PROJECT REPORT

An experimental Design and Analysis Project aimed at understanding the significant factors in the manufacturing process of Optical Fiber and find optimal settings to minimize attenuation

COURSE: ST 516

Rucha Girgaonkar
Chaitanya Rajeev
Vishakha Patil

Table of Contents

Note:

For the purpose of simplicity in analysis, we will be referring to every factor in the process whenever required using the following alphabetical notation:

S.No.	Factor	Factor Type	Notation	Range/Levels	Unit
1	Fiber draw speed	Continuous	A	20 - 30	m/sec
2	Furnace temperature	Continuous	B	1800 - 2200	Celcius
3	Draw tension	Continuous	C	0.5 - 1.0	N
4	Germanium concentration	Continuous	D	0.01 - 0.05	N/A
5	Fiber design	Categorical	E	1,2	N/A
6	Draw tower	Categorical	F	1,2	N/A
7	Raw material supplier	Categorical	G	1,2	N/A
8	Coating type	Categorical	H	1,2	N/A

1 Executive Summary

This project entails a detailed analysis of the optical fiber manufacturing process conducted using the principles of experimental Design in the JMP software package. We were provided with 8 factors that could potentially affect the attenuation of light. The objective was to minimize the response (attenuation of light) and identify significant factors and their nature of variation in the process.

We experimented with fractional factorial designs, folded designs and augmented designs with center and axial points to identify the significant effects. After our experiments we concluded that Furnace Temperature, Germanium concentration and Fiber Design were the significant factors. Few interactions between these factors were also significant. Both Furnace Temperature and Germanium Concentration showed quadratic variation with the response. The optimum response of our model was 0.397666.

2 Introduction

An optical fiber is a single, hair-fine filament drawn from molten silica glass. One of the most important attributes in the design of optical fiber is optical signal attenuation. Optical signal attenuation is the measure of light lost between input and output. The primary aim of this project is to minimize the optical signal attenuation of the fire to maximize the light carrying capacity of the fiber.

An experimental design is used when the most rigorous of all research design is to be obtained. Thus, the approach which was followed was to design an experiment to be run to determine the best combination of settings for the fiber drawing process and to minimize attenuation. A total of 8 factors are identified.

The project has two objectives:

1. Identify any significant main effects and interactions on attenuation from the list of given factors.
2. Build a model to predict attenuation from the significant factors and generate a list of optimal settings based on that model.

It is also important to determine optimal settings in a minimum number of runs, as each run costs $1000, and the budget assigned for the experiments is $50,000.

3 Experimental Design

The general philosophy and series of steps to decide the experimental designs were as follows:

i. Fit a fractional factorial experiment with full factorial combinations of just the continuous factors.
ii. Decide which folded design to run, in order to de-alias the effects that are significant in the first design.
iii. Obtain centre points and axial points to check if the response function has quadratic variations.
iv. Perform the steepest descent runs only IF the optimum response(minima) is outside the predictor space (outside factor settings range).

3.1 Description

3.1.1 The 2^{8-4}_{IV} Fractional Factorial design:

This design has a good resolution and allows for full factorial combinations of 4 factors. Hence, we chose this model with full factorial runs of the continuous factors A, B, C, D. We utilized 16 runs for this design. The generating functions are:

- E = BCD
- F = ACD
- G = ABC
- H = ABD

3.1.2 The Folded Design:

Performing the first experiment, we found that interactions BD and DE were significant along with main effects B, D and E. But BD was aliased with AH, CE, FG, and DE were aliased with AG, BC and FH. Hence the aim of our folded experiment was precisely to de-alias and isolate these 2 effects We utilized another 16 runs for this experiment.

After tweaking the generating functions, we found that the exact fold that would achieve the required de-aliasing was generated by the functions:

- E = -BCD
- F = -ACD
- G = ABC
- H = -ABD

3.1.3 Augmented Design with Center and Axial Points

Once we de-aliased the required effects, we had to check whether there were quadratic variations within our response function. The next step forward was to get experimental data on center points and axial points in the predictor space of the 2 largest continuous main effects and prepare an augmented design. We would try to fit a first or second-order model on this data. We utilized 8 runs consisting of 4 center points and 4 axial points

3.1.4 Steepest Descent Design (if required)

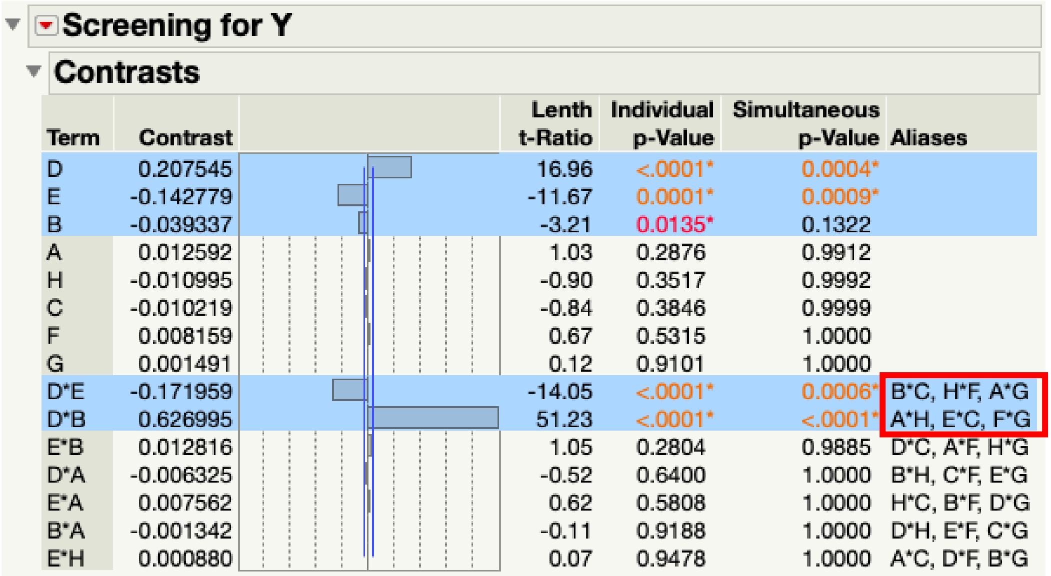
If the response surface shows signs that the optimum could lie outside the predictor ranges, then the steepest descent design would be needed to reach the vicinity of the optimum.

4 Analysis and Results

This section will be presented in a similar structure to the Experimental Design section, aligning with our general philosophy and steps.

4.1 The 2^{8-4}_{IV} Fractional Factorial Design (16 runs):

Since we had 8 factors that could be of consequence, we decided to fit a 2^{8-4} design, choosing the 4 continuous factors (fiber draw speed (A), furnace temperature (B), draw speed (C), germanium concentration (D)) and fitting a resolution IV model with 16 runs. This design was chosen since it allowed us to carry out full factorial runs of the 4 continuous factors A, B, C and D. The following image is the screening done for this design:



Screening for Y

Contrasts

Term	Contrast	Lenth t-Ratio	Individual p-Value	Simultaneous p-Value	Aliases
D	0.207545	16.96	<.0001*	0.0004*	
E	-0.142779	-11.67	0.0001*	0.0009*	
B	-0.039337	-3.21	0.0135*	0.1322	
A	0.012592	1.03	0.2876	0.9912	
H	-0.010995	-0.90	0.3517	0.9992	
C	-0.010219	-0.84	0.3846	0.9999	
F	0.008159	0.67	0.5315	1.0000	
G	0.001491	0.12	0.9101	1.0000	
D*E	-0.171959	-14.05	<.0001*	0.0006*	B*C, H*F, A*G
D*B	0.626995	51.23	<.0001*	<.0001*	A*H, E*C, F*G
E*B	0.012816	1.05	0.2804	0.9885	D*C, A*F, H*G
D*A	-0.006325	-0.52	0.6400	1.0000	B*H, C*F, E*G
E*A	0.007562	0.62	0.5808	1.0000	H*C, B*F, D*G
B*A	-0.001342	-0.11	0.9188	1.0000	D*H, E*F, C*G
E*H	0.000880	0.07	0.9478	1.0000	A*C, D*F, B*G

This model showed us that amongst the continuous effects, factors B and D (furnace temperature and germanium concentration) were the significant main effects; and among the categorical factors, factor E (design type 1 or 2) was the significant main effect. The effects B, D, E are aliased with 3-way interactions. As 3-way interactions tend not to be significant according to the sparsity of effects principle, when aliased with single factors, we can conclude that the effects B, D, and E are significant.

However, the 2-way interactions D*B and D*E are also shown as significant. But they are aliased with other 2-way interactions out of which B*C and E*C, which also contain significant main effects, could be significant. Hence, we can't be sure that D*B and D*E are significant in themselves. Therefore, we will require de-aliasing to isolate the effects.

4.2 The folded design (16 runs):

Aim: To isolate and de-alias effects DE and BD by selecting the correct fold												
Sr. No.	Effects Aliased w DE				Effects Aliased w BD				Generating Functions			
	AG	BC	DE	FH	AH	BD	CE	FG	E=BCD	F=ACD	G=ABC	H=ABD
1	+	+	+	+	+	+	+	+	+	+	+	+
2	+	+	-	+	+	+	-	+	-	+	+	+
3	+	+	+	-	+	+	+	-	+	-	+	+
4	+	-	+	-	+	-	+	-	-	-	-	-
5	+	+	-	+	+	-	+	+	-	-	+	-

Our aim now was to de-alias effects BD and DE which we found to be significant previously. The above table shows the working to determine the sign of aliases obtained from trying different +/- combinations for the generating functions E, F, G, and H.

We needed a combination which yielded opposite signs to the first design (Sr. No. 1) only for effects BD and DE. We can see that the 5th combination (Sr. No. 5) yields just that. Hence, we chose this fold. The fold is generated by generating functions E = -BCD, F = -ACD, G = ABC, H = -ABD

Screening for Y

Contrasts

Term	Contrast	Lenth t-Ratio	Individual p-Value	Simultaneous p-Value	Aliases
Block	-0.001254	-0.24	0.8169	1.0000	
D	0.209580	40.11	<.0001*	<.0001*	
E	-0.148004	-28.33	<.0001*	<.0001*	B*C*A, E*C*H, E*A*F, B*H*F
B	-0.061673	-11.80	<.0001*	<.0001*	D*C*H, B*A*H, B*C*F, D*A*F
C	-0.008482	-1.62	0.1164	0.9159	D*C*A, E*A*H, E*C*F, D*H*F
A	0.007639	1.46	0.1532	0.9702	D*B*A, D*E*H, E*B*F, A*H*F
H	0.003005	0.58	0.5849	1.0000	D*B*C, E*B*H, D*E*F, C*H*F
F	-0.001232	-0.24	0.8204	1.0000	D*E*C, E*B*A, D*B*F, C*A*F
G	-0.000498	-0.10	0.9264	1.0000	E*B*C, D*E*A, D*B*H, C*A*H
D*E	-0.187138	-35.82	<.0001*	<.0001*	
D*B	0.601553	115.13	<.0001*	<.0001*	
E*B	0.007529	1.44	0.1573	0.9749	C*H, A*F
D*C	0.005287	1.01	0.3023	1.0000	

Effect Summary

Source	LogWorth	PValue
B*D	29.851	0.00000
D	18.171	0.00000 ^
D*E	16.955	0.00000
E	14.489	0.00000 ^
B	6.517	0.00000 ^

Remove Add Edit FDR ('^' denotes effects with containing effects above them)

Summary of Fit

RSquare	0.995469
RSquare Adj	0.994598
Root Mean Square Error	0.051121
Mean of Response	1.830624
Observations (or Sum Wgts)	32

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	14.928607	2.98572	1142.478
Error	26	0.067948	0.00261	**Prob > F**
C. Total	31	14.996555		<.0001*

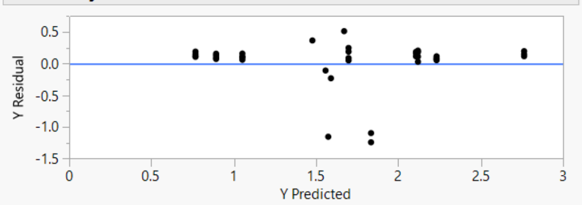
The folded design data was added to the previous fractional factorial data for a total of 32 data points. The same model was fit on this data. The model was statistically significant. The results show that BD and DE remain significant and are now de-aliased. This confirms that the effects of the significant interactions are indeed BD and DE.

4.3 Augmented design with the center and axial points (8 runs):

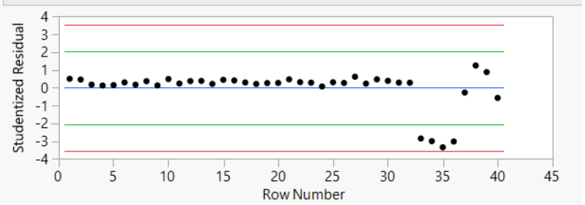
Even though the fitted reduced model was statistically significant, in order to capture any quadratic variations between effects, one needs to fit data for the center and axial points in the predictor space. Since, in our reduced model, B and D are the only two continuous factors, our predictor space is essentially 2-D. Hence, we can fit a spherical Central Composite Design (CCD) with the center and axial points.

We added the center and axial points data to the data table and re-ran the previous model while keeping the settings of the categorical factors at levels which minimized the attenuation response; for E this was level 1. We used 8 runs for the points.

Residual by Predicted Plot



Studentized Residuals



Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	8	6.3826630	0.797833	270.7933
Pure Error	26	0.0766033	0.002946	Prob > F
Total Error	34	6.4592663		<.0001*
				Max RSq
				0.9964

Residual Analysis and Lack of Fit test in 1st order model for Augmented Design with center/axial points.

From the above plots and tests, we can conclude that the linear model was insufficient to explain the new data which confirmed the presence of polynomial variation within the effects. The residuals and lack-of-fit tests for the model were analysed to make this conclusion.

Owing to this, we re-fit the model adding squared predictors, moving forward. The detailed analysis and reports of all our models, along with the final model are described below:

Effect Summary

Source	LogWorth	PValue
D(0.01,0.05)	33.637	0.00000
D*B	33.635	0.00000
B*B	27.620	0.00000
B(1800,2200)	27.620	0.00000 ^
D*D	21.156	0.00000
D*E	18.989	0.00000
E	16.379	0.00000 ^

Remove Add Edit ☐ FDR ('^' denotes effects with containing effects above them)

Model significance reports for the 2nd Order Model

Summary of Fit

RSquare	0.995121
RSquare Adj	0.994053
Root Mean Square Error	0.057118
Mean of Response	1.689638
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	21.291327	3.04162	932.3032
Error	32	0.104399	0.00326	Prob > F
C. Total	39	21.395726		<.0001*

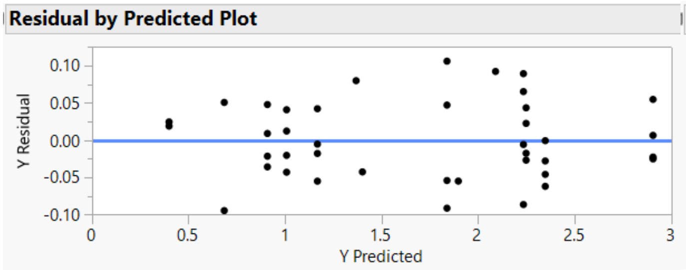
Lack Of Fit

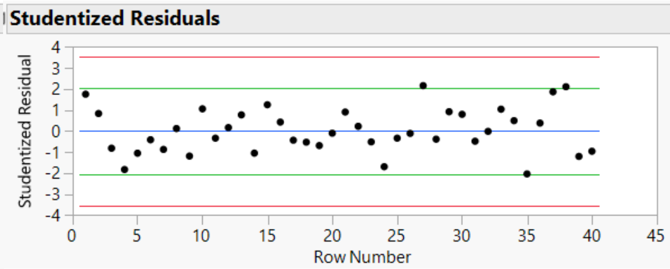
Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	6	0.02779599	0.004633	1.5724
Pure Error	26	0.07660328	0.002946	Prob > F
Total Error	32	0.10439927		0.1950
				Max RSq
				0.9964

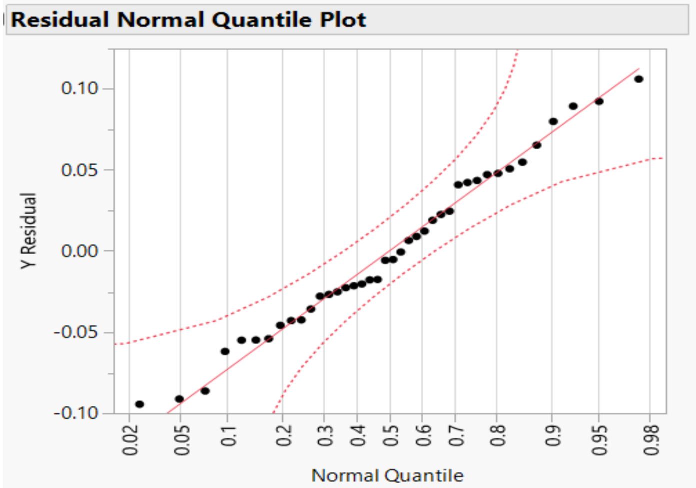
From the above reports, we can clearly conclude that the squared effects are highly significant, and the resulting model is a very good fit. It has an F-value of 932.3 with a p-value of <0.0001. The lack of fit has significantly improved from the 1st Order model and is now insignificant.

4.3.1 Model Adequacy

Before we move forward with using this model for predictions, it is imperative to conduct model adequacy checks using appropriate plots.

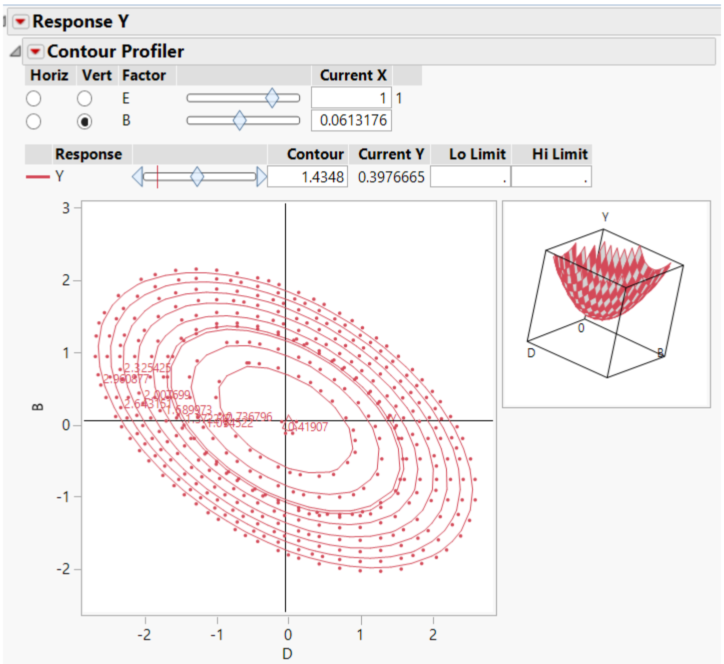




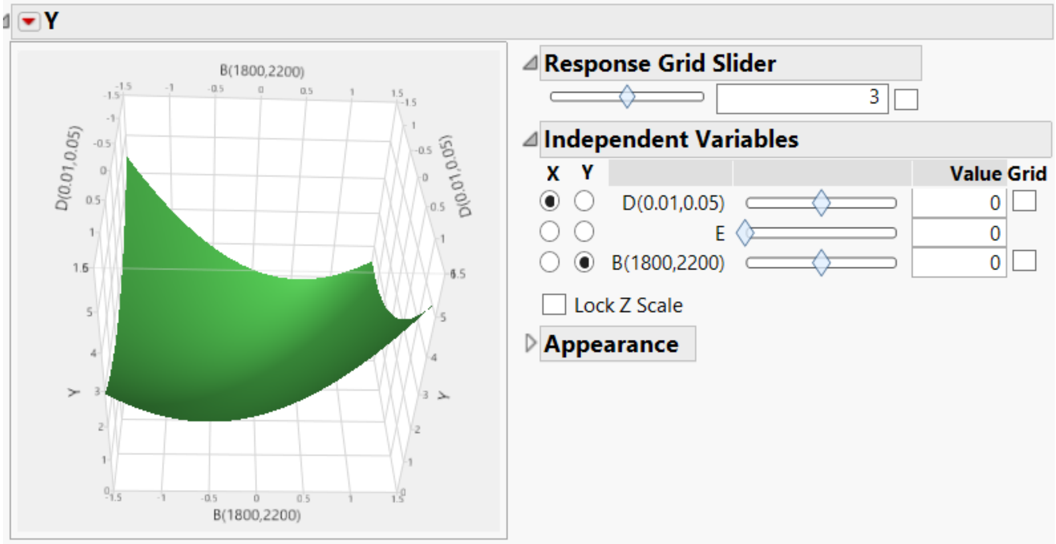


- The residual plots look adequate.
- The residuals have constant variance.
- The residuals are normally distributed according to the Normal Quantile Plot.
- Hence model adequacy is maintained.

4.3.2 Contour Profiles and Response Surface



- The contour plots hint at the optimum somewhere near the center.
- The corners seem to increase the response value.

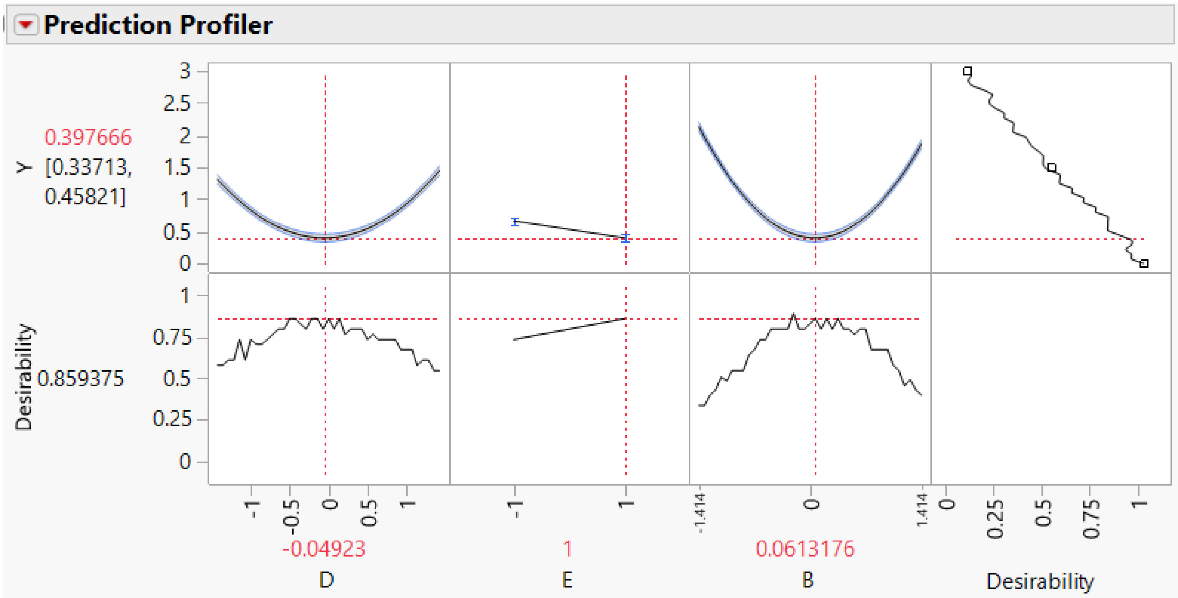


The contour profiles and response surface plots show that the optimum (minimum attenuation) is within the predictor space somewhere close to the center. The response function seems to be bowl-shaped.

This means that we will not need to conduct steepest descent runs to reach the vicinity of the optimum. Hence in the next step, we use the prediction profiler to find the optimum attenuation.

4.3.3 Prediction Profiler and Optimization

The prediction profiler finds the minimum value of y = 0.397666 at D = -0.04923 and B = 0.0613176. The E factor which is categorical is set to 1.



4.4 Notes

Note: Handling Blocks (Days)

In our analysis, we considered the blocks as different data requests which happened on different days. Since we use 4 data requests for our final model, Block is a categorical factor with 4 levels. We modelled block as a random effect factor in our final second-order model to check its significance.

Random Effect Predictions

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Wald p-Value	Pct of Total
Block	172.48891	0.611199	0.5191265	-0.40627	1.6286682	0.2391	99.424
Residual		0.0035434	0.0009327	0.0022454	0.0064135		0.576
Total		0.6147424	0.5191092	0.1919882	9.8765179		100.000

-2 LogLikelihood = -49.2083758

Note: Total is the sum of the positive variance components.

Total including negative estimates = 0.6147424

	Block
24	2
25	2
26	2
27	2
28	2
29	2
30	2
31	2
32	2
33	3
34	3
35	3
36	3
37	4
38	4
39	4
40	4

- The analysis tells us that for our model, the block factor is not significant at the 0.05 level.
- Also, the block factor was confounded with effect B.
- Hence the effect due to the block (days) can be ignored.

Note: Steepest Descent Runs

After fitting the central points, we erroneously obtained a path of descent from a previous response surface to which axial points had not been fit. We used 4 runs for experiments down that path. Hence, we ended up using a total of 44 runs instead of 40.

Highlights

- No. of runs used: 44
- Optimum Attenuation: 0.39766
- Final Model Type: 2nd Order Reduced

5 Conclusions

In order to achieve minimum optimal signal attenuation, we performed an experimental design procedure using fractional factorial, de-aliasing the main effects, using the center and axial points and then built a reduced model to minimize the attenuation. A model using quadratic effects was incorporated because of the presence of curvature in some of the significant terms. The significant main effects, their interactions and second-order model terms were identified using the experimental design. We then continued to find the optimal settings for significant factors to minimize response variable: furnace temperature, germanium concentration and fiber design type while keeping the other insignificant factors to optimal settings as per their parameter estimates in a complete 1st order model.

1) The recommended process settings to run the experiment are summarized in the following table:

Main Factors	**Significant (α = 0.05)**	**Optimal setting**
Fiber draw speed (A)	No	20 m/sec
Furnace temperature (B)	Yes	2012.26 C^{o}
Draw tension (C)	No	1.0 N
Germanium concentration (D)	Yes	0.029
Fiber design 1 or 2 (E)	Yes	Fiber design 2
Draw tower 1 or 2 (F)	No	Draw tower 2
Raw material supplier 1 or 2 (G)	No	Supplier 1
Coating type 1 or 2 (H)	No	Coating type 2

2) Significant Effects:

Effect Summary

Source	LogWorth	PValue
D(0.01,0.05)	33.637	0.00000
D*B	33.635	0.00000
B*B	27.620	0.00000
B(1800,2200)	27.620	0.00000 ^
D*D	21.156	0.00000
D*E	18.989	0.00000
E	16.379	0.00000 ^

Remove Add Edit FDR ('^' denotes effects with containing effects above them)

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	3189350.4	83028.4	38.41	<.0001*
D(0.01,0.05)	24.066798	0.403886	59.59	<.0001*
E[-1]	0.1487045	0.009086	16.37	<.0001*
B(1800,2200)	637880	16605.68	38.41	<.0001*
D*D	0.0001971	8.303e-6	23.73	<.0001*
D*E[-1]	0.003851	0.000192	20.10	<.0001*
D*B	2.4062124	0.040389	59.58	<.0001*
B*B	31894.501	830.2841	38.41	<.0001*

The significant main effects were B, D and E. B and D both have significant quadratic variations as well. The significant interaction effects in the analysis were B*D and D*E.

5.1 Reasons for Recommendations

- The 2nd Order Reduced model is simple and fits the responses well.
- Model has high R^2 = 0.995, high F ratio = 932.3 and is statistically significant
- Model can make an accurate prediction on the optimal response.
- Model is reduced to only include significant effects. Hence it is very interpretable.
- Model is adequate.

5.2 Analysis Limitations

- Highly reduced model. Insignificant factors could become significant outside the current experimental ranges.
- Confounding of Blocks with Higher-order effects in our analysis caused difficulty in interpretation.

5.3 Future Scope of Work

For performing the experimental procedure, we were limited by the number of runs that could be run by the manufacturing department per day. Hence the block size here was 16 runs per day. One could perform analysis on the current model by taking some additional runs at the predicted optimum to confirm our experimental design. This will help improve the precision of estimation of the model even further if required and predict the minimum attenuation level.