# Modeling properties of water-based paints dried by microwaves from the design and analysis of a mixture-process experiment

S. Barra[1,2], F. Boulanger[1], A. Trapani[1], and D. Cabrol-Bass[2]

[1]*Rohm & Haas European Laboratories, 06560 Valbonne, France*
[2]*Laboratoire Arômes, Synthèses, Interactions, Équipe Chimiométrie et Modélisation, Université de Nice-Sophia Antipolis, 06108 Nice Cedex 2, France*

Heightened environmental awareness has lead major paint companies to develop environmentally friendly waterborne coatings. Recently, microwave radiation has been used to accelerate the drying of industrial waterborne coatings.

The aim of this study was to develop statistical models to optimise the performance of microwave radiation dried paints. This paper explores the effects and interactions between the major components of the paints and examines how these interactions affect the properties of the paints. A combined mixture with a double factorial design, including several types of components, has been used in this work. Eight paint properties and three microwave energy levels have been considered on each point of the experimental design.

A computer simulation consisting of 12 separate models has been developed that accurately predicts the performance of paint.

## Introduction

Because of numerous environmental and legislative requirements the paint industry is increasing its interest in the development and production of water-based paints. Using water-based paints allows for the reduction of the emission of polluting, organic solvents found in traditional paints and varnishes. However the amount of time needed for drying these new, environmentally-friendly paints can sometimes drive down productivity. Since water is very efficiently heated by microwave energy, it seems logical that a microwave oven should be very effective for drying a water-based paint and this topic has gained much interest in the recent past. Still, drying water-based paints using microwaves requires adapting formulations to this new technology and presents a challenge for paint formulators. This study concerns paints used in the professional, wood-finishing industry (furniture, doors, windows). This industry requires very fast surface finishing lines and water-based paints are still a minor part of the coatings used.

The objective is twofold. On one hand it is desired to understand how the major components of a water-based paint affect, alone or in combination, the major properties characterizing a good quality paint when dried by microwave energy. To achieve this, a mixture design is planned in which the quantity of each of the components varies in a domain constrained by feasibility requirements. In addition, it is wished to determine if these properties depend also upon the chemical types of binder or coalescents used in the formulation as well as upon the energy dosage received at the drying stage. It is thus necessary to combine the mixture design with a double factorial design including several types of binders and in which blends will be tested with different energy dosages.

On the other hand, it is desired to use this understanding to determine which blends simultaneously optimize the eight measured properties. To achieve this, simple mathematical models of the polynomial type will be derived from the collected data and used to build response surfaces over the whole experimental domain. If such blends do not exist, an algorithm for modelling desirability will be used to come up with acceptable compromises.

## Approach and design

The effect of various components of water-based coatings and others factors are studied using a statistical experimental design. Not only the components (and the mixture of the components) of the paints but also the process used for drying the paints affect the final properties of the dry paint film. Such experiments where the process variables are factors that are not present in the mixtures to be studied, but whose levels, when changed, could affect the results obtained from these mixture are known as mixture experiments with process variables.

This paper presents the result and the strategy of a mixture experiment with process variables consisting of six

components, combined with two others factors which are the chemical nature of two of the six mixture components. In particular the fact that each paint mixture (or formulation) can include two types of binders (the main polymeric portion of the formulation) and one coalescent alone or two coalescents simultaneous to be chosen in a list of four candidates must be take into account.

### Mixture components

The six mixture components are listed below:

$X_1$ = binder (acrylic dispersion polymer);
$X_2$ = coalescent 1;
$X_3$ = coalescent 2;
$X_4$ = water;
$X_5$ = polyurethane dispersion;
$X_6$ = salt.

It is known that the properties of the mixture are dependent only upon the respective percentage of active ingredients present. In this design, every component is expressed in terms of active ingredient weight. In this case, the only way to compare formulations is to fix the total content of active material. In this experiment the sum of the components is 60 g. Thus:

$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 = 60$ g of active material.

The experimental region for the mixture components is constrained by the imposition of lower and upper bounds on the component proportions that determine the feasibility domain.

Individual constraints:

$$28 \leq X_1 \leq 35$$
$$0 \leq X_2 \leq 8.75$$
$$0 \leq X_3 \leq 8.75$$
$$0 \leq X_4 \leq 20$$
$$0 \leq X_5 \leq 7$$
$$0 \leq X_6 \leq 0.53.$$

Relational constraints:

$3.5 \leq X_2 + X_3 \leq 8.75$. This relationship means that either coalescent 1 or 2 can be equal to 0, but that a coalescent is always present in the mixture.

$$34.5 \leq X_1 + X_5 \leq 35.5.$$

The resulting region of this mixture experiment is very constrained due to the limits imposed on each component and the relationships holding between them.

### Indicator variables

Two indicator variables, $Z_1$ and $Z_2$ are defined to represent the chemical nature of respectively, the acrylic binder and the coalescent used in different formulations of this design.

The values of $Z_1$ and $Z_2$ are defined as follows:

*Binder type:*

$Z_1 = 0$ coded for the first type of acrylic binder;

$Z_1 = 1$ coded for the second type of acrylic binder that can used in the blend.

One binder only can be used in each formulation.

*Coalescent pair type:*

Four different coalescents are of interest, coded by $A$, $B$, $C$ and $D$.

Each coalescent ($A$, $B$, $C$, $D$) can be present in the formulation either alone (in which case either $X_2 = 0$ or $X_3 = 0$) or as a pair with a different coalescent in the list (in which case both $X_2 \neq 0$ and $X_3 \neq 0$).

For example $Z_2 = 1$ indicates that coalescents $A$ or $B$ can be present alone (in this case $X_2$ or $X_3$ is equal to zero), or can be both present ($X_2 \neq 0$ and $X_3 \neq 0$).

The values of $Z_2$ are defined as:

**$Z_2 = 1$** $A$ and $B$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

**$Z_2 = 2$** $A$ and $C$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

**$Z_2 = 3$** $A$ and $D$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

**$Z_2 = 4$** $B$ and $C$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

**$Z_2 = 5$** $B$ and $D$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

**$Z_2 = 6$** $C$ and $D$ can be present alone ($X_2 = 0$ or $X_3 = 0$) or together ($X_2 \neq 0$ and $X_3 \neq 0$).

### Process variable

The power of the microwaves used to dry the blends takes three levels and as such is a process variable.

These levels are coded $P_i$, $i = 1$ to 3.

### Building the model

In mixture experiments containing others factors, the typical model-fitting strategy is to fit a combined model containing terms in the mixture components only along with terms involving crossproducts between the mixture components and the other factors. The model must incorporate the blending properties of the mixture components and the effects of the others factors. Such a model form allows to measure the blending properties of the mixture components and to determine if the blending properties change when changing the setting of the process variables.

Polynomial models of different degrees can describe the blending properties of the mixture components and the effect of the process variables on the responses of interest.

The adopted approach was to express the blending properties of the mixture components with a reduced Scheffé-type polynomial model [1] of degree two.

In the present case, with six components whose values are denoted by $X_1$, $X_2$, $X_3$, $X_4$, $X_5$, $X_6$ the response is expressed by:

$$\text{Response} = \sum_{i=1}^{6} b_i X_i + \sum_{i=1}^{6} \sum_{j>1}^{6} b_{ij} X_i X_j \, .$$

The first six terms are called the linear blending terms, the remaining terms are called the non linear blending terms [2].

In this problem, it is assumed that water acts as a diluent, and blends additively, thus a reduced quadratic model in the mixture components should provide a reasonable approximation of the response surface. The reduction from the full 21 term quadratic model results from the elimination of all binary blending terms with $X_4$ (Water), thus leaving 16 terms. The resulting reduced quadratic model in the mixture components is:

$$
\begin{aligned}
Y = \ & b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + b_5 X_5 + b_6 X_6 + \\
& b_{1,2} X_1 X_2 + b_{1,3} X_1 X_3 + b_{1,5} X_1 X_5 + b_{1,6} X_1 X_6 + \\
& b_{2,3} X_2 X_3 + b_{2,5} X_2 X_5 + b_{2,6} X_2 X_6 + \\
& b_{3,5} X_3 X_5 + b_{3,6} X_3 X_6 + \\
& b_{5,6} X_5 X_6 + \varepsilon
\end{aligned}
\tag{1}
$$

where the $b_{i,j}$ are unknown coefficients to be estimated and $\varepsilon$ is the random error term.

The indicator variable model is expressed as:

$$\text{Response} = \alpha_0 + \{\text{Effect of } Z\} + \varepsilon$$

where $\{\text{Effect of } Z\} = \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_{1,2} Z_1 Z_2$.

The combined model containing both the mixture and the indicator variables is formed by expanding the mixture model with terms involving the indicator variables. If, when changing the level of $Z_i$, only the value of the response changes and the blending properties are not affected, the combined model can be written as [3]:

$$\text{Response} = \{1\} + \{\text{Effect of } Z\} + \varepsilon.$$

However changing the level of $Z_i$ might affect the blending properties, which means that the $Z_i$ factors are interacting with the $X_i$ factors. The combined model is therefore expressed as [3]:

$$\text{Response} = \{1\} + \{1\}\{\text{Effect of } Z\} + \varepsilon$$

where $\{1\}\{\text{effect of } Z\}$ represents the effect of changing the level of $Z$ on the blending properties of the components.

To reduce the number of experimental runs, the interaction between $Z_1$ and $Z_2$, was not taken into consideration, so the term $\{1\}Z_1 Z_2$ has been eliminated from the model.

So the combined model is expressed as:

$$Y = \{1\} + \{1\}Z_1 + \{1\}Z_2 + \varepsilon. \tag{2}$$

To even further reduce the number of experimental runs, the interactions between $X_4$ (Water) and $Z_1$ and between $X_4$

and $Z_2$ were dropped from the initial model (1) leading to be the following 15 term model (3):

$$
\begin{aligned}
Y = \ & b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots b_5 X_5 + b_6 X_6 + \\
& b_{1,2} X_1 X_2 + b_{1,3} X_1 X_3 + b_{1,5} X_1 X_5 + b_{1,6} X_1 X_6 + \\
& b_{2,3} X_2 X_3 + b_{2,5} X_2 X_5 + b_{2,6} X_2 X_6 + \\
& b_{3,5} X_3 X_5 + b_{3,6} X_3 X_6 + \\
& b_{5,6} X_5 X_6 + \varepsilon.
\end{aligned}
\tag{3}
$$

Then the final combined model, which serves as a basis to design the experiments is:

$$Y = \{1\} + \{3\}Z_1 + \{3\}Z_2 + \varepsilon. \tag{4}$$

This model is linear for $Z_1$ and $Z_2$ and does not include the $X_4 Z_1$ and $X_4 Z_2$ terms. Model $\{1\}$ includes 16 terms to estimate, the term $\{3\}Z_1$ includes 15 terms, and the term $\{3\}Z_2$ includes $5 \times 15$ terms (after coding $Z_2$) to be estimated. This model has a total of 106 parameters to estimate.

The effect of P, the power level of microwaves used to dry a formulation, was also included in the model. It is believed that power might have a quadratic effect on many of the responses. So the model becomes:

$$Y = \{(1)\} + \{(3)\}Z_1 + \{(3)\}Z_2 + (a_1{}^1 P + a_1{}^2 P^2) + \varepsilon. \tag{5}$$

When fully crossed with model (4), the power variable triples the number of formulations required.

To reduce the amount of work, each formulation is split into 3 samples. The 3 levels of power are randomly assigned to the samples. Each sample is dried at the correspondent power of the same length of time (2').

Therefore, the assignment of the experimental units to the power levels is restricted, thus leading to a design of the split-plot type [4]. Two sources of experimental error are present in the data. The first is the main plot error that occurs among the mixture blends of the formulations. It allows to estimate the error on each component of the mixture $(X_i)$. The second source of error is the sub plot error that exists among the three subsamples receiving the different levels of power. It allows to estimate the error of the microwaves power applied to every sample.

An example of the resulting Analysis of Variance (ANOVA) table is presented in section III-3, in table III.

The full model now includes 108 parameters (with the additional of two parameters for estimating the power effect). It was therefore decided to generate 110 experiments, plus an additional set of six genuine replicates.

### Technical details of design set up

For a discussion about optimality criteria see references [2,5].

To design the experiments, three packages were at our disposal, none of which at this time could, alone, handle such a complex task:

– Design Expert [6] version 5 offers D and U generation criteria, does not handle mixture process designs.

– Design Expert beta version 6. Offers D optimality criteria, handles mixture process designs but a bug (now corrected) would not allow the fixing of the multiple constraints on the components.

– SAS/QC Proc OPTEX [7] offers several kinds of optimality criteria, including D and U, but does not handle multiple constraints on the components proportions, does not generate replicates and requires more memory than was available to handle large sets of candidate points.

The final strategy adopted was:

1. To use DX-5 to generate a U optimal design (Modified Distance Based) containing 31 different points. These 31 points were chosen among a total of 879 candidates in an attempt to reduce the amount of memory required by SAS when running its D-optimal criteria selection in step 3. The goal was to give SAS a large choice of points to select from. All point types (including thirds of edge and triple blends) were included in the set of candidates.

2. To fully cross the design with the variables $Z_1$ (2 levels) and $Z_2$ (6 levels) thus obtaining a set of $2*6*31 = 372$ points.

3. To use SAS to select, using the D-Optimality criteria, 110 points to which 6 genuine replicates were artificially added.

The experimental matrix is given table I.

As can be observed by looking at the matrix, there are only 20 different mixtures (not counting changes in $Z_1$ and $Z_2$ levels), left in the experiment after the D-Optimal selection step.

These are:

– 3 vertices (out of 40) labeled 'V' in the matrix.
– 2 center of edges (out of 100) labeled 'CE' in the matrix.
– 6 thirds of edge (out of 200) labeled '1/3E in the matrix.
– 6 triple blends (out of 400) labeled 'TB' in the matrix.

**Table I. Experimental matrix.**

| | | X3=0 | X2=0 | X3=0 | | X3=0 | X2=0 | X3=0 | X2=0 | X2=0 | X3=0 | | X2=0 | X2=0 | X2=0 | X3=0 | | | | X3=0 | |
| $Z_1$ | $Z_2$ | V | V | V | CE | CE | 1/3E | 1/3E | 1/3E | 1/3E | 1/3E | 1/3E | TB | TB | TB | TB | TB | TB | PC | PC | PC |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | | | | | 57 | | 65 | | 37/115 | 16 | | | | | | 99 | 51 | 81 | | |
| 0 | 2 | 11 | 45 | | | 56 | | 66 | | 38 | 17/113 | | 85 | 72 | 23 | 98 | 93 | | | | |
| 0 | 3 | 10 | | 107 | | | 78 | | 48 | | 18 | | 86 | 71 | | | | 50 | | | |
| 0 | 4 | 9 | | 106 | 102 | | | | | | 19 | 2 | | 70 | 22/114 | | 94 | 49 | | | |
| 0 | 5 | 8 | | 105 | | | | 67 | 47 | 39 | | | 87 | 69 | | 97 | 95 | | | 31 | |
| 0 | 6 | 7 | 46 | 104 | | 55 | | 68 | | | 20 | | 88 | | 21 | 96 | | | 82 | 32 | |
| 1 | 1 | 13 | 40 | 110 | | | | | | | | 6 | 83 | 75 | 26 | | 89 | | | 27 | |
| 1 | 2 | 12 | | 109 | | 61 | 76 | | | 33 | 14 | 5 | | 74 | | | | 54 | | 28 | |
| 1 | 3 | | 41 | | | 60 | | 62 | | 34 | | | | | 25 | 101 | 90 | | 79/116 | 29 | |
| 1 | 4 | | 42 | | | 59 | 77 | 63 | | 35 | | | | 84 | | 100 | 91 | | | 30 | |
| 1 | 5 | | 43 | 108 | | 58 | | | | | 15/112 | 4 | | | 24 | | | 53 | 80 | | |
| 1 | 6 | | 44 | | 103 | | | 64 | | 36 | | 3 | | 73 | | | 92 | 52 | | | 1/111 |

Accidental replicates in coalescent A, B, C, D:

| A |
| B |
| C |
| D |

◯  Designed Replicates.

– 3 constraint plane centroids (out 98) labeled 'CE' in the matrix.

Because of the extremely severe constraints on the component proportions, the distance between vertices and centers or thirds of edges is small, and the resulting loss of optimality from this selection strategy should not be significant.

Close examination of the matrix, reveals that setting up $Z_2$ in this manner generates a lot of "accidental" replicates. For example, the first vertex (third column in the matrix) has $X_3 = 0$, which means only one coalescent, the first in pair is present in the formulation. As shown in section II-2, one can see that level 1, 2 and 3 of $Z_2$ all define the same first coalescent, coalescent A in amount given by $X_2$.

As a result, for example:

STD = 10 and STD = 11, represent accidental replicate for coalescent A ($X_3 = 0$).

STD = 56 and STD = 57, represent accidental replicate for coalescent A ($X_3 = 0$).

STD = 9 and STD = 8, represent accidental replicate for coalescent B ($X_3 = 0$).

STD = 43 and STD = 44, represent accidental replicate for coalescent C ($X_2 = 0$).

STD = 33 and SDT = 35, represent accidental replicate for coalescent D ($X_2 = 0$).

All together, there are 32 degrees of freedom (6 from the designed replicates plus 26 from the accidental replicates) to estimate the experimental error and to test the homogeneity of the experimental error in the experimental domain. Because of the discrete nature of the responses (see below) it was decided to keep these accidental replicates so as to obtain a better estimation of the experimental error.

## Analysis of the design

### The responses

Eight responses were measured in this experiment. All responses except hardness were visually assessed and ranked on a scale from 0 (poor) to 5 (excellent). The list of all responses is given below.

*Block:* This response measures how easy it is to separate two identical samples placed together face-to-face with a fixed pressure immediately after oven drying. A block value 0 signifies that the samples were almost permanently glued together by the freshly dried coating (obviously very undesirable on an industrial finishing line) and 5 signifies that there was no sticking what so ever. For this response the pure experimental error based on 32 degrees of freedom is estimated at 0.36.

*Damage:* This response measures the degree of damage caused by the above mentioned block test. Block and Damage relate to what might happen when freshly dried, painted pieces are stacked together at the end of a finishing line. The pure experimental error for this response is estimated at 0.35.

*Sanding:* This response measures the capability of the freshly dried paint film to be sanded in an efficient and productive manner. It is usual for a piece to receive two coats of paint. There is usually a sanding operation in between the two coats. The pure experimental error for this response is estimated at 0.27.

The three first-mentioned responses indirectly measure the hardness of the paint film immediately after drying *via* microwave irradiation.

*Hardness:* Measurement of the hardness of the film. This measurement is realized by an instrument known as a König pendulum on the scale 0 to 40. The higher the value of this response, the harder the film is. The pure experimental error for this response is estimated at 1.43.

*Appearance:* Assessment of the esthetic aspect of the painted surface and underlying support after microwave oven drying. The pure experimental error is 0.54.

*Scratch:* Measurement of the friability or toughness of the film. The pure experimental error for this response is 0.32.

*Mar:* Capability of the coatings to resist against abrasion. The pure experimental error is 0.44.

*Stains:* Measurement of the chemical resistance of the coated surface after microwave drying. The pure experimental error is 0.15.

Some of the responses are highly positively correlated and this correlation depends on the microwave power. One example of correlation between responses (at $P = 1300$ Watts) is shown in table II.

Block, Damage and Sanding are the responses most highly correlated, because they represent indirectly the same basic properties.

### Analysis

Before starting the large task of analyzing all eight responses, the response "Block" was examined in detail first. All the analysis done so far is in L-pseudo-Components.

We defined 5 indicator variables $Z_{2,i}$ with $i = 1$ to 5 as:

– when $Z_2 = j$  $j = 1$ to 5  $Z_{2,i} = 1,$  $i = j$

  $Z_{2,i} = 0,$  $i \neq j$

– when $Z_2 = 6$  $Z_{2,i} = 0,$  $i = 1$ to 5.

Consequently the term $\{(3)\}Z_2$ in model $\{4\}$ or $\{5\}$ becomes:

$$\{(3)\}Z_{2,1} + \{(3)\}Z_{2,2} + \{(3)\}Z_{2,3} + \{(3)\}Z_{2,4} + \{(3)\}Z_{2,5}.$$

The whole design $\{5\}$ is a split-plot, with P as the subplot effect. All other effects are whole plot effects [4].

**Table II.** *P* = 1300 W correlation between responses.

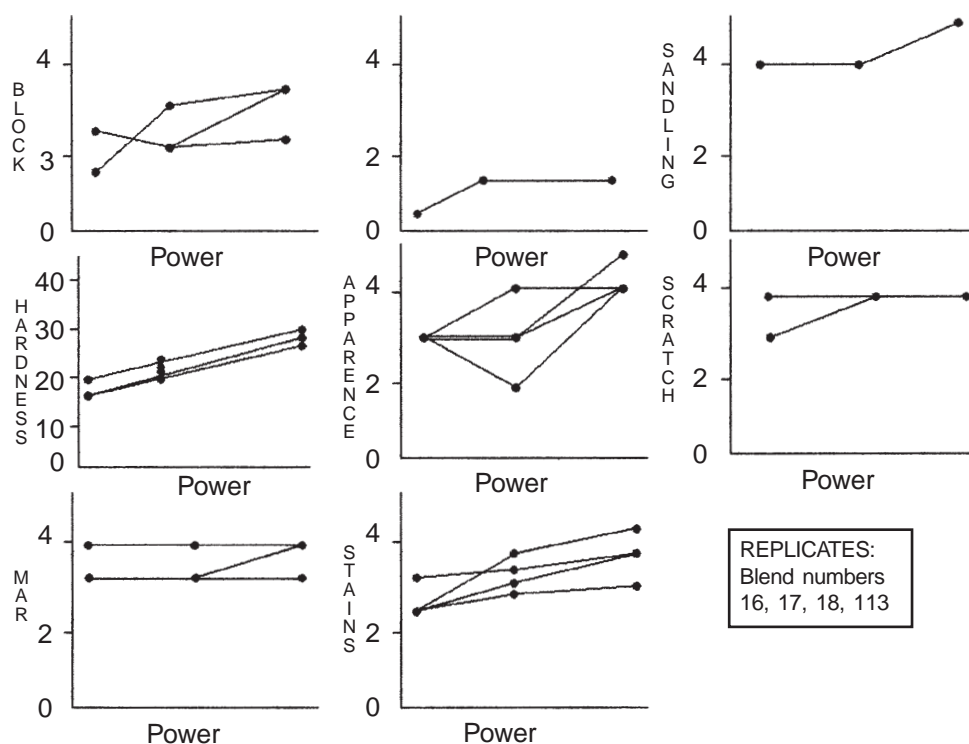|  | *Block* | *Damage* | *Sanding* | *Hardness* | *Apparence* | *Scratch* | *Mar* | *Stain* |
|---|---|---|---|---|---|---|---|---|
| **Block** | 1 | 0.87804 | 0.84066 | 0.63511 | 0.26609 | 0.57533 | 0.58162 | 0.51886 |
| **Damage** |  | 1 | 0.73890 | 0.54214 | 0.18101 | 0.51220 | 0.51093 | 0.42896 |
| **Sanding** |  |  | 1 | 0.64832 | 0.32902 | 0.58970 | 0.59660 | 0.46266 |
| **Hardness** |  |  |  | 1 | 0.31287 | 0.57157 | 0.57789 | 0.46956 |
| **Appearence** |  |  |  |  | 1 | 0.39318 | 0.41226 | 0.23105 |
| **Scratch** |  |  |  |  |  | 1 | 0.84054 | 0.58225 |
| **Mar** |  |  |  |  |  |  | 1 | 0.65124 |
| **Stain** |  |  |  |  |  |  |  | 1 |



**Figure 1. Visualisation of the power effect.**

The first step in the analysis of the data was to get a feeling of what kind of effect the microwaves power might have on the responses, particularly at the replicate points.

An example of the numerous graphs that were built is given in figure 1.

Nowhere it was possible to conclude that power had a quadratic effect, as was initially anticipated. This was confirmed by the result of the analysis of variance which was run as a second step. An example of such a table is given in table III for the response "Block".

The model built here has been simplified to show how the split-plot structure affects the degrees of freedom available for testing the significance of the effects. Note that there is no way so far to acknowledge in the analysis the presence of 26 accidental replicates. This was a real a problem that could not be solved, despite several attempts at "tricking" the package. Therefore, the number of degrees of freedom for error for testing the mixture component effect (main plot effect) is $116 - 6 = 110$. The number of degrees of freedom for error for testing both the linear and quadratic terms of power is $(116-1)*(3-2) = 230$.

When the model for the mixture components involves non linear blending terms as well as cross products between mixture and non mixture variables, estimation of the denominator degrees of freedom becomes extremely complex, but,

**Table III. Analysis of variance (ANOVA) showing the split plot structure.**

Solution for fixed effects

| Effect | Estimate | Sdt Error | Denominator DF | t | t Pr(t) |
|---|---|---|---|---|---|
| $X_1$ | 3.2634 | 2.8353 | 110 | 1.15 | 0.2522 |
| $X_2$ | 2.6484 | 2.1396 | 110 | 1.24 | 0.2184 |
| $X_3$ | 0.2969 | 2.0866 | 110 | 0.14 | 0.8871 |
| $X_4$ | 7.7392 | 2.1924 | 110 | 3.53 | 0.0006 |
| $X_5$ | 6.6902 | 8.4724 | 110 | 0.79 | 0.4314 |
| $X_6$ | −0.9488 | 2.9940 | 110 | −0.32 | 0.7519 |
| POWER | 0.3599 | 0.0336 | 230 | 10.69 | 0.0001 |
| POWER*POWER | −0.0898 | 0.0597 | 230 | −1.50 | 0.1338 |

Tests for fixed effects

| Source | Numerator DF | Denominator DF | Type III F | Pr > F |
|---|---|---|---|---|
| $X_1$ | 1 | 110 | 1.32 | 0.2522 |
| $X_2$ | 1 | 110 | 1.53 | 0.2184 |
| $X_3$ | 1 | 110 | 0.02 | 0.8871 |
| $X_4$ | 1 | 110 | 12.46 | 0.0006 |
| $X_5$ | 1 | 110 | 0.62 | 0.4314 |
| $X_6$ | 1 | 110 | 0.10 | 0.7519 |
| POWER | 1 | 230 | 114.23 | 0.0001 |
| POWER*POWER | 1 | 230 | 2.26 | 0.1388 |

in principle, there are always about a hundred, and two hundreds degrees of freedom for estimating mixture effects and power effects respectively.

The real difficulty occurred during the third step, when it was time to derive the "best model" that could both adequately describe the response surface and also lead to a chemical interpretation of the role that the types of binder and coalescent ($Z_1$, $Z_2$) were playing. Three regression procedures Backward, Forward and Stepwise were used in turn to select the fewer number of terms in model {5}.

All derived models have about 30 significant terms (out of the 106 possible terms in the model) and present an equivalent Adjusted RSquared (around 0.8). For all models, the residuals are approximately normally distributed, and some lack of fit is present for each model (based on a mean squared error with 32 degrees of freedom). Unfortunately, these models are different in the terms they contain as well as in the magnitude of the coefficient estimations. Thus, the asymmetric nature of the design as well as the fractionating on $Z_1, Z_2$ was making the interpretation of the model hazardous.

The main risk of fitting one single overall combined model lies in picking up non linear blending somewhere in the experimental region and projecting it, to other locations of the experimental domain where non linear blending is not present [8]. Instead a more accurate picture of the mixture surface would be obtained by fitting a mixture model to the data at each combination of $Z_1$, $Z_2$ variables. Because of this, twelve individual models have been fitted at each combination of the variables ($Z_1 \times Z_2$).

At this stage, the number of data points for each combination of $Z_1$, $Z_2$ (see experimental matrix) was rather low and there were not particularly well chosen, but no other alternative was available.

After validation of these twelve models, the following predictions and interpretations could be made.

## Interpretation

At this stage, different tools helped to extract the main interpretation and conclusions for this project. The first tool that we used was the perturbation plot for each separate model.

### Perturbations plots

For each separate model we used the perturbation plots (also called traces) to predict the formulation rules. The traces

$X_1$= Binder
$X_2$= First coalescent
$X_3$= Second coalescent
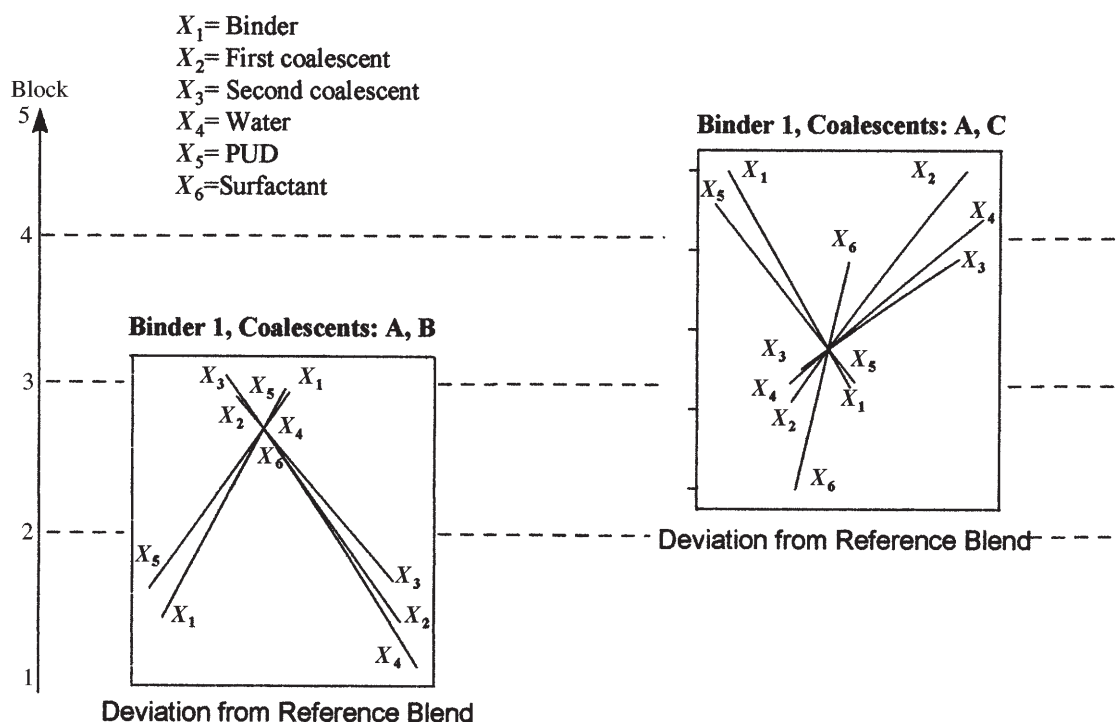$X_4$= Water
$X_5$= PUD
$X_6$=Surfactant

**Figure 2. Example of a perturbation plot (traces) for the response block.**

show the effect of changing each mixture component while holding all others in a constant ratio. The response is plotted while moving along an imaginary line from a reference blend towards the vertex of the component being incremented. The default reference blend is the centroid of the design. As the amount of a particular component increases, the amounts of all other components decrease, but their ratios to one another remain constant as defined in the reference blend.

In this case, traces allows to find those mixture components that affect the response most, but also to compare binder and coalescent types by juxtaposition, on the same scale, of the set of traces corresponding to each separate model. An example for Block is presented in figure 2.

The reference blend is defined as:

| | |
|---|---|
| Binder | 32.49 |
| Coalescent 1 | 3.04 |
| Coalescent 2 | 3.95 |
| Water | 17.44 |
| PUD | 2.77 |
| Surfactant | 0.31 |

For blends containing coalescent *A* and *B* (left most graph), decreasing the amount of binder 1 (other components move in ratio defined by the reference blend) decreases block. However for blends containing coalescent

*A* and *C* (right most graph), the behavior of the binder differs, and we observe an increase in block, instead. It was observed that the effect of the surfactant is negligible in blends containing coalescent *A* and *B*, but is very influential in blends containing *A* and *C*.

In the same manner, comparison between all combinations of sets of traces (one set per model) allows to determine the specific rules for all components for the twelve models. In particular this comparison leads the conclusion that, binder 1 behaved almost consistently better than binder 2.

### Effect of the microwave drying power on paint properties

A split-plot analysis has been conducted on the separate models to identify and quantify those properties that are most affected by the microwave drying.

Influence of microwaves drying on each response (scale 0 to 5) is reported in table IV. The numbers in the table (except for hardness) represent expected changes in the response when, for a given blend, drying from mid power level to high power level. A decrease of the same magnitude is expected when drying from mid power level to low power level.

**Table IV. Effect of the microwave power on each responses.**

| Z1 = 0 | Z2 | Block | Damage | Sanding | Hardness | Apparence | Scratch | Mar | Stains |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 0.69 | 0.44 | 0.34 | 2.49 | None | 0.19 | 0.25 | 0.09 |
| | 2 | 0.54 | 0.5 | 0.42 | None | None | 0.25 | 0.21 | 0.17 |
| | 3 | 0.53 | 0.69 | 0.25 | 2.35 | None | None | 0.19 | 0.13 |
| | 1 | 0.45 | 0.52 | None | 1.79 | None | None | 0.25 | 0.14 |
| | 5 | None | 0.62 | 0.5 | 1.77 | None | 0.25 | 0.37 | 0.15 |
| | 6 | 0.27 | 0.5 | 0.27 | 1.74 | None | 0.45 | 0.23 | 0.17 |
| **Z1 = 1** | | | | | | | | | |
| | 1 | 0.47 | 0.72 | 0.58 | 1.74 | None | 0.33 | 0.42 | 0.19 |
| | 2 | 0.35 | 0.45 | 0.3 | 1.73 | None | 0.15 | 0.35 | 0.18 |
| | 3 | 0.37 | 0.5 | 0.6 | 1.45 | None | 0.35 | 0.47 | 0.20 |
| | 1 | None | 0.28 | 0.38 | 1.58 | None | 0.22 | 0.3 | 0.15 |
| | 5 | 0.28 | 0.22 | 0.36 | 1.71 | None | None | 0.28 | 0.21 |
| | 6 | None | None | None | 1.2 | None | None | 0.15 | 0.22 |

All responses assessed on a scale from 0 (poor) to 5 (excellent) except hardness, measured on scale from 0 to 40.
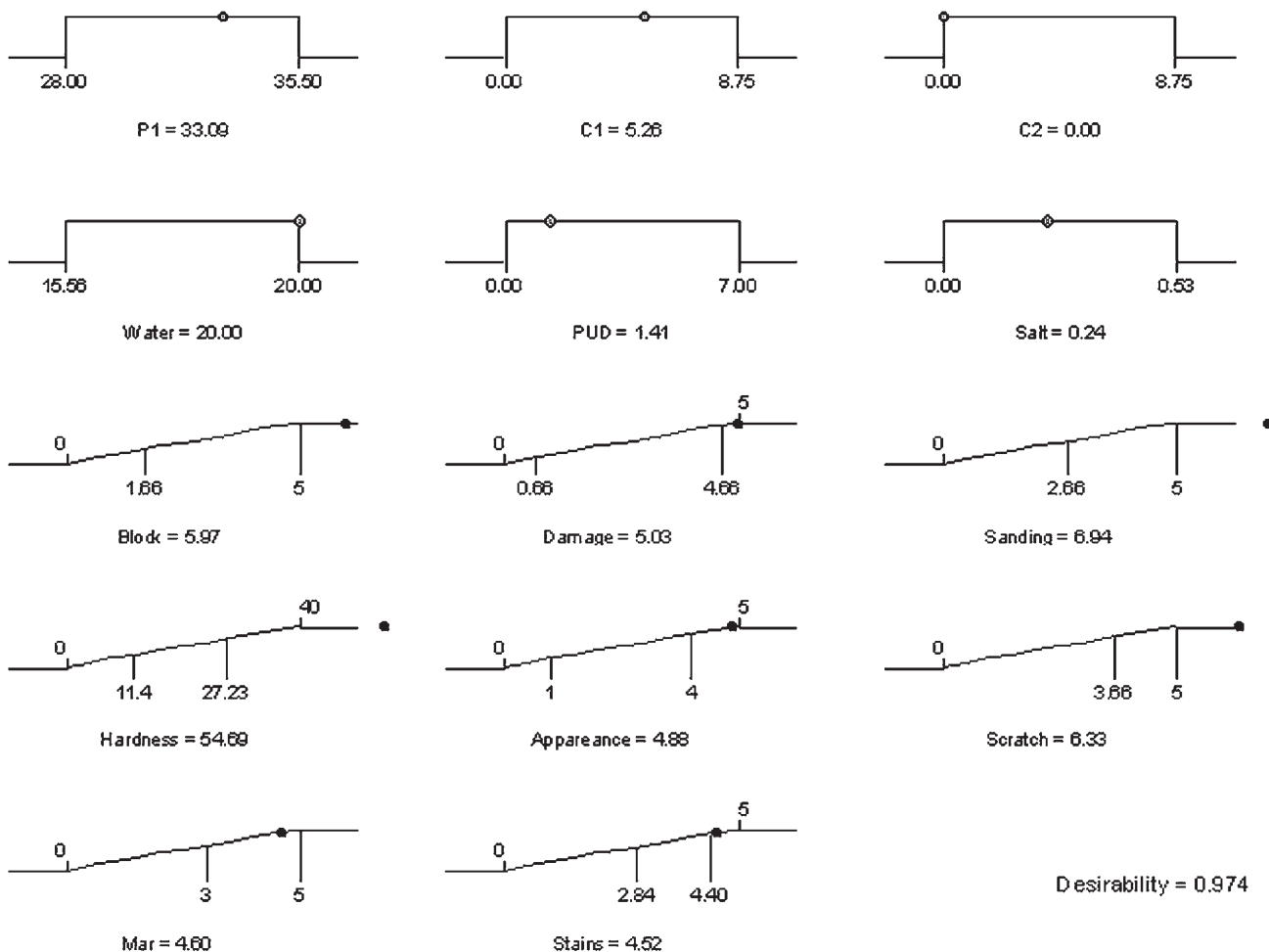


**Figure 3. Best formulation following the result of the desirability algorithm.**

### Definition of the optimal region. Desirability algorithm

Because the individual response surfaces do not present the same optimum regions, a "desirability algorithm" has been used.

Our aim was to find which blend most closely simultaneously optimises the eight measured properties. The optimisation searches for a combination of mixture component levels that simultaneously satisfy the requirements placed on each response. Numerical optimisation was used rather than a graphical one. Box and Draper (empirical Model Building and Response Surface) describe a method due to Derringer and Suich [9] which can determine the best combination of responses.

Following their example, an objective function is defined: $D$, corresponding to the desirability function, which reflects the desirable ranges for each response ($di$). The desirable ranges are from 0 to 1 (least to most desirable, respectively). The simultaneous objective function is a geometric mean of all the responses:

$$D = (d_1 \times d_2 \times d_3 \ldots d_n)^{1/n} = (\Pi\ d_i)^{1/n}$$

where $n$ is the number of responses.

In this problem, the target for all responses is the maximum. An additional parameter, called "weight", affects the transformation of each response. The weight is used if one wants to give added emphasis to the upper/lower bound, or to the target value. A weight of 1 corresponds to "the partial desirability" $d_i$ varing from 0 to 1 in a linear fashion. Weights greater than 1 give more emphasis to the goal, on the other hand, weights less than 1, give less emphasis to the goal.

In this case, a linear increase was selected for partial desirability to reach the target (maximum) and different importance were put on each response in order to achieve the best reasonable compromise. Block was the most important response, stains was the least important one. The desirability algorithm has been used with each model (or combination of process variables) to find the best compromise. One example of the use of the desirability algorithm is shown below.

Example for binder 1 ($Z_1 = 1$) is given in figure 3. The first two lines indicate the actual composition of the recommended blends. The three last lines indicate the predicted value of the eight properties of interest.

By doing so, one can find the best mixture based on each model which simultaneously optimizes the eight responses. This allows for creation of customized paint formulations based on the specific needs of a given industrial finishing line.

## Conclusion

Numerous difficulties in this study were encountered, not all of them were anticipated.

In particular, we had in mind a rather generic definition of a process variable as being a factor in the experiment that does not form any portion of the mixture but whose levels, when changed, could affect the blending properties of the experiment [10]. Thus, binder type and the coalescent type were incorrectly classified as process variables. They should have been instead defined as categories of components (or major-components) including respectively two and four components each (or minor components) [10].

This study can be defined as a categorized-component mixture study where some complexity arises from the presence of a large number of major-components as well as from the three restrictions placed on the blending, namely:

– only one binder can be present in any blend, and one binder must be present;
– at most two coalescents can be present in any blend and at least one must be present in all blends;
– of the 9-components at most 6 are present in any blend.

To our best knowledge no reference work describes a study as defined above. Nor was there a software package available to handle both the design and the analysis of the experiment. This will eventually be addressed in the future, meanwhile we recommend to design from the start as many "small" mixture experiments as there are combinations of minor component types. One advantage of such an approach over fitting one unique reduced model is to ensure no curvature will be picked up for some given binder or coalescent types that will be carried over for some other binder or coalescent types. One disadvantage is that, to limit the total number of blends to make, simpler models will have to be assumed adequate. The amount of work to invest is to be put in balance, as always, with the quality of the resulting information.

This study has demonstrated that the usage of different types of binders (acrylic emulsion polymers) and coalescents in combination with microwave drying could have a profound impact on a coating formulation's performances.

The major effects of all components have been established in this investigation, and can be both negative and positive. We have been explained the interaction between the two types of components (binder and coalescent) and with the understanding of this complex relationship, paint formulators can balance the choice of components and additives and develop formulations offering optimal performances in association with microwave drying with minimal detrimental effects from additives.

## Acknowledgements

## References

1. Scheffé, H. *Journal of the Royal Statistical Society B* **1958,** *20*, 344-360.

2. Cornell, J. A. *Experiments with Mixtures*; New York: John Wiley & sons, 2$^{nd}$ edition, 1990.

3. Piepel, G. F.; Cornell J. A. *Technometrics* **1985**, *27*, 219-227.

4. Cornell, J. A. *Journal of Quality Technology* **1988** *20*, (1), 2-23.

5. Myers, R. H.; Montgomery, D. C *Response Surface Methodology*; New York: John Wiley & sons, 1995, p 364.

6. Stat-Ease Incorporated, Minneapolis, MN 55413 USA.

7. SAS Institute Incorporated, Cary, NC, USA.

8. Cornell, J. A. *Journal of Quality Technology* **1995**, *27*, (1), 13-33.

9. Derringer, G. C.; Suich R. *Journal of Quality Technology* **1980**, *12*, (4), 373.

10. Cornell, J. A.; Ramsey, P. J. *Technometrics* **1998**, *40*, (1), 48-61.