

Generating Models and Rules for Spheronization

Introduction

In their paper “Studies of parameters important in the spheronisation process”, Baert, Vermeersch, Remon, Smeyers-Verbeke and Massart¹ have investigated a binary mixture of Avicel PH 101 and water, looking how changing the Avicel/water ratio, the spheronization time and the spheronization speed affects the yield and roundness of the spheronized particles. They used a design-of-experiments approach, with 3 spheronizer speeds, 5 spheronization times, and 3 different water/Avicel ratios. This led to 45 unique combinations of formulation and process conditions.

These 45 results were used in **FormRules** to examine the key relationships and to extract the rules that govern the interactions between the ingredients and process conditions, and the 2 measured properties. The inputs were chosen to be Avicel%, spheronization speed and spheronization time. The amount of water was not considered; since water% + Avicel% = 100, these are not independent variables.

In an initial examination of this problem, Structural Risk Minimization was used as the model selection criterion; this should give relatively simple, robust models provided that the data contain sufficient cause-and-effect information (i.e. there are no important variables that have remained unmeasured). The auto-select feature was used to choose the value of C1; **FormRules** suggested a value of 0.89 for this parameter and this value was used in all subsequent studies. This gave a complicated model for Roundness, as shown in Figure 1. In this case, spheronization speed contributes to *both* submodels. This makes it very difficult to interpret the rules; although the models give a good fit to the training data, it can be at expense of meaningful rules.

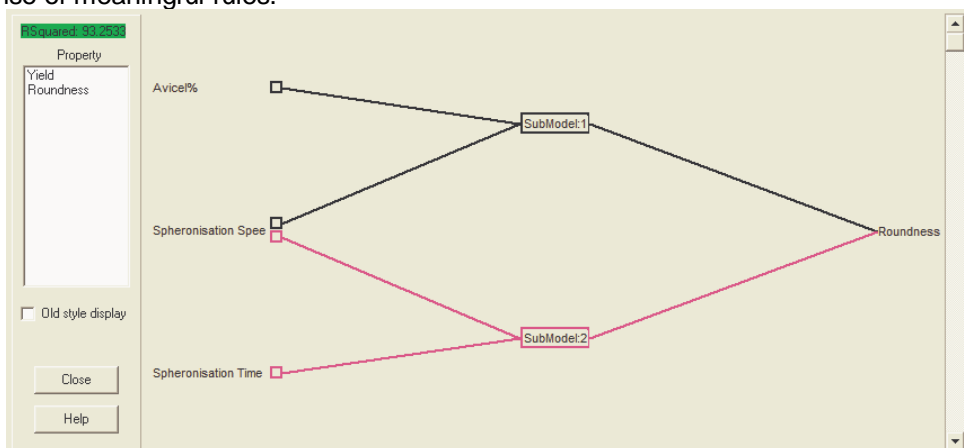


Figure 1. SRM Model for Roundness, calculated using default parameters

This situation whereby the same input contributes to two submodels should be avoided if possible. For the present data set, this could be achieved in one of two ways: reduce the C1 parameter from the default 0.89 to 0.84, or else use Minimum Descriptor Length (MDL). For both of these cases, the model for Roundness involved a single submodel, to which all 3 inputs contributed. The rest of this report will discuss the MDL results.

Results

Good models were found for both properties, as assessed by ANOVA (analysis of variance) statistics.

Models and Rules for Yield

For Yield, the ANOVA statistics showed that the R^2 value was 92.6%. Interestingly, spheronization time did not contribute to the model for the Yield, as Figure 2 illustrates. There is an interaction between the Avicel% and the spheronization speed, shown by the fact that both contribute to the same submodel. If there was no interaction, then there would have been two separate submodels.

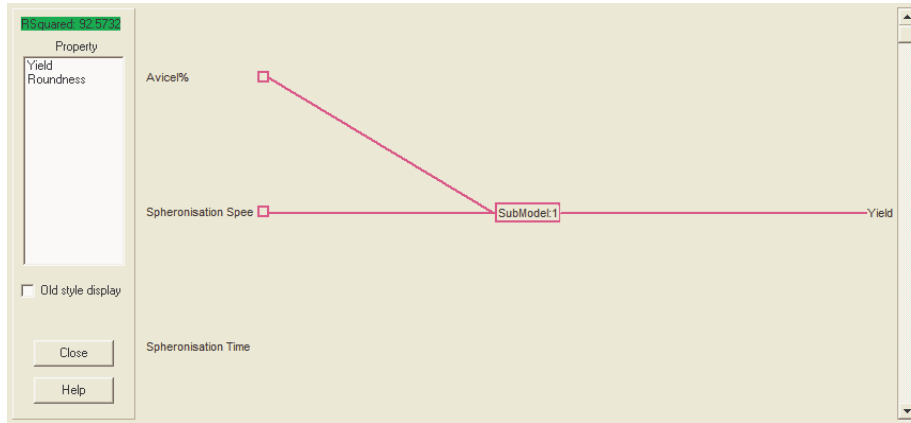


Figure 2. Graphical display of model for Yield

More detail about the models can be found readily from the **FormRules** output. The neurofuzzy approach used in **FormRules** also allows the models to be expressed as rules, and these are summarized in Table 1 for Yield.

Spheronization Speed	Avicel %	Yield (Confidence)
Low	Low	High (0.89)
Low	Mid	High (0.87)
Low	High	High (0.73)
Mid	Low	High (0.77)
Mid	Mid	High (0.65)
Mid	High	Low (0.84)
High	Low	High (0.61)
High	Mid	Low (0.55)
High	High	Low (0.96)

Table 1. Summary of rules for Yield

In Table 1, the rule that contributes the most to the High value is shown in blue, while the rule that makes the smallest positive contribution (or largest negative contribution) is shown in red.

When the 'confidence levels' (shown in parentheses) have values in the mid-range (e.g. 0.50) it means that the value of yield is neither low nor high – it is in the middle. For example, a confidence level of 0.60 that the yield is Low means that the value lies somewhat below the mid-point, i.e. that it is 60% towards the 'low' value along a straight line connecting the Low and High values, as illustrated in Figure 3.

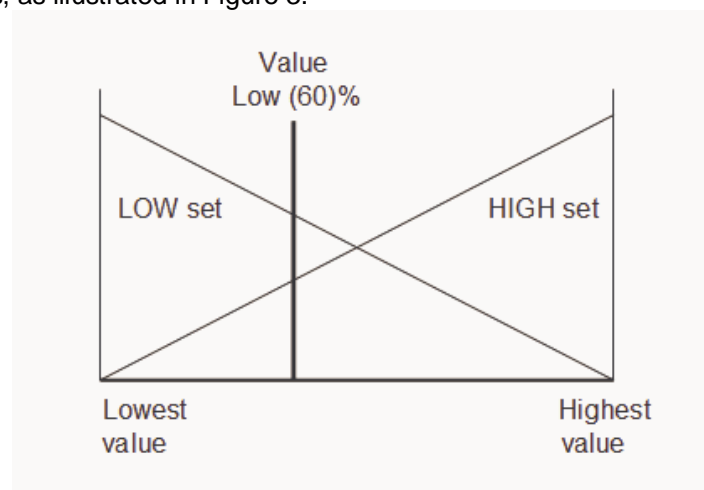


Figure 3. Definition of 'confidence level' for a particular value

The rules in Table 1 show that in all cases, increasing the Avicel percentage (i.e. increasing the amount of water) decreases the yield. The yield is higher when the spheronization speed is low, and lower when spheronization speed is high. The results are summarized in the 3D plot given in Figure 4, which shows how this submodel contributes to the overall value of the Yield.

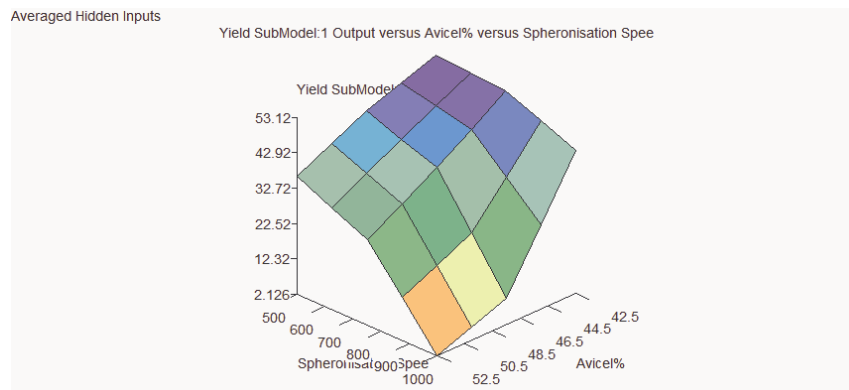


Figure 4. Effect of Spheronization Speed and Avicel percentage on Yield

The statistical study of Baert *et al* found that spheronization time was also important in determining the yield, a finding which is not upheld in the results presented here. They treated each formulation separately, rather than dealing with them all together as has been done here. Their R^2 values were not better than those found in the present study, though. When they did use all data together, they found that Avicel % times speed explained about 64% of the variation in the yield. Our results show, however, that the interaction is more complex than a simple product of the variables, and our models give a better fit to the data.

For the individual formulations, Baert *et al* also found that the square of the speed (rather than the speed itself) was important, with increased speed leading to smaller yields. This behaviour has been picked up by **FormRules**, with the effect of spheronization speed on Yield shown in Figure 5. In a subsequent study with **FormRules**, undertaken specifically to investigate this behaviour, $Speed^2$ was used as an input parameter, and this study confirmed that there was a linear decrease with $Speed^2$. However, in this alternative scenario, R^2 decreased slightly to 90.5%.

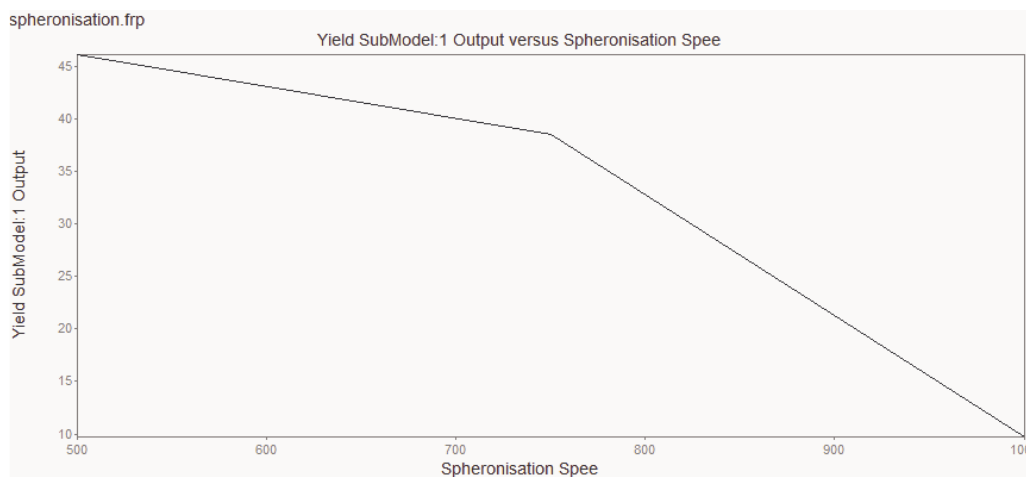


Figure 5. Effect of spheronization speed on Yield

Models and Rules for Roundness

Roundness is defined as the ratio of the largest to the smallest diameter of the spheres (i.e. aspect ratio), so that the most spherical particles have Roundness values close to 1. For the Roundness model, the R^2 value using the MDL model selection criterion was 92.0%, indicating a very reliable model. Not surprisingly, Spheronization Time was important here,

as Figure 5 shows. In fact, there is an interaction between all three variables Avicel %, spheronization time and spheronization speed.

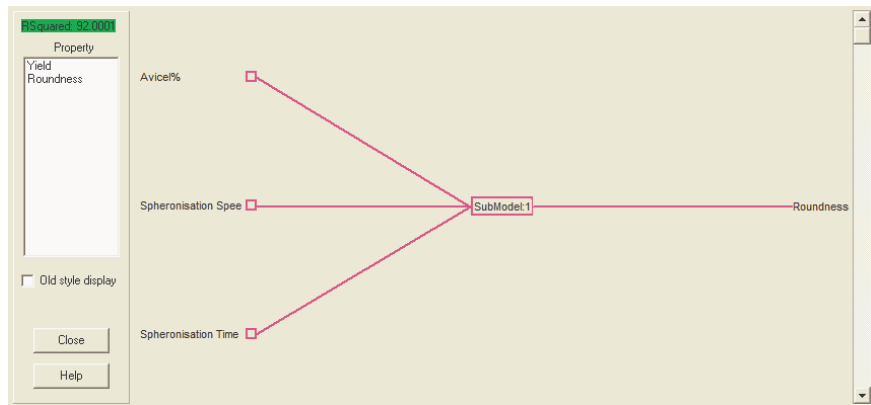


Figure 5. Graphical display of model for Roundness

Spheronization speed has its greatest effect when the Avicel % is high, as Figure 6 shows. Figure 6 assumes that the spheronization time is about 15 minutes – i.e. in the middle of the range of experimental values used. If the spheronization speed is low and the Avicel% is high, then the particles are markedly aspherical. If Avicel % is low (i.e. water content is high) then the spheronization speed has only a small effect.

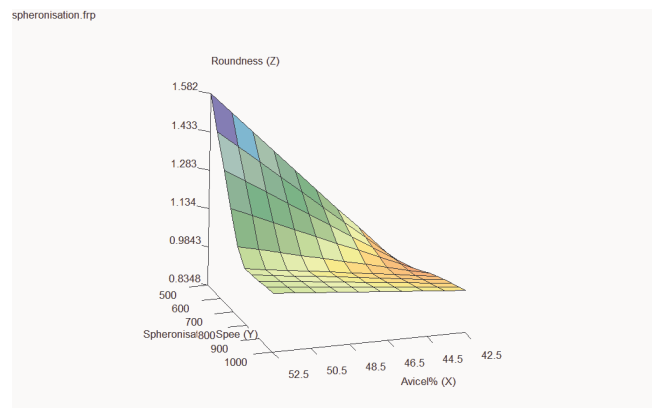


Figure 6. Effect of Spheronization speed and Avicel % on Roundness, for spheronization time of 15 minutes

When the spheronization time is low (1 minute), different behaviour is observed as Figure 7 shows. In this case, the spheronization speed is important regardless of the Avicel %.

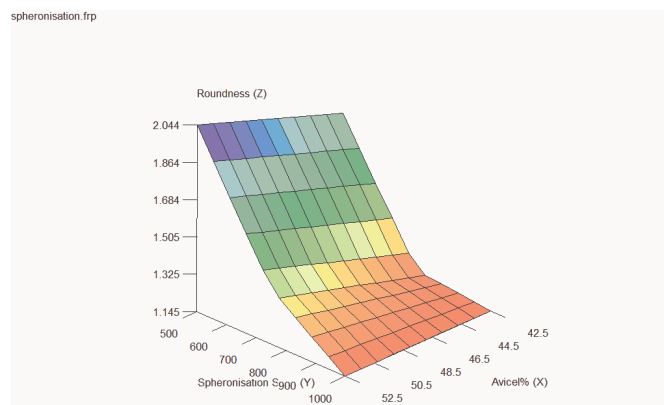


Figure 7. Effect of Spheronization speed and Avicel % on Roundness, for spheronization time of 1 min

The models can be presented in the form of rules, shown in Table 2. Here, because there are 3 inputs to the submodel, the rules are of the form IF spherization speed AND spherization time AND Avicel% then Roundness...

Spherization Speed	Spherization Time	Avicel %	Roundness (Confidence)
Low	Low	Low	High(0.72)
Low	Low	High	High (0.97)
Low	Mid	Low	Low (1.0)
Low	Mid	High	High (0.53)
Low	High	Low	Low (1.0)
Low	High	High	Low (0.51)
Mid	Low	Low	Low (0.85)
Mid	Low	High	Low (0.63)
Mid	Mid	Low	Low (1.0)
Mid	Mid	High	Low (0.98)
Mid	High	Low	Low (0.96)
Mid	High	High	Low (0.96)
High	Low	Low	Low (0.88)
High	Low	High	Low (0.88)
High	Mid	Low	Low (1.00)
High	Mid	High	Low (0.96)
High	High	Low	Low (0.99)
High	High	High	Low (0.99)

Table 2. Effect of spherization speed and Avicel % on Roundness

The rule that makes the largest positive contribution is shown in blue, while the rule that makes the smallest positive (or largest negative) contribution is shown in red.

Conclusions

Neurofuzzy data mining has pulled out the key relationships between the formulation and process variables and the 2 measured properties. Baert *et al*¹ invoked the fact that temperature and relative humidity may be important to explain the lack of fit in their models. However, the fact that good models (evidenced by the high R² values) were obtained using the neurofuzzy approach indicates that these parameters are not actually of great significance.

The models extracted here were better than those found from the statistical study. Furthermore, they were more easily obtained, since it was not necessary to transform the data (as was required for roundness, in the work by Baert *et al*) nor was it necessary specially to investigate combinations of variables. These interrelationships were determined automatically within the neurofuzzy study.

References

(1) L Baert, H Vermeersch, J P Remon, J Smeyers-Vereke and D L Massart, International Journal of Pharmaceutics 96 (1993) 225-229