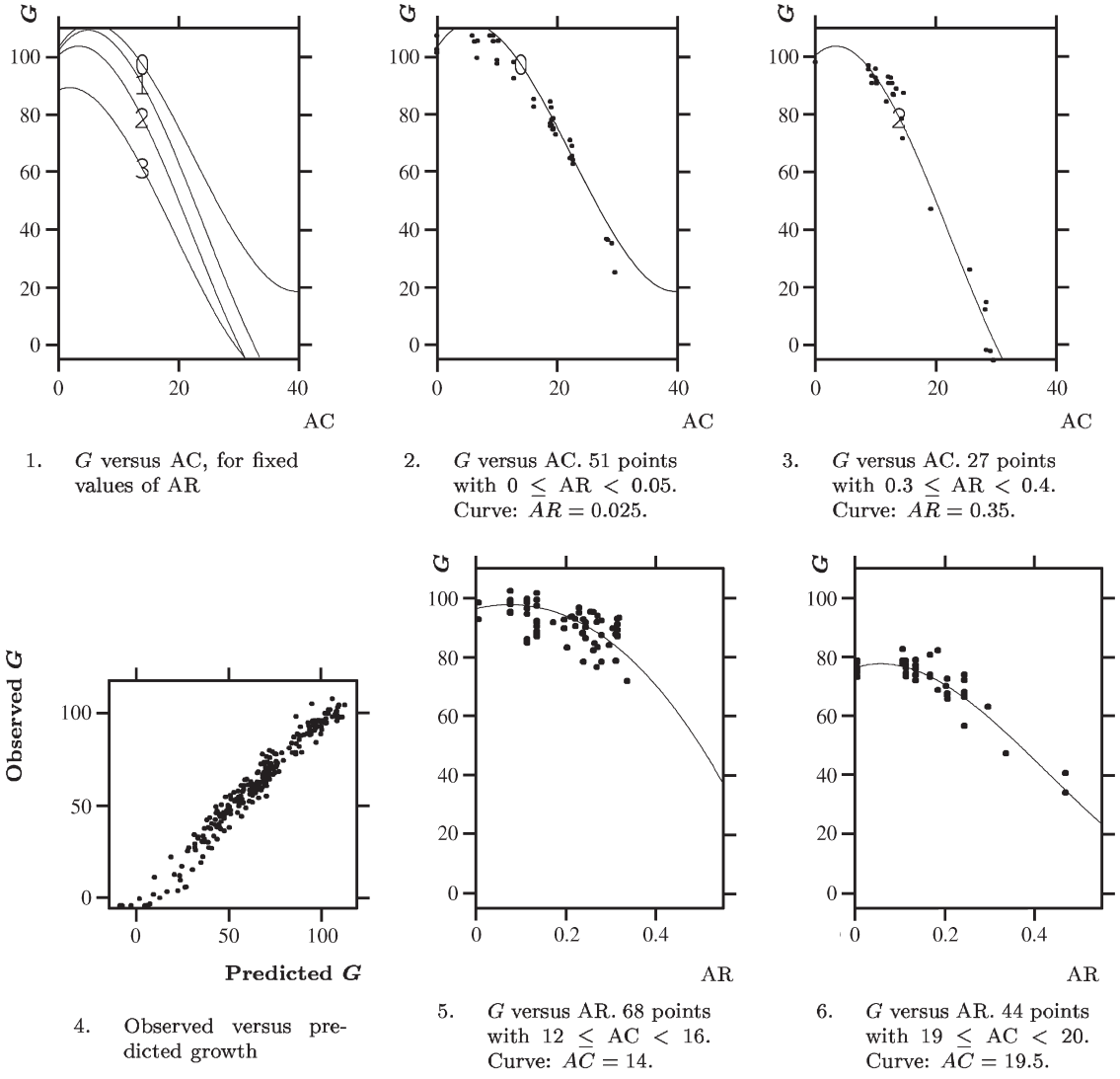
Orthonormal polynomials

lin acetic: 0.2767 (acetic − 4.2857)
lin citric: 0.6287 (citric − 2.2917)
lin lactic: 0.20949 (lactic − 6.875)
lin paba: 4.899 (paba − 0.25)
lin lacso: 0.05497 (lacso − 22.28)
lin shmp: 0.29345 (shmp − 5.3986)
lin stpp: 0.90055 (stpp − 1.36)
lin carvac: 27.574 (carvac − 0.030625)
lin thymol: 14.643 (thymol − 0.096)
lin AC: 0.12085 (AC − 16.833)
lin AR: 7.7339 (AR − 0.2125)
lin AS: 0.0866 (AS − 10)
quad carvac: − 1.3536 − 28.0860 (carvac − 0.030625) + 1029.1570 (carvac − 0.030625)^2
quad AC: − 0.84895 + 0.01194 (AC − 16.833) + 0.0124 (AC − 16.833)^2
cub AC: 0.07190 − 0.38368 (AC − 16.833) + 0.00120 (AC − 16.833)^2 + 0.00234 (AC − 16.833)^3
quad AR: − 1.2082 − 2.2795 (AR − 0.2125) + 72.2682 (AR − 0.2125)^2
cub AR: − 0.08076 − 30.87453 (AR − 0.2125) − 29.56758 (AR − 0.2125)^2 + 1090.54897 (AR − 0.2125)^3

Reference measures

AC	3	12	16	19	21	30		
AR	0.05	0.15	0.25	0.4				
AS	0	0	0	10	20	30		
acetic	0	0	2	4	6	8	10	
carvac	0	0	0	0	0.25	0.05	0.07	0.1

Supports of the discrete measures used to define the orthogonal polynomials. The same weight is given to each value, but some values are repeated to increase their weight (for example 0 for carvac). For the other factors, the support was selected as the list of taken values. It is easy to check that the vectors of values taken from one of these lists by the corresponding polynomials (1, lin, quad, cub) are orthonormal.



1. G versus AC, for fixed values of AR

2. G versus AC. 51 points with $0 \leq AR < 0.05$. Curve: $\overline{AR} = 0.025$.

3. G versus AC. 27 points with $0.3 \leq AR < 0.4$. Curve: $\overline{AR} = 0.35$.

4. Observed versus predicted growth

5. G versus AR. 68 points with $12 \leq AC < 16$. Curve: $\overline{AC} = 14$.

6. G versus AR. 44 points with $19 \leq AC < 20$. Curve: $\overline{AC} = 19.5$.

Fig. 8. Some graphs illustrating the global adjustment of the growth percentage G. Curves. 0: AR=0.025, 1: AR=0.25, 2: AR=0.35, 3: AR=0.45. The points in graphs 2, 3, 5, 6 are represented with the corrected growth percentage G_c defined by Eq. (3). The aim of this correction is to bring all data to the same conditions, except for AC and AR. The common conditions are defined by: AS=0, paba=0, lacso=0, shmp=0, stpp=0, acetic=3.5 mM, citric=2 mM, lactic=5 mM, carvac=0.025 mM, thymol=0.07 mM.

the one predicted by adding the inhibition created by the different compounds alone (Lachowicz et al., 1998; Tassou et al., 1995; Periago et al., 2002). These definitions are generally based on the hypothesis that the inhibitory effect of each compound varies linearly with concentration when its acts alone.

As explained in Section 4.2, this latter hypothesis is wrong here and this clearly explains why combining several organic acids, or several aromatic compounds, leads to an apparent synergy. But this is not a real synergy if the reference is the action of each isolated inhibitor.

Appendix A. Normalized factorial effects

A good way to make all factorial effects comparable is to normalize them. We explain below how this is done in a 2^2 factorial design studying two factors A, B at two levels numbered -1 and 1 each. We denote the levels by the same letters and let $\tau(A, B)$ be the expected mean of the observation for the treatment A, B. The general mean denoted by $e(1)$, the main effects of factor A, B denoted by $e(A)$, $e(B)$ and the interaction denoted by $e(AB)$ are then defined by:

$$\begin{aligned} e(1) &= [\tau(1,1) + \tau(1,-1) + \tau(-1,1) + \tau(-1,-1)]/4 \\ e(A) &= [\tau(1,1) + \tau(1,-1) - \tau(-1,1) - \tau(-1,-1)]/4 \\ e(B) &= [\tau(1,1) - \tau(1,-1) + \tau(-1,1) - \tau(-1,-1)]/4 \\ e(AB) &= [\tau(1,1) - \tau(1,-1) - \tau(-1,1) + \tau(-1,-1)]/4. \end{aligned} \tag{4}$$

Inverting these relations (4) gives

$$\begin{aligned} \tau(1,1) &= e(1) + e(A) + e(B) + e(AB) \\ \tau(1,-1) &= e(1) + e(A) - e(B) - e(AB) \\ \tau(-1,1) &= e(1) - e(A) + e(B) - e(AB) \\ \tau(-1,-1) &= e(1) - e(A) - e(B) + e(AB). \end{aligned} \tag{5}$$

So with this definition (4) the expected mean is got from the general mean $e(1)$ by adding or subtracting the main effects $e(A)$ and $e(B)$and their interaction $e(AB)$ according to the levels A, B and their product AB. The preceding equalities can be summed up by the following equality

$$\tau(A,B) = e(1) + Ae(A) + Be(B) + ABe(AB).$$

With this definition a main effect such as $e(A)$ is the half difference between the means at the level 1 and the level −1. This differs from the usually adopted definition where the main effect is twice the value above, that is the full difference between these two means. Similarly the division by 4 is often omitted in the definition of the interaction which is then 4 times the value above. That is the general mean and factorial effects are often defined by

$$\begin{aligned} e(1) &= [\tau(1,1) + \tau(1,-1) + \tau(-1,1) + \tau(-1,-1)]/4 \\ e'(A) &= [\tau(1,1) + \tau(1,-1) - \tau(-1,1) - \tau(-1,-1)]/2 \\ e'(B) &= [\tau(1,1) - \tau(1,-1) + \tau(-1,1) - \tau(-1,-1)]/2 \\ e'(AB) &= [\tau(1,1) - \tau(1,-1) - \tau(-1,1) + \tau(-1,-1)] \end{aligned} \quad (6)$$

which is inverted in

$$\begin{aligned} \tau(1,1) &= e(1) + e'(A)/2 + e'(B)/2 + e'(AB)/4 \\ \tau(1,-1) &= e(1) + e'(A)/2 - e'(B)/2 - e'(AB)/4 \\ \tau(-1,1) &= e(1) - e'(A)/2 + e'(B)/2 - e'(AB)/4 \\ \tau(-1,-1) &= e(1) - e'(A)/2 - e'(B)/2 + e'(AB)/4. \end{aligned} \quad (7)$$

The dispersion of the expected means τ (A, B) can be measured by the variance:

$$\mathrm{var}(\tau) = \sum_{A,B} (\tau(A,B) - e(1))^2/4$$

which with the factorial effects defined in Eq. (4) takes the form

$$\mathrm{var}(\tau) = e(A)^2 + e(B)^2 + e(AB)^2.$$

Thus all these factorial effects contribute in the same way to this dispersion. They can therefore be directly compared to each other, which is not the case with the alternative definition (6) which leads to

$$\mathrm{var}(\tau) = \frac{1}{4}\left(e'(A)^2 + e'(B)^2 + \frac{1}{4}e'(AB)^2\right)$$

and in which, if $e'(AB) = e'(A)$, the contribution of $e'(AB)$ is much smaller than that of $e'(A)$.

This normalization is similar to what is often done in linear regression. The explicative variables are reduced so as to have the same standard deviation, which makes the regression coefficients comparable. In Eq. (5), the vectors of coefficients (1 1 −1 −1) of $e(A)$, (1 −1 1 −1) of $e(B)$, (1 −1 −1 1) of $e(AB)$ have the same standard deviation 1 while in Eq. (7) their standard deviations are respectively 1/2, 1/2 and 1/4. That is if $e'(AB) = e'(A)$, the variability induced by $e'(AB)$ on the response is in standard deviation twice smaller. So $e(AB)$ and $e(A)$ can be directly compared which is not the case of $e'(AB)$ and $e'(A)$.

Appendix B. Polynomial effects

When fitting polynomials, it is highly recommended to introduce orthogonal polynomials instead of monomials. We explain why in the case of one explicative quantitative variable x, then tell how to proceed in the case of several explicative quantitative variables.

Assume the observed response y has expectation $f(x)$, that is, it follows the model

$$y = f(x) + \varepsilon,\ E(\varepsilon) = 0 \quad (8)$$

and that we approximate $f(x)$ by a polynomial $P(x)$ of degree 3 in x:

$$f(x) \approx P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3. \quad (9)$$

The successive derivatives of P are $P'(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2$, $P^{(2)}(x) = 2\alpha_2 + 6\alpha_3 x$, $P^{(3)}(x) = 6\alpha_3$. So $\alpha_0 = P(0)$, $\alpha_1 = P'(0)$, $\alpha_2 = P^{(2)}(0)/2$ and finally $\alpha_3 = P^{(3)}(x)/6$ for any value of x.

While α_3 gives therefore a global indication on the behaviour of P, the three first parameters α_0, α_1, α_2 cannot be given any meaning if the origin is selected arbitrarily as they depend from the behaviour of P at a value of 0 which can be far away from the experimented values of x. Moreover their estimates may be strongly correlated and very imprecise. Assume for instance the values of interest for x are in the interval [30, 33] and that the experimented design is equireplicated on each of the four equispaced values 30, 31, 32, 33. Then the correlations between the estimates $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$ are greater than 0.999 and a very slight variation of the observations can generate a very big change in them.

Things become better if x is centered. Let m be a mean value of x, for instance $m = 31.5$ in the example. The expression of $P(x)$ as a function of $x - m$ is

$$P(x) = \alpha_0' + \alpha_1'(x-m) + \alpha_2'(x-m)^2 + \alpha_3'(x-m)^3 \quad (10)$$

$$\text{where } \begin{cases} \alpha_0' = \alpha_0 + \alpha_1 m + \alpha_2 m^2 + \alpha_3 m^3, \\ \alpha_1' = \alpha_1 + 2\alpha_2 m + 3\alpha_3 m^2, \\ \alpha_2' = \alpha_2 + 3\alpha_3 m, \\ \alpha_3' = \alpha_3. \end{cases}$$

The coefficient $\alpha_2' = \alpha_2 + 3\alpha_3 m$ of the term $(x-m)^2$ is one half of the mean value of the second order derivative $P^{(2)}(x) = 2\alpha_2 + 6\alpha_3 x$, hence it gives, as $\alpha_3' = \alpha_3$ as a global indication on the behaviour of P, that is the mean curvature.

But α_0', α_1' are the value of $P(x)$ and $P'(x)$ at the particular point m. They are not necessarily representative of the global behaviour of P. It is more meaningful to consider instead the constant β_0 best approximating $P(x)$ and β_1 slope of the best fitting straight line. To get homogeneous notations, we then let $\beta_2 = \alpha_2'$, $\beta_3 = \alpha_3$.

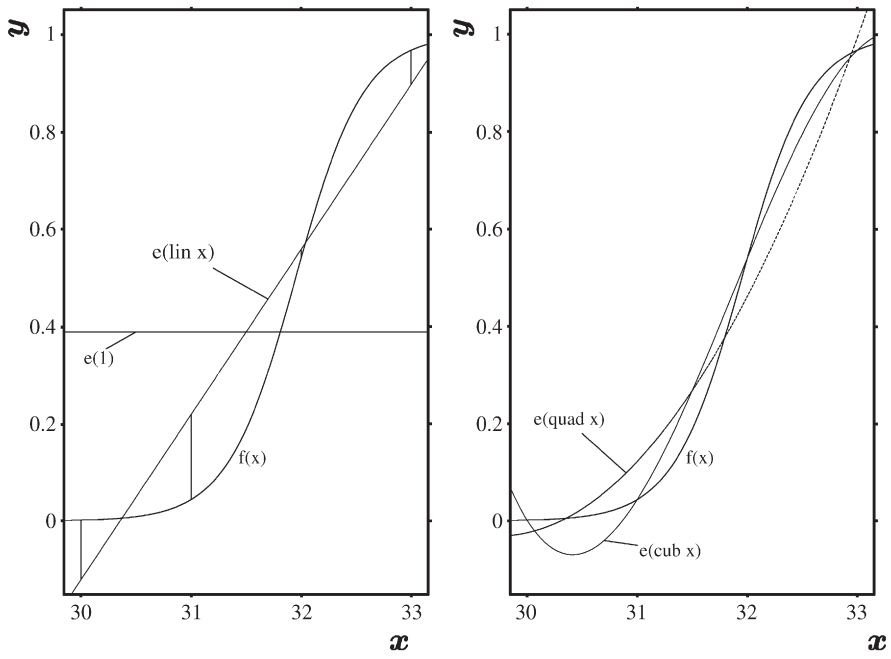


Fig. 9. Definition of polynomial effects when the fitting quality is measured by Eq. (11). Up to normalization, each one is the coefficient of the higher degree of the best fitting curve to each it is connected. The best fitting constant (horizontal line) and straight line are on the left, the best fitting quadric and cubic curve on the right. So the mean effect $e(1)$ is the ordinate of the horizontal line, $e(\text{lin } x)$ the slope of the straight line, $e(\text{quad } x)$ the curvature of the quadratic curve, and finally $e(\text{cub } x)$ the coefficient of the term in x^3 in the cubic curve. The quality of the fit is in that case measured by Eq. (11), the sum of the squares of the vertical distances at the four abscissa 30, 31, 32, 33 which for the straight line is the length of the reported vertical segments.

To define the best fitting constant, or straight line, or polynomial of any order, a measure of the quality of an approximation is needed. The measure used here is the $L^2(\mu)$ distance associated with the scalar product $\langle Q, P\rangle = \int QP\mathrm{d}\mu$. It takes the form

$$\mathrm{d}(Q,P) = \frac{1}{4}\Big([Q(30)-P(30)^2] + [Q(31)-P(31)]^2 + [Q(32)-P(32)]^2 + [Q(33)-P(33)]\Big)^2 \quad (11)$$

if the measure μ gives a weight of 1/4 to each of the four points 30, 31, 32, 33 studied, that is if the associated scalar product is

$$\langle P,Q\rangle = \frac{1}{4}[P(30)Q(30) + P(31)Q(31) + P(32)Q(32) + P(33)Q(33)]. \quad (12)$$

It takes the form

$$\mathrm{d}(P,Q) = \frac{1}{3}\int_{30}^{33}[P(x)-Q(x)]^2\mathrm{d}x \quad (13)$$

if μ is the uniform measure of density 1/3 on the interval [30, 33] and the associated scalar product therefore

$$\langle P,Q\rangle = \frac{1}{3}\int_{30}^{33}P(x)Q(x)\mathrm{d}x. \quad (14)$$

Note that β_2 can be shown to be the coefficient of the term of degree 2 in the parabol best fitting $P(x)$, so that $\beta_0, \beta_1, \beta_2, \beta_3$ are the coefficients of the terms of highest degree in the polynomials $P_0(x)$, $P_1(x)$, $P_2(x)$, $P_3(x)$ of respective degrees 0, 1, 2, 3 best fitting $P(x)$ (hence $f(x)$ too if $P(x)$ is the best approximation of f with the same distance).

Fig. 9 illustrates these best fitting constant, straight line, parabol and cubic giving the polynomial effects up to the normalization coefficient, in the case of distance (11).

These best approximations of P by polynomials of degrees 0, 1, 2, 3 are easily taken from the associated orthogonal polynomials O_0, O_1, O_2, O_3 of degrees 0, 1, 2, 3 in x. The parameters β_0, β_1, β_2, β_3 are those appearing when $P(x)$ is expressed as a function of these polynomials

$$P(x) = \beta_0 O_0(x) + \beta_1 O_1(x) + \beta_2 O_2(x) + \beta_3 O_3(x).$$

They are

$$\beta_0 = \langle P,O_0\rangle/\langle O_0,O_0\rangle,\ \beta_1 = \langle P,O_1\rangle/\langle O_1,O_1\rangle,$$
$$\beta_2 = \langle P,O_2\rangle/\langle O_2,O_2\rangle,\ \beta_3 = \langle P,O_3\rangle/\langle O_3,O_3\rangle.$$

In this little example the reference measure is symmetric around the central point $m=31.5$. If m_k denotes the central moment of order k, the orthogonal polynomials take then the following form

$$O_0(x) = 1,\ O_1(x) = x-m,\ O_2(x) = (x-m)^2 - \frac{m_2}{m_0},$$
$$O_3(x) = (x-m)^3 - \frac{m_4}{m_2}(x-m). \quad (15)$$

With the measure μ giving a weight 1/4 to each of the four points 30, 31, 32, 33, the central moment m_k is

$$m_k = \frac{1}{4}\sum_x (x-m)^k = [(30-m)^k + (31-m)^k + (32-m)^k + (33-m)^k]/4 \quad (16)$$

hence $m_0=1$, $m_2=1.25$, $m_4=2.5625$, $m_6=5.703125$ and so

$$O_0(x)=1,\ O_1(x)=x-31.5,\ O_2(x)=(x-31.5)^2-1.25,$$
$$O_3(x)=(x-31.5)^3-2.05(x-31.5).$$

In the analysis of Section 5.3, the reference measures are chosen to follow more or less the empirical distributions obtained from the 273 points. As a consequence, they are not symmetric and the orthogonal polynomials are found numerically. The interest of such a choice is to lead to effects with more precise estimates. Table 6 gives these reference measure and associated orthonormal polynomials for 5 factors. For the other factors, the reference measure was selected as the one giving the same weight to each of the values taken by the factor. For instance, since citric takes the values 0, 1.25, 1.67, 2.5, 3.33, 5, it is the measure giving a weight of 1 to each of these 6 values which was selected as a reference measure.

B.1. More on the reparameterisation through orthogonal polynomials

The new parameters β_0, β_1, β_2, β_3 have a clear definition. Moreover they have uncorrelated estimates at least in the case where the actual design coincides with the reference measure, for instance in the example just considered where the distance is defined by Eq. (11) if each of the points $x=30$, 31, 32, 33 is equireplicated.

To allow a direct comparison between the parameters as in Appendix A, it is judicious to normalize these orthogonal polynomials. The *orthogonal* polynomials O_0, O_1, O_2, O_3 are then replaced by the *orthonormal* polynomials

$$R_0=O_0/||O_0||,\ R_1=O_1/||O_1||,\ R_2=O_2/||O_2||,$$
$$R_3=O_3/||O_3||.$$

If the reference measure is symmetric as in the example it is easy to check that the denominators are

$$||O_0||=1,\ ||O_1||=\sqrt{m_2},\ ||O_2||=\sqrt{m_4-\frac{m_2m_2}{m_0}},$$
$$||O_3||=\sqrt{m_6-\frac{m_4m_4}{m_2}}$$

and so with the measure μ giving the same weight of 1/4 to each of the four points 30, 31, 32, 33, $||O_0||=1$, $||O_1||=\sqrt{1.25}$, $||O_2||=1$, $||O_3||=\sqrt{0.45}$. The expression as a function of these orthonormal polynomials is

$$P(x)=\gamma_0R_0(x)+\gamma_1R_1(x)+\gamma_2R_2(x)+\gamma_3R_3(x)$$

and the *normalized* parameters are obtained from P by the equalities

$$\gamma_0=\langle P,R_0\rangle,\ \gamma_1=\langle P,R_1\rangle,\ \gamma_2=\langle P,R_2\rangle,\ \gamma_3=\langle P,R_3\rangle.$$

The parameter $\gamma_0=||O_0||\beta_0$ is called the *general mean* and $\gamma_1=||O_1||\beta_1$, $\gamma_2=||O_2||\beta_2$, $\gamma_3=||O_3||\beta_3$, are respectively called the *linear*, *quadratic* and *cubic* effects of x. We use the abbreviations lin x, quad x, cub x for the last three ones. So up to the normalization coefficients, lin x is the slope of the better approximation of $P(x)$ by a straight line, quad x the coefficient of the term of degree 2 (curvature) in the best approximation by a quadratic curve, cub x the coefficient of the term of degree 3 in the best approximation by a cubic curve.

In many cases, it is more convenient to use lin x, quad x, cub x to denote the orthonormal polynomials $R_1(x)$, $R_2(x)$, $R_3(x)$ and then e(lin x), e(quad x), e(cub x) to denote the corresponding effects (Cliquet et al., 1994). This last notation is more in agreement with the one used for factorial effects in Appendix A and we use it in Fig. 9.

The extension to polynomial in several factors x_1, x_2,... is easy as the orthogonal (resp. orthonormal) polynomials are the products of those in one factor. For instance the orthonormal polynomial of degree 1 in x_1, 2 in x_2 is the product $R_1(x_1)R_2(x_2)$. It is known as the tensor product $R_1 \otimes R_2$. The coefficient of this polynomial is denoted lin x_1 quad x_2, or e(lin x_1 quad x_2) in the alternative notation. The orthogonality of such polynomials is with respect to the product measures, hence the estimated parameters are uncorrelated at least in the reference design associated with this product measure.

For instance, with two factors x_1 and x_2 and a design equireplicating each of the 8 points of the grid defined by $x_1=30$, 31, 32, 33, $x_2=6$, 7, the orthogonal polynomials of degree less than 2 in x_1, 1 in x_2 are

$$O_{10}\otimes O_{20},\ O_{11}\otimes O_{20},\ O_{12}\otimes O_{20},\ O_{10}\otimes O_{21},$$
$$O_{11}\otimes O_{21},\ O_{12}\otimes O_{21},$$

where O_{10}, O_{11}, O_{12} are the orthogonal polynomials in x_1 defined as O_0, O_1, O_2 in Eq. (15), distance 11, by

$$O_{10}(x)=1,\ O_{11}(x)=x-31.5,\ O_{12}(x)=(x-31.5)^2-1.25$$

and O_{20}, O_{21} the orthogonal polynomials in x_2 defined in that case by

$$O_{20}(x)=1,\ O_{21}(x)=x-6.5.$$

The orthonormal polynomials are defined in the same way

$$R_{10}\otimes R_{20},\ R_{11}\otimes R_{20},\ R_{12}\otimes R_{20},\ R_{10}\otimes R_{21},\ R_{11}\otimes R_{21},$$
$$R_{12}\otimes R_{21}.$$

Note that the orthonormal polynomials and hence the polynomial effects are defined through a reference measure,

which in the case of several factors is the product between the measures associated with each factor. It follows that the orthogonal polynomial in several factors is the product of those associated with each factor and that one does not have to worry about the order of introduction of the factors in defining these orthogonal polynomials. Moreover, these polynomials are not dependant on the particular set of data processed and thus allow comparison between different sets of data.

References

Ankenman, B.E., 1999. Design of experiments with two- and four-level factors. Journal of Quality Technology 31 (4), 363–375.

Cliquet, S., Durier, C., Kobilinsky, A., 1994. Principle of a fractional factorial design for qualitative and quantitative factors: application to the production of *Bradyrhizobium japonicum* in culture media. Agronomie 14, 569–587.

Diamond, W.J., 1981. Practical Experiment Design for Engineers and Scientists. Lifetime Learning Publications, Belmont.

Draper, N.R., Mitchell, T.J., 1967. The construction of saturated 2^{k-p}_R designs. The Annals of Mathematical Statistics 38, 1110–1126.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and its Applications. Chapman & Hall.

Fisher, R.A., Yates, F., 1957. Statistical Tables for Biological, Agricultural and Medical Research. Oliver and Boyd, Edinburgh, UK.

Goos, P., Kobilinsky, A., O' Brien, T.E., Vandebroek, M., 2005. Model-robust and model-sensitive designs. Computational Statistics and Data Analysis 49 (1), 201–216.

Gould, G.W., Abee, T., Granum, P.E., Jones, M.V., 1995. Physiology of food poisoning microorganisms and the major problems in food poisoning control. International Journal of Food Microbiology 121–128.

Gould, G.W., 1996. Industry perspectives on the use of natural antimicrobials and inhibitors for food applications. Journal of Food Protection. Supplement 82–86.

Kobilinsky, A., 1988. Tactiques en analyse de variance et en régression. In La revue de MODULAD 1, p 25–58. Ed INRIA.

Kobilinsky, A., 1994. PLANOR: program for the automatic generation of regular experimental designs. Version 2.2 for Windows, 1999, Biometrics Unit (MIA). INRA, Jouy-en-Josas.

Kobilinsky, A., 1997. Les plans factoriels. In: Droesbeke, J.J., Fine, J., Saporta, G. (Eds.), Plans d'expériences: applications à l'entreprise. Technip, Paris, pp. 69–209.

Kobilinsky, A., Monod, H., 1995. Juxtaposition of regular factorial designs and the complex linear model. Scandinavian Journal of Statistics 22 (2), 223–254.

Lachowicz, K.J., Jones, G.P., Briggs, D.R., Bienvenu, F.E., Wan, J., Wilcock, A., Coventry, M.J., 1998. The synergistic preservative effects of the essential oils of sweet basil (*Ocimum basilicum* L.) against acid-tolerant food microflora. Letters in Applied Microbiology 26, 209–214.

Lambert, R.J.W., Pearson, J., 2000. Susceptibility testing: accurate and reproducible minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. Journal of Applied Microbiology 88, 784–790.

Moll, M., Moll, N., 2002. Salmonelles et salmonelloses en France. Sécurité alimentaire du consommateur. Collection Sciences et Techniques Agroalimentaires, pp. 1–33.

Montgomery, D.C., 1997. Design and Analysis of Experiments, 4th ed. John Wiley & Sons, New-York, NY.

Nazer, A.I., Kobilinsky, A., Tholozan, J.L., Dubois-Brissonnet, F., 2005. Combinations of food additives at low levels to inhibit the growth of *Salmonella* sv. Typhimurium: a synergistic effect? Food Microbiology 22, 391–398.

Pearson, E.S., Hartley, H.O., 1976. Biometrika Tables for Statisticians. Charles Griffin and Company Ltd for the Biometrika Trust, London, UK.

Periago, P.M., Palop, A., Fernandez, P.S., 2002. Combined effect of nisin, carvacrol and thymol on the viability of *Bacillus cereus* heat-treated vegetative cells. International Journal of Food Science and Technology 7, 487–492.

Plackett, R.L., Burman, J.P., 1946. The design of multifactorial experiments. Biometrika 33, 305–325.

Ray, B., 2001. Introduction. In: C.P. LLC (Ed.), Fundamental Food Microbiology, pp. 336–343.

Tassou, C.C., Drosinos, E.H., Nychas, G.J.E., 1995. Effects of essential oil from mint (*Mentha piperita*) on *Salmonella enteritidis* and *Listeria monocytogenes* in model food systems at 4 °C and 10 °C. Journal of Applied Bacteriology 78, 593–600.

Zwietering, M.H., Jongenburger, I., Rombouts, F.M., van't Riet, K., 1990. Modeling of the bacterial growth curve. Applied and Environmental Microbiology 56 (6), 1875–1881.