# ST 516 Final Project Report

**Team members:** Rezvan Mahdavi Hezaveh, Sajjad Taghyieh, Hossein Tohidi, Mojtaba Sardarmehni, Kimia Vahdat

## **Executive summary**

In this project, we are working for an optical fiber manufacturer and trying to find the best operational specs for minimizing the optical signal attenuation. There are eight operational factors, namely fiber draw speed, furnace temperature, draw tension, germanium concentration, fiber design, draw tower, raw material supplier, and coating type. The first four factors are continuous and the rest of them are categorical. Fifty experiments in a sequential procedure in four days were performed to find the significant factors and develop a model for predicting the optical signal attenuation. Results indicate that coating type, raw material suppliers, draw tower, draw tension and fiber draw speed are not significant while furnace temperature, germanium concentration and fiber design are significant main effects. The significant interactions derived from the optimal model are furnace temperature\*germanium concentration, draw tension\*raw material supplier, germanium concentration\*fiber design and furnace temperature^2. The optimal design suggests furnace temperature, germanium concentration and fiber design should be at 2051.62 C, 0.01 and 1, respectively in order to minimize the optical signal attenuation, which is 0.17. The other factors and their interactions are not significant and can be set to either levels according to economic or safety factors.

#### Introduction

We are hired as statisticians to work for a manufacturer of optical fiber which is going to introduce a new fiber manufacturing design that can minimize the optical signal attenuation. The field experts introduced 8 factors (Fiber draw speed, Furnace temperature, Draw tension, Germanium concentration, Fiber design, Draw tower, Raw material supplier, and Coating type) that they believe might impact the objective function. Our goal in this project is to design an effective experiment (with a minimum number of runs) to determine the significant factors and interactions on the attenuation, and the optimal setting to minimize the attenuation. Given we have only the budget for 50 runs overall (50,000\$ budget), we need to manage the budget, carefully. The rest of this report is as follows: In experimental design section, we will go through the sequential procedure taken to find the significant factors and their interactions. The analysis and results section will cover the final results of our design of experiment. We will state a summary of our experiments and results in the conclusion section along with recommendations to improve the model.

## **Experimental design**

Due to our limited budget (\$50,000) and our relatively large number of factors, (4 continuous and 4 categorical) designing a full factorial DOE will not be feasible. Therefore, we need to carefully design a fractional DOE to represent a full factorial DOE. Note that by reducing the number of runs, fractional DOE will be unable to evaluate the impact of some of the factors (or their interactions) independently, as we might have many confounding interactions (usually higher-order interactions). As shown in the Sparsity of Effects Principle, the higher-order interactions are rare, therefore, we can obtain meaningful results from fractional factorials DOE.

The other approach to reduce the number of runs, is to design screening experiments. Screening results moves us from the point of having no idea about the effect of each factor to a point where some of the significant factors have been realized. The goal here is to identify the most important factors that might minimize the attenuation.

On the other hand, we are told that we can only run 16 times each day which might be a source of unexpected variability in the model. To overcome this issue, we need to block the experiments by day. Then, we can analyze whether or not the day is significant in the final model.

Having said that, two different designs have been considered for the screening experiments:

- 1- 16 runs with one block (resolution IV)
- 2- 32 runs with two blocks (because of the limit of 16 runs per day, resolution IV).

As it is shown in figure below, 16 runs in one day result in aliasing four two-ways interactions which might cause to exceed our budget in the next steps, whereas running 32 runs will decrease the aliasing to couples of two-way interactions for all of them except one, which is aliased with 2 others.

Aliasing of Effects				
Effects		Block Aliases		
Fiber draw speed*Draw tension Fiber draw speed*Germanium concentra Fiber draw speed*Fiber design Fiber draw speed*Raw material supplier Fiber draw speed*Coating type Draw tension*Germanium concentration Draw tension*Raw material supplier	= Fiber design*Coating type  = Fiber design*Raw material supplier  = Draw tension*Coating type = Germanium concentration*Raw material supplier  = Germanium concentration*Fiber design  = Draw tension*Fiber design  = Raw material supplier*Coating type  = Germanium concentration*Coating type			
Aliasing of Effects				
Effects	Aliases			
Fiber draw speed*Furnace temp Fiber draw speed*Draw tension Fiber draw speed*Germanium concentration Fiber draw speed*Fiber design Fiber draw speed*Draw tower Fiber draw speed*Raw material supplier Fiber draw speed*Coating type	= Draw tension*Coating type = Germanium concentration*Raw material supplier = Fiber design*Draw = Furnace temp*Coating type = Germanium concentration*Draw tower = Fiber design*Raw material = Furnace temp*Raw material supplier = Draw tension*Draw tower = Fiber design*Coating type = Furnace temp*Draw tower = Draw tension*Raw material supplier = Germanium concentration*Coa = Furnace temp*Fiber design = Draw tension*Germanium concentration = Raw material supplier*Coi = Furnace temp*Germanium concentration = Draw tension*Fiber design = Draw tower*Coating type = Furnace temp*Draw tension = Germanium concentration*Fiber design = Draw tower*Raw material	supplier iting type ating typ		

Furthermore, since we can only have 16 runs per day, and we wanted to consider the effect of blocking and see if it has a significant effect on the response, we ended up choosing 32 runs within 2 days.

After conducting an analysis on the results obtained from resolution IV and running 32 experiments in 2 days, we observe that block is not significant in the effect tests, where the ANOVA shows that the whole model is significant, and the REML test verifies this observation by showing us that only 11.92% of the total variation is explained with blocks and the confidence interval of blocks are very tight around zero. Aforementioned tables are given below:



In the screening experiments, we tolerate more type I error, so at the  $\alpha=0.2$  level, we removed the insignificant factors. We found 2 significant interactions which were confounded with another interaction. Confounded interactions were as follows:

- Fiber Draw Speed\*Raw Material=Germanium Concentration\*Fiber Design
- Draw Tension\*Raw Material=Germanium Concentration\*Coating type

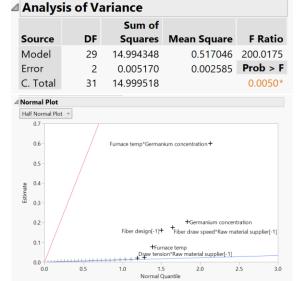
As we can see from the above aliased interactions, Raw material and Germanium Concentration were common factors in both confounded lines. We have two options for dealiasing these interactions. We can either multiply raw material or Germanium concentration with -1 for a new set of experiments, and we chose raw material for this purpose. Again, using resolution IV, a set of 16 new experiments were run to dealias the aforementioned confounded interactions, and 2 additional runs were requested for center points to help us in optimization process, resulting in a total of 50 runs. Since we have 4 categorical factors, we need to set them to one level and get the center points with regard to other continuous factors. However, we only have one categorical factor that was significant, Fiber design, so we needed to get one center points for each level of this factor.

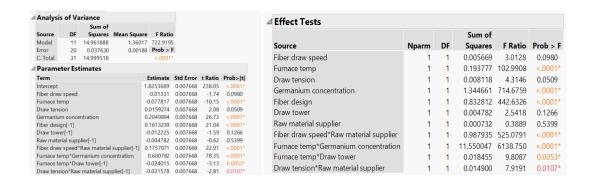
## **Analysis and Results**

As discussed in experimental design section, we started with 32 screening runs in two days. Tables below show the results of effect test and ANOVA. As we can see block is not a significant factor but the overall model is significant. It is in line with the results from REML test given in the previous section. In addition, if we consider a significance level of 0.2 (just for the purpose of screening), we will have Furnace temperature, Germanium concentration and Fiber design as significant factors among the main effects. Fiber draw speed, Draw tension, Draw tower and Coating type are marginally significant and Raw material supplier does not seem important here (it is also shown in the normal half plot given below). Significant interactions are Fiber design\*Raw material supplier (which Germanium aliased with concentration\*Fiber design), **Furnace** temperature\*Germanium concentration, Furnace temperature\*Draw tower and Draw tension\*Raw material supplier (which is aliased with Germanium concentration\*Coating type).

Now that we know which factors and interactions are important, we can remove the rest from the model. The updated effect tests, parameter estimates and ANOVA table are as follows:

			Sum of		
Source	Nparm	DF	Squares	F Ratio	Prob > I
Fiber draw speed	1	1	0.005669	2.1929	0.2768
Furnace temp	1	1	0.193777	74.9619	0.0131
Draw tension	1	1	0.008118	3.1404	0.2184
Germanium concentration	1	- 1	1.344661	520.1771	0.0019
Fiber design	1	- 1		322.1703	0.0031
Draw tower	1	- 1	0.004782	1.8500	0.3068
Raw material supplier	1	1	0.000732	0.2831	0.6479
Coating type	1	- 1	0.003911	1.5130	0.3437
Block	1	1	0.008182	3.1651	0.2172
Fiber draw speed*Furnace temp	1	- 1	0.000453	0.1752	0.7162
Fiber draw speed*Draw tension	1	- 1	0.001518	0.5874	0.5235
Fiber draw speed*Germanium concentration	1	- 1	0.001485	0.5743	0.5277
Fiber draw speed*Fiber design	1	- 1	0.000712	0.2755	0.6520
Fiber draw speed*Draw tower	1	- 1	0.001189	0.4601	0.5675
Fiber draw speed*Raw material supplier	1	- 1	0.987935	382.1790	0.0026
Fiber draw speed*Coating type	1	- 1	0.000024	0.0091	0.9325
Furnace temp*Draw tension	1	- 1	0.002395	0.9265	0.4373
Furnace temp*Germanium concentration	1	- 1	11.550047	4468.092	0.0002
Furnace temp*Fiber design	1	- 1	0.003912	1.5134	0.3437
Furnace temp*Draw tower	1	- 1	0.018455	7.1393	0.1162
Furnace temp*Raw material supplier	1	- 1	0.000888	0.3435	0.6171
Furnace temp*Coating type	1	- 1	0.000526	0.2035	0.6961
Draw tension*Germanium concentration	1	- 1	0.000031	0.0120	0.9226
Draw tension*Draw tower	1	- 1	0.000312	0.1208	0.7613
Draw tension*Raw material supplier	1	1	0.014900	5.7639	0.1384
Germanium concentration*Draw tower	1	- 1	2.98888e-6	0.0012	0.9760
Fiber design*Draw tower	1	1	0.002576	0.9965	0.4233
Draw tower*Raw material supplier	1	- 1	0.002040	0.7890	0.4681
Draw tower*Coating type	1	1	0.002303	0.8910	0.4449

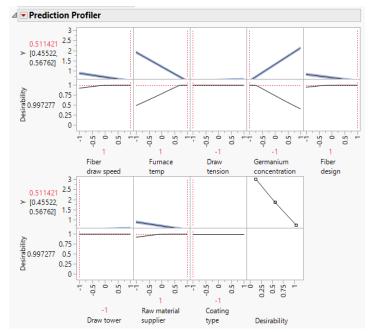




As above figures suggest, the overall model is significant, showing that we are able to explain more than 99% of the variability in the response. Also, after removing insignificant ones, we see that fiber draw speed, furnace temp, draw tension, germanium concentration, fiber design, draw tower, fiber draw speed\*raw material supplier, furnace temp\*germanium concentration, furnace temp\*draw tower and draw tension\*raw material supplier are significant at  $\alpha=0.2$ , but Raw material supplier is still insignificant. In addition, since the whole model is significant, based on ANOVA table, we can conclude that these factors are important. We found 2 significant interactions which were confounded with another interaction. Confounded interactions were as follows:

- Fiber Draw Speed\*Raw Material=Germanium Concentration\*Fiber Design
- Draw Tension\*Raw Material=Germanium Concentration\*Coating type

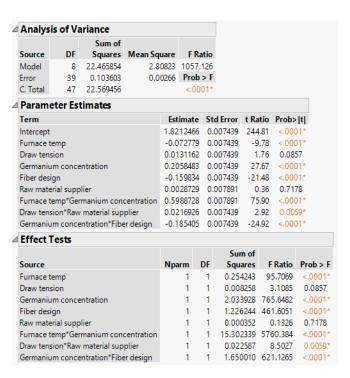
The optimal design for our first experiment is fiber draw speed, furnace temp, fiber design and raw material supplier at high level, and draw tension, coating type, germanium concentration and draw tower at low level and the optimal response for this design is 0.511. We can also see that coating type, draw tension and draw tower have almost zero slope here, which is another confirmation for considering them as insignificant.



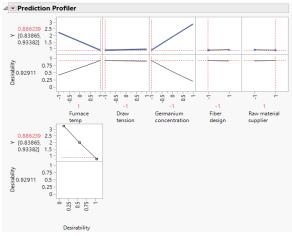
As explained in experimental design section, to dealias the confounded interactions, we have multiplied raw material by -1 and requested 16 more runs. The full model after adding 16 more runs is displayed below. We can see that from the first aliased interactions, both Germanium concentration\*fiber design and Fiber Draw Speed\*Raw Material are significant (the latter is significant at 0.2 significance level) and from the second one only Draw tension\*Raw material supplier is marginally significant at 0.2 level. Furthermore, block is not again significant, whereas the ANOVA table is telling us that the whole model is.

Analysi	s of Va	riance					
		Sum of					
Source	DF	Squares	Mean Square	F Ratio			
Model	32	22.508366	0.703386	172.7091			
Error	15	0.061090	0.004073	Prob > F			
C. Total	47	22.569456		<.0001*			
Parame	ter Es	timates					
Term				Estimate	Std Error	t Ratio	Prob>
Intercept				1.8212466	0.009211	197.72	<.000
Fiber drav	v speed			-0.007939	0.00977	-0.81	0.4292
Furnace to	emp			-0.075218	0.00977	-7.70	<.000
Draw tens	ion			0.0128054	0.00977	1.31	0.2097
Germaniu	m conce	entration		0.2069505	0.00977	21.18	<.000
Fiber desi	gn			-0.159266	0.00977	-16.30	<.000
Draw tow	er			0.0091009	0.009211	0.99	0.3388
Raw mate	rial supp	lier		0.0028729	0.00977	0.29	0.7727
Coating ty	/pe			0.0027736	0.00977	0.28	0.7804
Block[1]				0.0201123	0.013027	1.54	0.1434
Block[2]			-0.011868	0.013027	-0.91	0.3767	
Fiber draw speed*Furnace temp		)	0.0017035	0.00977	0.17	0.8639	
Fiber draw speed*Draw tension			-0.001636	0.00977	-0.17	0.8693	
Fiber draw speed*Germanium concentration		oncentration	-0.001135	0.00977	-0.12	0.9090	
Fiber draw speed*Fiber design			-0.007317	0.00977	-0.75	0.4655	
Fiber draw	speed*	Draw tower		0.0016745	0.00977	0.17	0.8662
Fiber draw	v speed*	Raw material	supplier	0.0187895	0.011281	1.67	0.1166
Fiber drav	v speed*	Coating type		0.0039619	0.00977	0.41	0.6908
Furnace to	emp*Dra	w tension		-0.000369	0.00977	-0.04	0.9703
Furnace to	emp*Ger	manium con	centration	0.5988728	0.00977	61.30	<.000
Furnace to	emp*Fib	er design		0.0056862	0.00977	0.58	0.5692
Furnace to	emp*Dra	w tower		0.0076326	0.009211	0.83	0.4203
Furnace to	emp*Rav	v material sup	plier	-0.003307	0.00977	-0.34	0.7397
Furnace temp*Coating type			-0.000932	0.00977	-0.10	0.9252	
Draw tension*Germanium concentration		entration	0.0034353	0.00977	0.35	0.7300	
Draw tension*Draw tower		-0.002552	0.00977	-0.26	0.797		
Draw tension*Raw material supplier		plier	0.0172639	0.011281	1.53	0.1468	
Germanium concentration*Draw tower		w tower	-0.004947	0.00977	-0.51	0.6200	
Fiber desi	gn*Draw	tower		0.008972	0.011281	0.80	0.4389
Draw tower*Raw material supplier		ier	-0.003162	0.00977	-0.32	0.7507	
Draw tower*Coating type			-0.008484	0.011281	-0.75	0.4637	
Germanium concentration*Coating type		ting type	0.0043143	0.011281	0.38	0.7075	
Germaniu	m conce	entration*Fibe	er design	-0.194497	0.011281	-17.24	<.0001

As a result we can now remove all insignificant factors and interactions and we will end up with the following results. The reason why we decided to choose 0.2 significance level was that we wanted to include all marginal important factors so we could find the best ones at the end. So, after doing our screening analysis we can switch to 0.05 significance level again. Using 0.05 type I error, we can see that Furnace temperature, Germanium concentration and Fiber design are significant from the main effects and Draw tension is marginally significant. From the interactions, Furnace temperature\*Germanium concentration, Draw tension\*Raw material supplier and Germanium concentration\*Fiber design are effective.

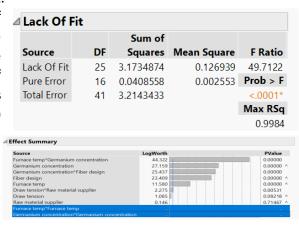


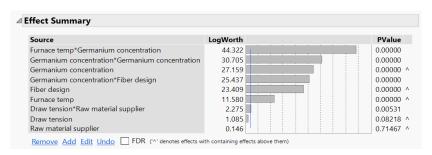
Using the prediction profiler, we can find the optimal setting with this data. Notice that, since we removed some of the main effects from the model, we do not have them in the profiler. The optimal setting here is furnace temp and raw material supplier at high level and draw tension, germanium concentration and fiber design at low level, other factors (Coating type, Fiber draw speed, etc.) can be set to either low or high because they are not significant here. Note that, Raw material supplier itself is not significant but its interaction is, so it does matter what setting we use for it.



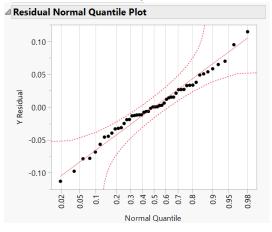
After finding out the significant effects, the next step is to optimize the attenuation using these factors. To do this, we can first look at the prediction profiler plot we have above; we see that furnace temp and germanium concentration can get better, but we did not have enough range of data for them, so we need to run the experiment again by adding two center points which will end us up with 50 total runs and maximum budget limit.

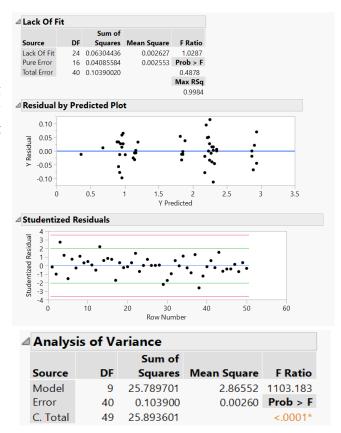
After fitting the model, the results from the lack of fit table tells us that there is a curvature in the model and we need to add the square terms of the main effects in the model. However, if we add all of the square terms of the significant main effects (germanium concentration^2 and furnace temp^2) to the model, we end up with 2 singular effects that are aliased with each other. To solve this issue, we ran two additional analysis, each with only one of them in model. First, we removed furnace temp^2 and the results are shown below.



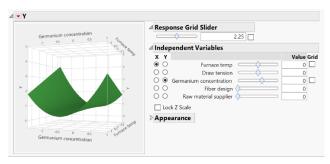


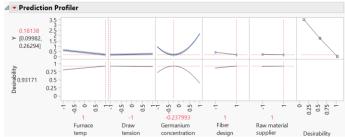
Lack of fit is no longer significant showing, by adding the quadratic term, we were able to explain the curvature in the model. The residual plots and studentized residuals look fine and the constant variance assumption holds. Also, the normality assumption seems to hold as well in the qq plot. At the end, the ANOVA table confirms that the quadratic model is significant.



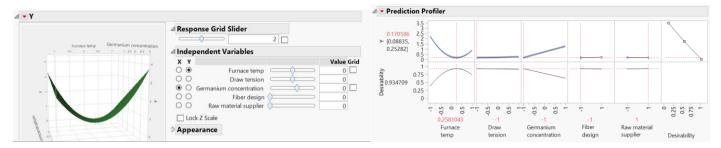


The optimized desirability is shown in the following prediction profiler and surface plot. The optimal setting in this experiment is furnace temp, fiber design and raw material supplier at the high level, draw tension at the low level and germanium concentration at -0.2379 which results to 0.18138 for the response.





If we remove Germanium^2 and keep Furnace temp^2 in the model, we end up with the similar parameter estimates, diagnostic plots and ANOVA. However, on prediction profiler we find different result as follow:



The optimal setting in this experiment is raw material supplier at the high level, draw tension, germanium concentration and fiber design at the low level and furnace temp at 0.2581 which results to 0.170586 for the response. So, this is the best result we got for this project.

#### Conclusion

In this project we were asked to design an experiment to optimize attenuation for an optical fiber under budget restrictions. Eight factors (Fiber draw speed, Furnace temperature, Draw tension, Germanium concentration, Fiber design, Draw tower, Raw material supplier, and Coating type) have been considered in this project.

The results show that furnace temp, germanium concentration, fiber design, furnace temp\*germanium concentration, draw tension\* raw material supplier, germanium concentration\* fiber design, furnace temp\* furnace temp are the significant factors at 0.05 significance level.

The optimal setting of factors that minimize the attenuation is shown below:

- Furnace temp = 0.258 (2051.62 C)
- Draw tension = -1 (note: it is not within our significant factors)
- Germanium concentration = -1 (0.01)
- Fiber design = -1 (index 1)
- Raw material supplier = 1 (note: it is not within our significant factors)

The given setting gives us the 0.170 for the optimal attenuation.

In this project we had 8 factors that potentially affect the attenuation with a possibility of existing variability for runs within different days (block). Therefore, to investigate the effect of each factor in detail, we need to have more runs. However, we have been limited by 50 runs. On the other hand, we chose 32 runs for the screening phase, and we were unable to double them for the dealiasing phase. So, we ended up with an extra 16 runs in one more day which might possibly make our analysis biased. Consequently, there were only two runs left for the optimization step which again might be source of biases in the final conclusion. To resolve orthogonality of factors, we need 8 more runs (2 runs for each continuous factor) plus the center points, but due to limitation of budget we compromised with 2 runs for center points. Having said that, the main limitation of this project was budget restrictions.