

Finding Rules for Shampoo Foam Height with FormRules

Background

When using personal care products like shampoos, customers often gauge the efficacy of the product by properties like the amount of foam that is generated. This is because this is easy to assess visually, while other relevant properties may be less immediately obvious.

In a study reported in the book *Response Surface Methodology: Process and Product Optimization using Designed Experiments*, R H Myers and D C Montgomery give data from a D-optimal design. Three different ingredients were varied, subject to a constraint that the total of all three has to add up to 50% of the total.

The three ingredients were:

Lauryl sulphate: between the limits 20% and 30% (X1)

Cocamide: between the limits 7% and 10% (X2)

Lauramide: between the limits 13% and 20% (X3)

Only one property – foam height in mm – was measured.

Montgomery then examined his data using statistical techniques. Here, we have used **FormRules**, a program based on neurofuzzy logic, to discover the relationships within the data. We note that it is not a prerequisite for **FormRules** that designed experiments be used; however, that was what was available in the literature.

Only 9 unique experiments were available, since the 13 reported by Montgomery contained several replicates. In our study, the replicate values were averaged, so as not to bias some regions of the design space unnecessarily.

Model for Foam Height

The 9 different data points, using the 'actual' values from Montgomery's book, were put into a spreadsheet and copied into **FormRules**. The default model assessment criterion, Structural Risk Minimization, did not detect useful relationships within the data when the default value of C_1 (calculated by FormRules to be 0.81) was used. However, reducing C_1 to 0.7 (which allows a more complicated model to be developed) gave a good model with an R^2 value of 0.885 given by the ANOVA statistics.

Changing the model selection criterion to Minimum Descriptor Length (which again will allow a somewhat more complex model) gives the same results as the Structural Risk Minimization with C_1 value of 0.7.

These results show that a good model can be developed from the relatively limited amount of data available. The fact that R^2 is not overly high ($>.99$) gives a good indication that over-training (that is, fitting to the noise in the data) is unlikely to have occurred.

Results for Foam Height

The model showed that the most important property in controlling foam height was the amount of cocamide. This is illustrated in the figure below, where the purple line denotes the variable that contributes most strongly to the model.

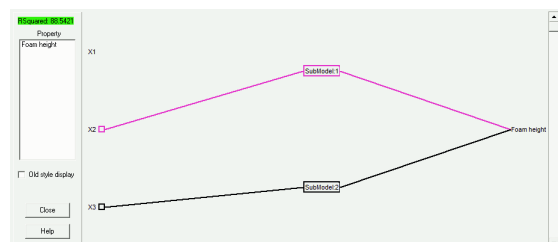


Figure 1. Graphical representation of model for foam height

There is no interaction between the two variables, as indicated by the fact that each contributes via a separate submodel. Not surprisingly, only two variables contribute – this can be expected, since the variables sum to a total of 50% so are not totally independent.

The rules for foam height that **FormRules** gives from this model are:

For the first submodel, involving the amount of cocamide (X2),

IF X2 is LOW THEN Foam height is LOW (0.52)
IF X2 is MID THEN Foam height is HIGH (1.00)
IF X2 is HIGH THEN Foam height is LOW (1.00)

while for the second submodel, involving lauramide (X3)

IF X3 is LOW THEN Foam height is LOW (0.77)
IF X3 is MID THEN Foam height is HIGH (1.00)
IF X3 is HIGH THEN Foam height is LOW (0.88)

The rule given in blue makes the greatest positive contribution to the model, while the rule given is red makes the largest negative contribution.

The relationship described by first submodel can be expressed graphically as shown in Figure 2.

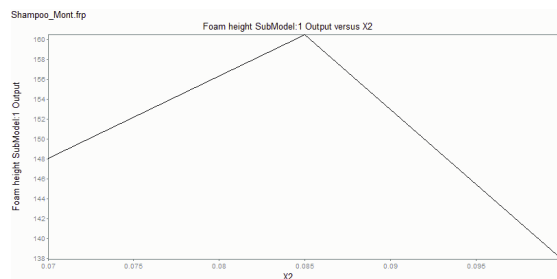


Figure 2. Relationship between amount of cocamide and foam height

Clearly there is a non-linear relationship, with foam height at its maximum when the amount of cocamide is in the middle of the range. Large amounts of cocamide will lead to the lowest foam height.

As mentioned above, the amount of lauramide has a lesser effect on the foam height. The functional form of the relationship is described in Figure 3.

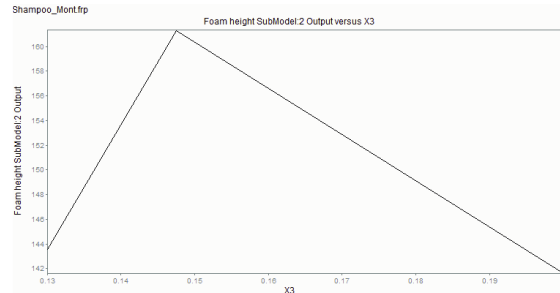


Figure 3. Relationship between the amount of lauramide and foam height

As for cocamide, the foam height is highest when the amount of lauramide lies towards the middle of the range – although by the time it reaches mid-range, the foam height is already decreasing.

The straight-line representations in both Figures 2 and 3 arise because of the way that **FormRules** works, fitting B-spline functions to the data.

Conclusions

Despite the relatively small amount of data, **FormRules** has been able to find relationships within the data. These show that the most important variable is the amount of cocamide, followed in importance by the amount of lauramide.

There is no interaction between the two variables, and maximum foam height is achieved when amounts of both cocamide and lauramide lie towards the middle of their ranges.

The rules were discovered automatically with little intervention on the part of the user – the only requirement was to change the model selection criterion from the default values. This means that it was considerably easier to find the key relationships using **FormRules** instead of a statistical study.

For further information on FormRules, and applying neurofuzzy logic to your problems, contact us at the address below.

© 2009 Intelligensys Ltd
All rights reserved.