

Computer aided rubber compound formulation using design of experiments

Ransi Devendra^{1,2}

¹ Uroxsys Ltd., Auckland, New Zealand.

² The University of Auckland, The school of Chemical Sciences, Auckland, New Zealand.

Abstract:

The article explains how the technique of design of experiments (DOE) can be used to formulate rubber compounds with the aid of commercial software. Some background information in theory followed by advantages of the factorial method over conventional methods that are practised in the industry are discussed. Then an example of a two-level factorial study related to rubber compounding is explained. Finally the best combination of the ingredients in the formulation for a set of desired properties are determined using an optimization method available in the software.

Introduction

In the rubber industry the products are made by using different elastomers, compounding chemicals and processing methods. In the rubber product manufacture it involves an accurate combination of all of these factors. These are generally practised in the rubber industry using trial and error methods or with the knowledge from previous experience. Sometimes traditional scientific method which dictates changing one factor at a time (OFAT) is also used. Most of the time these techniques can take many man hours to produce a product. Unfortunately, at the end very little is learned about how these factors (ingredients and processing conditions) related to a selected response (property) of the manufactured item which is important in product development. These drawbacks can be overcome through taking a systematic design of experiments^{1,2} multifactor approach.

Design of experiments is an established and proven technique for product and process development in all sectors of the industry³. Design of experiment is an invaluable method to identify critical parameters in optimization of chemical processes and identify robust operational regions for processors. It will have the advantage of getting the product's knowledge early in the development cycle and reduce the time to market the product.

Using design of Experiments protocol³ product and process development are done by making a series of controlled intervention according to a specified plan. A well designed experiment will use the smallest number of experimental runs to get the maximum amount of information. These experiments should be able to clearly separate the effective parameters from the noise (experimental error) of the system. This way it is possible to screen out the important few parameters from the trivial many.

The experimental design process involves three stages i.e. planning, implementation and analysis. Planning involves some knowledge on the system that is being investigated when it comes to the selection of factors and responses. This means for such action can only be done by persons with previous experience related to the similar systems. Implementation involves carrying out experiments by changing factors in a systematic fashion and minimising the effects of uncontrollable factors (noise). The changes in the factors should cover the entire design space. Once the experiments are carried

Ransi Devendra is the Science and Technology Manager at Uroxsys Ltd., New Zealand. He had been developing products based on urethane, epoxy and silane chemistry for the past 25 years. Before migrating to New Zealand he had been working at CISIR in the Rubber Technology section as a Research Officer and as a Resin Chemist in the paint industry in Sri Lanka. Ransi's professional career has involved research and development work in polymer related topics for the past 40 years. He also has expertise in multiscale modelling of material. Ransi has a B.Sc. in Natural Sciences, M.Sc. (Polymer Science and Technology) and a Ph.D. in Chemistry (Investigation on urethane catalysts using computational chemistry).



out, mathematical modelling is used to find the relation between the input factors and the response for further analysis.

Planning of experiments

The first step in any experiment is to have a clear objective. The objectives should be translated into precise questions that the experiments can answer. It is important to find the factors that will affect the performance of the final product. As a general rule is to include in the design any borderline factors, since to eliminate them early in the process is much cheaper than to identify and introduce them further down the line.

Implementation

Next is to identify the factors and response and check the levels of the factors that give the best response. If the responses are linear the factor levels can be two. If there are nonlinear factors more than two levels are selected.

Scientific learning is iterative, the results from each stage of investigation generate questions to be answered during the next stage. It may even turn out a different factor, not initially suspected of being important, will dominate future investigations. Therefore never spend more than 25% of your experimental resources on the initial experiments².

Mathematical modelling of factors- response relation

The relation between the recipe ingredients and process conditions on selected properties of a product can be investigated using mathematical modelling methods. The aim of mathematical modelling is to derive a function or a formula that closely reproduces the behaviour of the system under investigation. This mathematical function when presented with input values, it predicts the output values of a property close to what is observed experimentally. Modern day software that is used for such applications also has the ability to visualise such data in graphical form when subject to various analyses. One of the graphical outputs is the response surface⁴ which is a surface that is generated in a 3D coordinate space using values of a property that are calculated by varying the values of two factors (ingredients or process conditions). Use of such visual aid enhances the identification of critical ingredients and process conditions. Through this method it will be possible to design products with optimum performance efficiently.

Statistical⁵, neural network⁶ and multiple linear iterative³ are three mathematical modelling techniques

that are used in the above mentioned investigations. All give very promising results but each has its own advantages. In statistical methods the data points (factor levels) in the design space are selected using theory to minimise experimental errors. Thereafter with experimental results a mathematical model is built using regression techniques. In neural network methods the model building is data driven. Therefore more and more data is needed to build an accurate model. However, this method can detect very complex interactions of the factors (ingredients or process conditions) to a selected property. Neural network methods also have the capability of optimising a higher number of factors more effectively compared to statistical methods. One of the advantages of this method over statistical methods is that it can use data that is already available in a manufacturing environment. This becomes very handy for rubber companies since they have thousands of formulas with evaluated properties in their databases. Both statistical and neural network modelling can produce recipes after optimization of selected desired properties. Mathematical modelling based on multiple linear iteration techniques has been recently implemented⁷ and it is getting popular in the rubber industry for recipe development. This is similar to neural networks where historical data can be used to generate new recipes. In this paper application of statistical methods in design of experiment will be discussed.

Statistical method

Design of experiments using statistical methods generally follow two steps⁸. Initially a screening test is done on factors that are not sensitive to a selected response to eliminate them. This is done using a two level factorial method. The advantage of the screening process will reduce the number of experiments to generate the response surface by limiting the number of factors to a lower number. Once the sensitive factors are identified additional levels for the factors are introduced to generate response surfaces. The additional data points will help in mapping the curvature of the response surface accurately for optimization.

Design of experiments -Two level factorial

Two level factorial analysis based on F-test was developed by Sir Ronald Fisher, a geneticist¹. He developed this theory to compare treatments that are used in agricultural experiments. F-test compares the variance among the treatment means versus the variance of individuals within specific treatments. High values of F-test indicates that one or more of the means differ

from another. This is a valuable information to find factors that are sensitive to a selected property.

In design of experiments each factor value is changed and the response of a selected property is determined. In a set of experiments where the factors are varied between two levels can be done by two methods. One method is to change one factor at a time. Here only one factor is change while keeping all other factors at one level. This is called the OFAT (one factor at a time) method. Many people in industry are familiar with this technique. In the second method, factors are changed simultaneously. This is called factorial study. If the change is between two levels it is called a two-level factorial study. This method has an advantage over the OFAT method, where complete design space is covered during experimentation. Therefore factorial study is more effective than the traditional OFAT method. In OFAT replicate runs must be done to get the equivalent power as factorial study. The result for a two factor two level study is that to get the same equivalent power, OFAT required six runs versus only four for the two-level factorial design.

The advantage of factorial design becomes more pronounced as more factors are added. In a three factorial study, the factorial design requires only eight runs versus 16 for an OFAT experiment with equivalent power⁹. In both designs, the effect estimates are based on an average of four runs each: right to left, top to bottom and back to front design is now twice for OFAT for equivalent power. In this case the OFAT design covers twice as many data points compared to factorial design to gain the same power. The relative efficiency of the factorial method increases with every additional factor. The factorial design has another two advantages; it covers a broader volume of design space, and it reveals interaction between factors. Figure 1 shows a summary of the technique. When carrying out experiments, it is important to minimise the effect from uncontrolled factors by randomization. The only downside of this method is that the experimenter as mentioned earlier should have a basic understanding of the behaviour of the system i.e. to select the factors and corresponding response for the experimentation. This method becomes important in screening of factors in optimization of a set of properties. This is done by a response surface method which will be discussed subsequently.

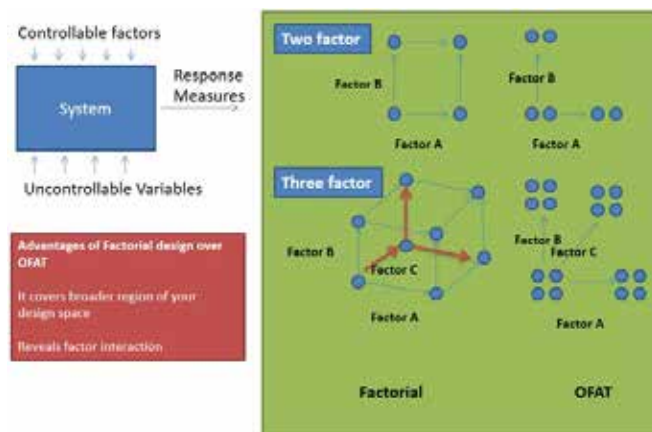


Figure 1 Two level factorial method. Number of experiments comparison between OFAT and two level factorial method⁵

Example of a two level factorial study

In this example a set of data obtained for rubber compounds¹⁰ consist of five ingredients were investigated. The five ingredients are CBS (cyclohexyl benzothiozyl sulphenamide), thiuram disulfide, silica, naphthenic oil and sulphur. In these compounds silica, naphthenic oil and sulphur had been varied and CBS and thiuram disulfide had been kept at 1.2 and 0.6 phr respectively. The final vulcanised films had been tested for range of properties but only tensile strength and 300% modulus are used for the purpose of this example. Table 1 summarises the experimental changes. The overall object of this exercise is to identify factors that are sensitive to a selected property.

Table 1: Factor changes

Factor	Name	Units	Low level (-1)	High level (+1)
A	Sulphur	phr	2	3
B	Silica	phr	30	50
C	Naphthenic oil	phr	6	18

As mentioned previously, full factorial experiments are conducted using all combinations of factors at all levels. Since this is a two level factorial 2^3 experiments are carried out. The following Table 2 shows the results for experiments carried out. Standard order gives how the experiments are performed to minimise noise.

Table 2: Results from the experiment

Standard Order	Run	A: Sulphur (phr)	B: Silica (phr)	C: Naphthenic oil (phr)	Y1: Tensile Strength (MPa)	Y2: 300% Modulus (MPa)
2	1	3	30	6	9.41	2.44
6	2	3	30	18	8.13	1.64
8	3	3	50	18	13.68	1.73
1	4	2	30	6	14.14	2.84
5	5	2	30	18	11.48	1.23
3	6	2	50	6	14.8	2.38
7	7	2	50	18	14.41	1.22
4	8	3	50	6	17.13	2.7

In the first investigation let us analyse the factors that contribute to tensile strength. Using the results from the experiments full factorial main effects and interaction effects are arranged for each response. The

following Table 3 gives all effects on Tensile Strength for each type of factor. Since a log transformation is used the tensile is converted to its log values.

Table 3 Effects for each type of factors

Standard	Main Effects			Interaction Effects				Response
	A	B	C	AB	AC	BC	ABC	Ln(Y1)
1	-	-	-	+	+	+	-	2.65
2	+	-	-	-	-	+	+	2.24
3	-	+	-	-	+	-	+	2.70
4	+	+	-	+	-	-	-	2.84
5	-	-	+	+	-	-	+	2.44
6	+	-	+	-	+	-	-	2.10
7	-	+	+	-	-	+	-	2.67
8	+	+	+	+	+	+	+	2.62
Effect	-0.165	0.348	0.152	0.212	-0.034	0.026	-0.065	2.54

Effects are calculated by averaging the values of the response for (+) and deducting the average values of the response for (-). Then the absolute values of the effects are arranged in an ascending order and plot them on an equally spaced cumulative half normal probability chart. If the effects are random then they will fall on a straight line passing through the origin. However, if the effects that are not random will fall outside the line as seen from Figure 2. Each factor or factor combinations that falls outside the line will contribute to the response compared to the others. Therefore to build a mathematical model for this response only these factors are selected. Sometimes depending on the data that is subject to analysis, it will be necessary to use mathematical transformation for better interpretation^{2,5,11}. In the present case as mentioned previously a logarithmic transformation is taken.

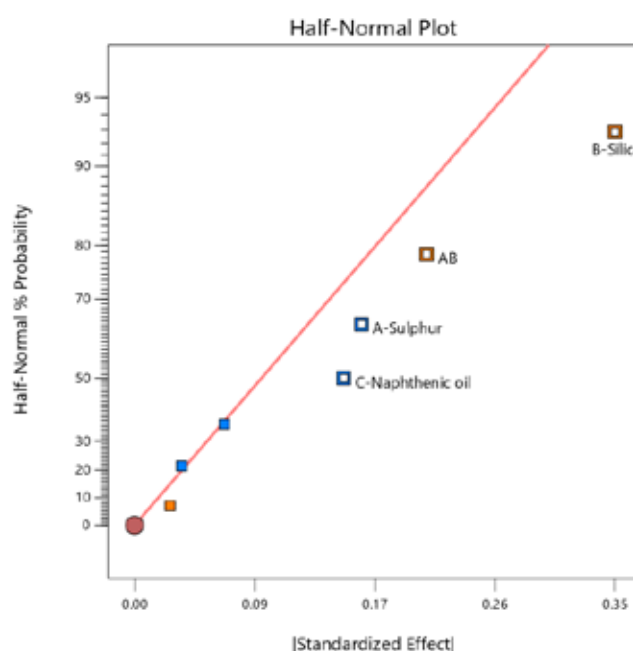


Figure 2. Half normal plot for effects for Tensile Strength

From the half normal plot it can be seen that there are three points that fall on the straight line while A, B, C and AB that fall outside. The effects that fall on the straight line are called trivial effects and they will be used as residual effects in the analysis of variance (ANOVA).

Similarly the same procedure can be done for response 300% modulus. The half normal plot for 300% modulus is shown in Figure 3.

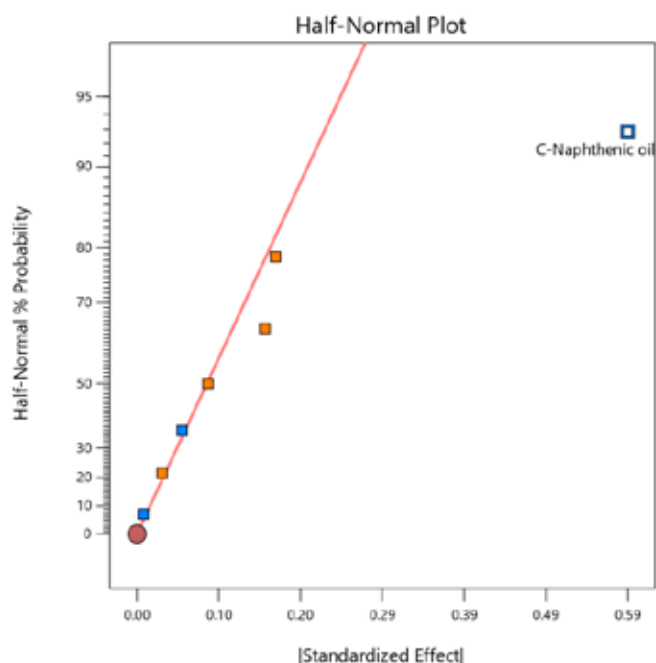


Figure 3. Half normal plot for 300% modulus.

From the plot for 300% modulus only Factor C (naphthenic oil) contributes. It can be seen that from the two analysis each response is sensitive to different factors.

Factor Interaction plot

One of the advantages in factorial design is the factor interaction. It can be seen from Figure 2 the effect AB is significant which indicates an interaction between sulphur (A) and silica (B). Table 4 summarises the data for interaction of sulphur (A) and silica (B) when naphthenic oil content at 12 phr.

Table 4. Summary of factor interactions for tensile strength.

Standard	Sulphur(A)	Silica (B)	Tensile Strength MPa (Y,Avg.)
1	-	-	12.77
3	-	+	14.63
5	+	-	8.72
7	+	+	15.34

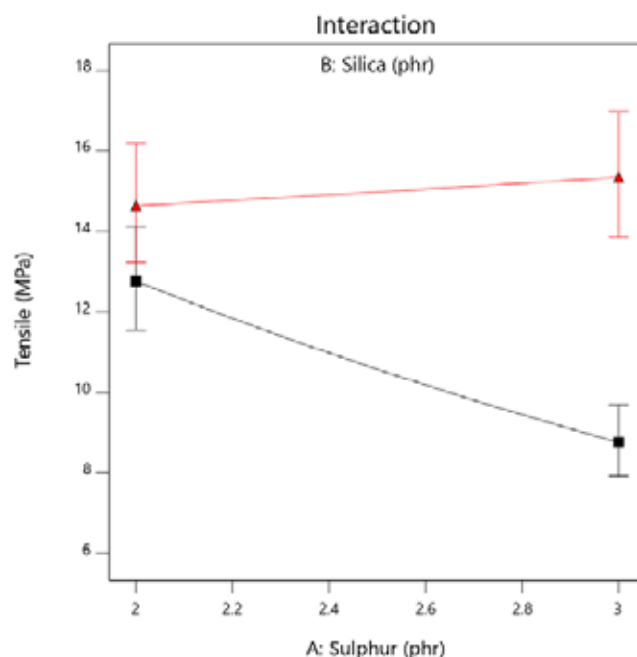


Figure 4. Factor interaction for Tensile Strength

The interaction plot (Figure 4) indicates that with increasing sulphur the tensile strength is reduced at a lower silica loading (black). This can be considered a breakthrough. Similar effects are claimed in literature¹². This information can be only obtained by considering simultaneous changes to the factors during experimentation. OFAT does not give such information.

Analysis of Variance (ANOVA)

Calculating the analysis of variance (ANOVA) for tensile strength using two level factorial is shown in Table 5. The sums of squares is calculated⁵ as follows:

$$SS = N/4 * \text{Effects}^2$$

Where N is the number of runs, in this experiment it is 8.

$$SS_{\text{Model}} = SS_A + SS_B + SS_C + SS_{AB}$$

$$SS_{\text{Residual}} = SS_{AC} + SS_{BC} + SS_{ABC}$$

$$\text{Corrected Total} = SS_A + SS_B + SS_C + SS_{AB} + SS_{\text{Residual}} \text{ (the total sum of squares (SS) corrected for the mean)}$$

Table 5. ANOVA for Tensile Strength

Source	Sum of Squares (SS)	Df	Mean Square (MS)	F Value	Prob > F
Model	0.4320	4	0.1080	26.75	0.0111
A-Sulphur	0.0541	1	0.0541	13.41	0.0352
B-Silica	0.2423	1	0.2423	60.02	0.0045
C-Naphthenic oil	0.0459	1	0.0459	11.38	0.0433
AB	0.0896	1	0.0896	22.18	0.0181
Residual	0.0121	3	0.0040	-	-
Corrected Total	0.4441	7	-	-	-

The F-Value is calculated as the mean squares of the factor against the mean squares of the residuals. For the model (MS Model/MS Residual)⁵ F-Value is 26.75. Once the F-Values are calculated for appropriate factors it is checked from the F-Value table of 95% confidence. When looking at the values in the table the Df (degrees of freedom) of the numerator (top) is taken as the column and the Df of the denominator (bottom) is taken as the row. If the value in the table is less than the value calculated then the effect for that factor is significant. Using this method it is possible to confirm the selected factors are correct. The procedure is clearly shown in Design Expert software¹³. Once the factors for the model is selected a mathematical relation can be built

using regression method for a selected response.

For the tensile strength analysis mathematical relation will be as follows:

$$\text{Ln (tensile)} = 4.5136 - 1.01101A - 0.035501B - 0.012639C + 0.021162AB$$

From this equation predicted values for Tensile Strength response can be calculated. Since a log transformation is used, the tensile is calculated as Ln (Tensile). Table 6 summarised the calculated values. Since we used the log transformation the residuals from the model are calculated by taking the difference between the log values of the actual and predicted values.

Table 6. Prediction Vs actual values for tensile strength

	A	B	C	AB	TS actual MPa	TS predict MPa	Resid Ln(Pred.)-Ln(Act.)
1	-	-	-	+	14.14	13.75	-0.028
2	+	-	-	-	9.41	9.44	0.003
3	-	+	-	-	14.8	15.75	0.062
4	+	+	-	+	17.13	16.52	-0.036
5	-	-	+	+	11.48	11.81	0.028
6	+	-	+	-	8.13	8.11	-0.002
7	-	+	+	-	14.41	13.54	-0.062
8	+	+	+	+	13.68	14.19	0.037

Figure 5 shows the half normal plot for the residual.

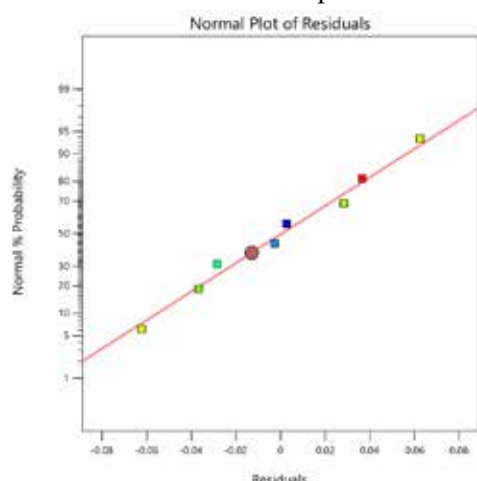


Figure 5. Half normal plot for Tensile strength residuals

From the Figure 5 it can be seen that the residuals fall along a line in a half normal plot which means the residues are normally distributed. This further confirms the analysis is acceptable.

Fractional Factorial

In our previous investigation only three factors were used. However, in most of the investigations it will be necessary to use more than 3 factors. In two level full factorials the number of experiments will be given by the formula 2^k where k is the number of factors that are under investigation. It can be seen that beyond four factors the number of experiments for full factorial run becomes not practical. To overcome this problem statisticians have come out with a method called

fractional factorial where fewer number of experiments are conducted for the analysis. Fractional factorial consists of a set of carefully chosen experimental runs from a full factorial. The set of experimental runs are selected so as to exploit the most important factors of the problem under the study with a fraction of an effort of full factorial in terms of experiments and resources. Due to the fewer number of experiments in fractional

factorial design it estimates the main effects and the two factor interactions only, but not the three factor or higher order interactions². The following Table 6 shows the number of experiments that are necessary for these runs⁵. Fractional factorial is mainly used in screening of factors when investigating a situation where there are higher number of selected factors. With this method it is possible to select a vital few from trivial many factors.

Table 6. Fractional factorial experiments⁵

Factors	Main effects and Two factor interactions	Full factorial	Fraction for Screening	Savings
4	10	16	12	25%
5	15	32	16	50%
6	21	64	32	50%
7	28	128	32	75%

Since the problem is investigated with a smaller number of experiments, fractional factorial has to be handled carefully. Details of the analysis for fractional factorial can be found in statistical references^{1,5}.

Response surface methodology

In the previous section two level factorial methods were explained to select factors that will be sensitive to a selected response (property) of a product. Once these factors are selected it will be necessary to find the best combination of the factors to give the optimum response. However, this cannot be done with existing information since the information obtained from two

level factorials give mainly linear relationship between factors to a selected response. Therefore it is necessary to use more data points in the design space to find the curvature of the surface. There are many methods¹⁴ to obtain these additional data points. All these methods will lead to map the curvature of the surface to find the optimum value. The flow chart¹⁵ to the procedure is shown in Figure 6. In this additional data points to the two level data are done by using a central composite design space. This includes data point at the centre point plus additional points away from the centre. Central composite design space for a three factorial experiment is shown as in Figure 7.

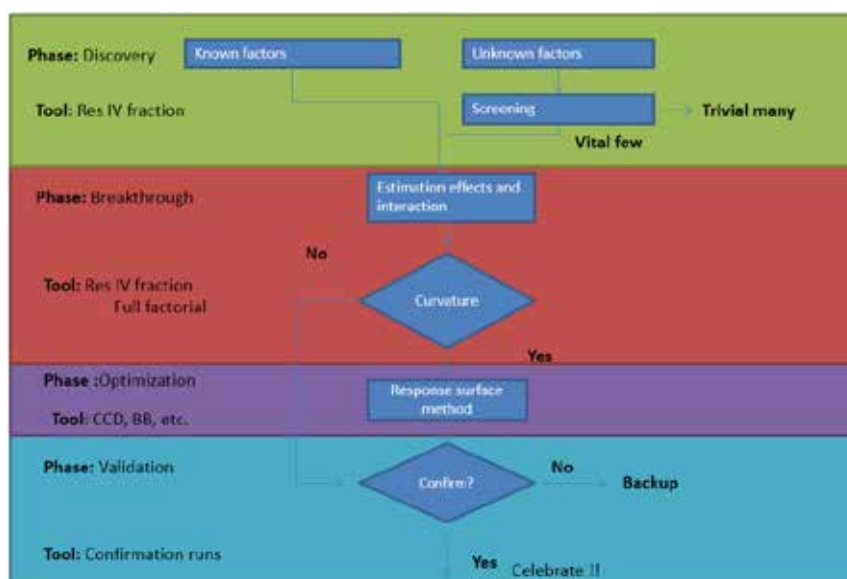


Figure 6. Flow diagram for response surface analysis

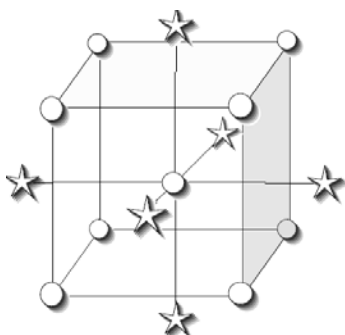


Figure 7. Three factor central composite design.

Using the additional data points experiments are conducted for a selected response. Then sensitive factors and factor combinations for a selected response are determined as before using F-test. Using these factors and factor combinations a mathematical equation is derived to generate the response surface. These equations can be linear, two factor interaction (2FI) or quadratic. Similarly the best relation is found for all other responses. Then optimization can be done for a desirable combination of responses (properties) to find the corresponding all factors (ingredients). Since these operations need many mathematical procedures, computational methods are used.

Response surface for each property is developed using additional data points to what were used in the factorial study. Table 7 shows the design space that was generated with Design Expert software for the set of experiments to determine the response surface. It can be seen in Table 7 that there are experiments that are replicated to determine the pure errors to include in the calculation.

Table 8 shows the set of properties that were obtained for the rubber compounds with different combinations of factor input as in Table 7.

Table 7 Design space of the experiments conducted

Experiments	Sulphur phr	Silica phr	Naphane oil phr
1	2	30	18
2	2.5	40	12
3	2	30	6
4	3	50	6
5	3.34	40	12
6	2.5	40	12
7	2.5	40	22.09
8	2.5	40	12
9	3	50	18
10	2.5	40	12
11	2.5	40	1.91
12	2	50	6
13	2.5	56.82	12
14	1.66	40	12
15	2.5	23.18	12
16	2.5	40	12
17	3	30	18
18	3	30	6
19	2.5	40	12
20	2	50	18

Table 8 Data on difference properties for the experiments in Table 7

Experiments	Cure_ Time min.	Mooney Scorch min.	300% Modulus MPa	Tensile MPa	Elongation %	Hardness IRHD	Die_ tear kN/m
1	16.35	32.72	1.23	11.48	740.2	45.97	17.11
2	21.34	33.57	1.65	14.48	721.2	53.57	29.35
3	16.66	31.83	2.84	14.14	624.61	51.03	24.35
4	24.22	37.14	2.70	17.13	753.17	67.23	39.88
5	23.5	40.52	1.66	14.57	830.9	51.19	25.22
6	20.4	34.54	2.00	14.48	728.84	53.02	31.36
7	26.78	32.03	1.26	9.77	811.12	49.63	20.18
8	20.62	33.44	1.70	14.50	739.16	52.72	31.62
9	33.48	35.67	1.73	13.68	847.07	54.15	39.68
10	19.9	34.93	1.73	14.34	723.3	52.93	29.72
11	21.49	36.00	3.55	14.78	610	55.32	28.30
12	23.58	23.70	2.38	14.80	804.1	63.41	39.09
13	25.56	25.90	2.35	15.64	841.61	65.39	47.44
14	20.31	22.52	2.84	14.76	682.57	50.53	33.44
15	14.56	32.89	2.06	7.20	524.39	48.52	12.75
16	19.9	34.93	1.73	14.34	723.3	52.93	29.72
17	19.21	35.01	1.64	8.13	736.909	49.69	15.78

18	17.63	39.17	2.44	9.41	481.96	55.4	21.99
19	21.34	33.57	1.65	14.48	721.21	53.57	29.36
20	24.96	22.23	1.22	14.41	933.86	51.66	28.75

Design-Expert® Software
Factor Coding: Actual
Original Scale

Tensile (MPa)
○ Design points below predicted value
7.20177 17.1272

X1 = A: Sulphur

X2 = B: Silica

Actual Factor
C: Naphthenic oil = 12

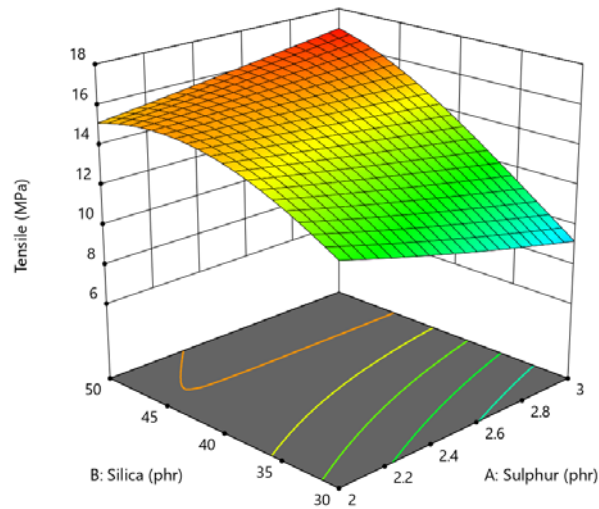


Figure 8. Response surface

Figure 8 shows the response surface for tensile strength with the changes to Silica and Sulphur. This gives a deeper understanding of the recipe.

An optimization for the recipe was done by selecting the properties with desirability i.e. cure_time - minimum, Mooney scorch - maximum, tensile - maximum, elongation - maximum, hardness - optimum, tear

-maximum. The Figure 9 shows the amount of sulphur, silica and naphthenic oil for this condition. Figure 10 shows the contour plot of the optimization. Yellow region shows the area where it satisfies the desired properties. Since the study covers a wide design space it will be possible to make changes in a very short time frame to the product's properties, depending on market preference.

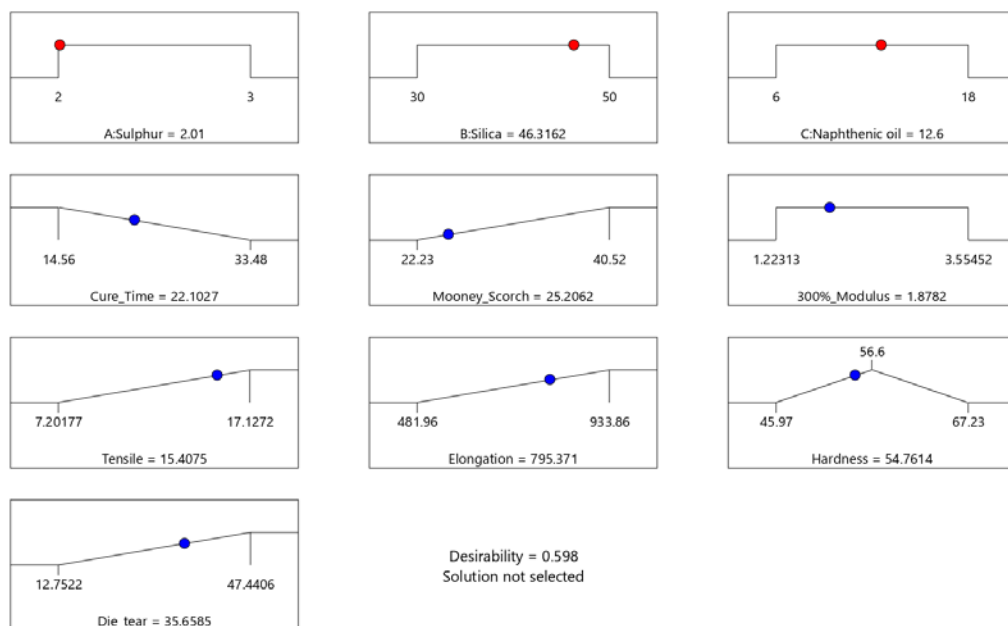


Figure 9 Design Expert optimization output.

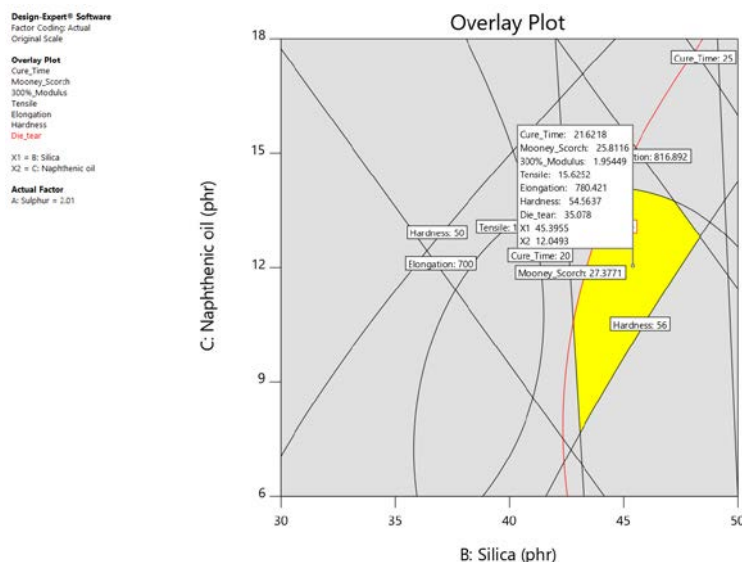


Figure 10 contour plot of the optimization.

In rubber compound recipe, ingredients are used as amounts per hundred parts of rubber. Therefore with such information it will be possible to carry out an analysis as discussed above. Mixture designs² are not generally used in this type of experiments because ingredients in a rubber compound recipe are not considered as proportions.

Although the above discussion was on a software based on statistical methods, software based on neural networks⁶ and multiple linear iteration methods⁷ can be used to predict new formulations. These methods will be an advantage for companies where there are lots of historical data available. These will be discussed in a separate paper.

Conclusion

Design of Experiments is an efficient method to use in development of rubber compounds in product manufacture. Using this method it will be possible to understand how each ingredient in the recipe affects the final properties of the vulcanised product. Using commercial software available today, it has become very effective to optimise a recipe to give the best performance in the final product. This technique will allow a chemist to arrive at a recipe much faster with desired properties. Therefore the time to market a product will be reduced dramatically.

References

- (1) Montgomery, D. C. Design and Analysis of Experiments; Wiley, 2012.
- (2) Tranter, R. L. Design and Analysis in Chemical Research; Sheffield Academic Press, 2000.
- (3) ASTM-E1325-16; ASTM International: West Conshohocken, Pennsylvania 19428-2959, USA, 2016.
- (4) Myers, R. H.; Montgomery, D. C.; Anderson-Cook, C. M. Response Surface Methodology: Process and Product Optimization Using Designed Experiments; Wiley, 2009.
- (5) Anderson, M. J.; Whitcomb, P. J. DOE Simplified: Practical Tools for Effective Experimentation; Taylor & Francis, 2000.
- (6) Roberts, R. J.; Rowe, R. Intelligent Software for Product Formulation; Taylor & Francis, 1998.
- (7) Graf, H. J. In Rubber Fibre Plastic Int. ; Dr. Gupta Verlag: Germany, 2013; Vol. 08, p 2
- (8) Weissman, S. A.; Anderson, N. G. Organic Process Research and Development 2015, 19, 1605.
- (9) Ledolter, J.; Kardon, R. H. Investigative Ophthalmology & Visual Science 2020, 61.
- (10) CAD/Chem; www.aiware.com: 1997.
- (11) Lee, D. K. Korean J Anesthesiol 2020, 73, 503.
- (12) Han, S.; Gu, B.; Kim, S.; Kim, S.; Mun, D.; Morita, K.; Kim, D.; Kim, W. Polymers 2020, 12, 2815.
- (13) Design-Expert www.statease.com, 2013.
- (14) Myers, R. H.; Montgomery, D. C.; Anderson-Cook, C. M. Response Surface Methodology: Process and Product Optimization Using Designed Experiments; Wiley, 2016.
- (15) Anderson, M. J.; Whitcomb, P. J. RSM Simplified: Optimizing Processes Using Response Surface Methods for Design of Experiments; Taylor & Francis, 2005.