MITOCW | Lec 21 | MIT 2.830J Control of Manufacturing Processes, S08

DUANE BONING: OK, so what I'm going to try to do today is kind of a fun case study on spatial modeling. And then we will have time afterwards, especially with the Singapore folks, to try to at least have brief meetings on a couple of the projects. I've met with at least one group. I think one group you guys met with Dave Hardt. At least talked, email. Did you meet also?

AUDIENCE: [INAUDIBLE]

DUANE BONING: OK. Because it looked like you guys are in good shape to, right?

AUDIENCE: I don't know.

DUANE BONING: He said do more.

AUDIENCE: [INAUDIBLE]

DUANE BONING: OK. So we can have some follow up on that as well. But what I'm going to do is go through this case study. My preferences actually this is-- I would have liked to have been able to talk about this a little bit earlier, because this is some very neat stuff on spatial modeling that in the past has sometimes been a part of some projects. So maybe you can look at this and see if there's something extra on spatial modeling you can do.

Essentially, this is going to be based on three papers that are also on the website. There's a paper by Davis, one by Mozumder, and one by Guo and Sachs. So you can grab those papers and learn more. But my basic agenda is as shown here.

First I want to sort of talk a little bit about applying the same response surface modeling and regression methodology that we've talked about for things like process conditions also to spatial parameters. So if you were spatially sampling, could you build a model as a function of spatial coordinates and get a sense of spatial uniformity across a wafer, across a part, across some spatial dimension?

And of course, the answer is yes. What's interesting is then when you go and calculate things like uniformity or uniformity metrics, there's some special issues that come up. And in particular, your sampling plan can affect dramatically what number you get for uniformity. And so we'll talk about ways to not fool yourself on that.

The second part will be then to go back and say, OK, if I can do a response surface model on spatial coordinates and the process is changing, how do I put those two things together to model spatial variation as a function of the process? And there's two different ideas, sort of opposite cuts at the same problem, one based on the Mozumder paper and the other by Guo and Sachs that I think are a very interesting case study in that they show an evolution in the thinking about how to efficiently and effectively do that.

So let me talk first about just purely spatial modeling. And of course, we know that there can be a spatial trend, systematic trend in many, many processes. Most of the methods we've been talking about so far kind of we're looking at random variation, ways of assessing as if we were sampling from a Gaussian or some other process. But underlying that, you may have that noise, measurement noise, or whatever, but underlying that, there may be a systematic spatial trend.

And these may result from inherent equipment or process asymmetries, be highly repeatable. You would like to be able to capture and express them. So I've sort of pictured here an example where you might have a concave or convex shape. If I were measuring some parameter across the wafer, like a film thickness, other geometric parameters, or some material property like film resistivity. And very often maybe gas flows are a little bit different in the center of the wafer and so on. You get these sort of bowl shaped patterns.

But there may be other kinds of spatial patterns, and you'd like to be able to capture that. So our goal is how do we model that? And then how do we boil that down into an overall metric that says how good this is and then drive the process to try to improve that? Decrease non-uniformity, improve uniformity.

So what I'm going to do is show an example. And what I'm going to do here is use synthetic data. By the way, what I mean by synthetic data is I'm going to generate data that I know it's properties of. I am going to generate a data with a known statistical and systematic, a known random and a known systematic component to it. This is a useful technique that you might consider somewhere in one of your team projects as well. You may be dealing with real data. And very often you don't know what ground truth is. And that's the whole interesting thing about a project is you don't really know the truth. What we're trying to do is infer things about it from experimental data.

If you're also exploring some methodology, you want to see does this new methodology that I've come up with that I think solves some problem or applies to some aspect of this problem? You can also test that with synthetic data that you generate yourself. And it's a nice complement in some cases to the other project data that you're doing.

So the data that I'm generating has a systematic component that is a circular or elliptic kind of wafer map. Plus it's got some superimposed random noise on it. And what I'm going to do is sample it in two different ways. So I'm going to generate two sampled data sets off of a more basic data set that I create. And the sampling patterns are going to be two that are commonly used in semiconductor manufacturing.

One is a circular wafer map. Let's see. Do I have a picture of it? Yeah. So here's a radial or circular sampling plan. The basic idea that I take a point at the center, then I take points around some radius, I take additional points around another radius, and I take additional points at a third radius so that I've got these rings of concentric circles as a sampling plan.

An alternative one is to simply use a rectangular or square sampling plan. So those are going to be two different sampling plans. What you would hope is you get the same uniformity metric off of the two, that it shouldn't be sensitive to sampling. I'm going to show you, in fact, that it can be quite sensitive.

OK, so what's my synthetic data? It's shown here. This is in fact the elliptic component, the systematic component, that I create some output. Call it resistivity here just to be concrete. Plus random Gaussian 0 mean noise. So I've got, say, some measurement noise, process noise superimposed on this.

What I'm going to try to do is two things. One is build a response surface model. Maybe, hopefully, recover something close to the underlying ground truth that I've imposed here. And we can now assess how good a response surface model fit is, because I know what the truth is. Second, I'm also going to calculate a non-uniformity metric. And the basic metric is going to be something like a noise to signal ratio. That is to say, the normalized standard deviation across my measurement points. So I calculate the mean. I calculate the standard deviation, normalize or divide. So I'm getting something, if I multiply by 100, something like a percentage non-uniformity across this wafer.

Here's the data, just so you have it. What I get in this sampling plan is 25 total data points, one at the center and then eight at each of 25, 50, and 75 millimeters away from the center. This is assuming sort of a 200 millimeter wafer, if you will. And here's the actual data based on my underlying synthetic data model. Here's the square sampling plan.

In this case, you'll notice that a few of the points, these corner points, actually fall off the edge of a round wafer. So they are discarded. And those are the ones in bold here. Those ones are removed, because they're outside of the wafer boundary. That should be a plus minus. So this actually has, in fact, slightly fewer data points. It has only 21 data points, whereas the previous sampling plan actually has more data. It's 25. Yeah. 3 times 8 plus 1. 25 data points.

OK. One thing I want to-- I'll be showing some jump plots, some surface plots, contours, and so on. And the contouring does show a little bit of difference because it interpolates just to plot things. But in both of these cases, they pretty much look similar except for some adjustment because of the location of the data points. And if anything, the radial 25 points looks like it actually kind of interpolates a little bit more to the roundish or elliptic underlying pattern that I said I was imposing.

This one because of the square grid looks like it may have some artifacts. But that gives you a feel for what's going on. What we've got is a not quite center high. It's not a pure bowl right in the center. It's a little bit off center, and it's not perfectly circular. It's kind of elliptic. So that gives you a feel for what the underlying pattern is.

So I say I want to fit a model. I'm going to fit a generic second order model in x and y. So it's got linear coefficients in x and y, an interaction term, and second order terms, x squared and y squared. So six total coefficients. What I'm doing first here is the radial pattern where I had 25 data points. And you can sort of see what happens here. What we've got is an r squared that looks like it's about 0.65. Doesn't sound great, but I know there's noise in the process.

So I got to look at it in a little bit more detail at the ANOVA and see is my model significant. The r squared is make me a little bit worried. It's not great. But I look, and overall the model compared to the error terms is significant. So I've got an f that is much, much larger than I would expect by chance alone. It's 99.94% significant.

OK, great. So I've got a model that is capturing something going on. Then I can go in and look at the parameter estimates, look at each of the individual terms, look at the t ratio, and look at the probability of those being significant. And they all look pretty significant except this one for sure, a y term. It says no, 0.5 chance that would have occurred by chance alone.

And I've only got maybe 92 %or 91% chance or level of confidence in this xy term. But if I was looking at, say, 95% confidence interval cutoffs, I might reject that term. And we'll see later, that's correctly rejecting that term if you look back. So knowing ground truth, actually this raises already a worry. How come on correctly rejecting the xy term but incorrectly rejecting another term? Interesting. OK.

Let's look at the other case. I'm using a rectangular pattern fitting the same model in this case. What happens now? Well, now we only have 21 data points. Notice that the r squared appears to be better. That's interesting. Again, the model is highly significant. And now if I look at the individual terms, the x is significant, the y is significant, the xy is not significant. That's rejected. The x squared and the y squared look-- I guess the y squared is kind of right marginal right at about 95%.

So it seems like it's coming up with different terms in the model. And also we'll see in a second when I compare them, it's coming up with some different estimates. So maybe this isn't a surprise. The sampling plan is affecting the model that I'm building. Doesn't sound too bad so far. But what I also want to do-- let me compare them first. And then what I want to do is also calculate a uniformity metric across all of this data and see what happens.

So here's our truth. I've expanded out that original model of the systematic component contaminated by noise out to the x terms, the xy, and the x squared terms. And here is the radial sampling plan and the square sampling plan. And the coefficients, some of them are not radically different. They're kind of in the ballpark. But in terms of deciding which terms are kept and which are rejected, we know that this one, in fact, really is 0, whereas this one tries to keep it. This one is very much smaller, and we actually get rid of it. This one's small and this one's small but real. Both of the models fit parameters for it but make different cutoff points on what's significant and what's not.

Now, what's really interesting, I think, is I now take that same data and I go back-- ah, great. Come on in, Charlene. What?

AUDIENCE: She's setting up right now.

DUANE BONING: Oh, Charlene, you can set up in here. Thank you. Is if I take those two sampling plans and I look back at the mean and the standard deviation just across my sampling points and calculate that non-uniformity metric. In one case, it's about a factor of two different than the other. The question is why? And you can ponder that question as you get coffee, because that's what I'm going to do.

So now I think everybody's had a chance to ruminate on this question of non-uniformities or at least read the next two sub-bullets at the bottom of the page, and then everybody has a good idea of what might be going on. One is, of course, we've got systematic curvature. I'm not random sampling, so I may, in fact, hit different points on that curvature. So that's one hypothesis is just because of the particular places where my points happened to fall, that's one idea.

What we'll find out, and I'll show that by more densely sampling, that's actually not what's going on necessarily in this problem. Actually, it's the structure of the sampling plan. And underlying that is how much data or what spatial region is each data point representing of the underlying wafer surface. So let me try to explain that.

So what I've done here is gone back and tried to say maybe I just wasn't sampling enough with just 25 or 21 data points. What's my best guess at the true non-uniformity? Which one of these two should I believe? Which one is closer, the 1.9%, the 3.2%? What's a better sampling plan?

Actually, before I go into that, what do you think? I need a vote here. How many people think that the radial sampling plan is probably closer to the truth in terms of both the model, but more importantly, the non-uniformity number that comes out? And how many people think it's a rectangular sampling plan?

So all those think the radial sampling plan is going to be better, raise your hand. I got I got two here and a half. three. We got two, three, three in Singapore, four in Singapore. Oh, we're talking it up. We're talking it up. OK, how about the rectangular sampling plan? About evenly divided. I got three here, one in Singapore. And OK, that must leave four asleep people who didn't vote, something like that.

OK. What I'm going to try to do is get a close estimate to what maybe better truth is. And the way I'm going to do that is try to take a very dense spatial sample so that hopefully it's not just sort of the random location of points. I'm going to kind of get really close to almost a full representation of the surface. And this is almost a 900 point-- slightly under 900 point sample. What is it? 665, I guess. A 29 by 29 spatial sample and then the points that fall within the radius of the wafer.

And as you do that, you actually get a, quote, a "true non-uniformity" that's about 3% closer to the other number. And if one does a response surface model fit, there's a couple of important points to come up with this. I go in and I fit the model. In this case, the coefficients that come out of the model are actually quite close to the true systematic piece. But notice that the r squared is still only about 0.77, something like that. Why is the r squared so bad? I got tons of data. Why is the r squared so bad?

And the answer is kind of highlighted right here. I've still got lots of random noise in the process. I was adding magnitude variance 0.49 or 0.7 standard deviation values to all of my measurements. I cannot model the random noise. That's my residual that's leftover. I could include it as an estimate of the random noise and know what that is, but I can't pull that into the systematic model. So there's always a ground if there's noise and randomness that you can never capture.

OK, so this seems to be suggesting that maybe that square sample with the smaller number of points was a little bit better. What's going on? This finally gets me to the Davis paper. And what they're going to propose is a method for looking at systematic trends. They're first going to make the point that the signal to noise ratio, this standard deviation over mu or mu over standard deviation, if I turn it into a signal to noise, is sensitive to both the location and the number of measurements. And they're going to propose a difference statistic rather than just the sample standard deviation to try to get at what a better, truer measure of overall non-uniformity will be.

Turns out this integration statistic is based on fitting of splines to your data, and not a lot of people go in for that. What I'm going to show is that a relatively simple approximation achieves a lot of the goals of this integration statistic, which is worrying about uniform sampling spatially, or alternatively, by weighting the importance of each of your measurement points by the amount of area that it's representing. So that's the core of the problem.

So what Davis does is looks at, first off, a sort of a smaller radial kind of picture compared to the one I showed. It's only got 13 data points. Here's the 13 data points with just linear interpolation between them. So that's the surface example that he's using. And if you just purely linearly interpolate between them, you can see that's a pretty coarse approximation to that surface. One can easily be worried that you might get bias or variance errors in trying to calculate with just those data points. What they're proposing is using thin plate splines, these TPS methods, which are conceptually, basically fitting localized polynomials such that you have minimum curvature over the data points or knots in the data.

All right. I'm sure there are statistical packages and others that can help fit those thin plate splines. I found when I've tried this it's very tricky, because what localized means can be a little bit tricky. You've got other kind of human intervention going on in trying to fit these things. So it can be a little bit tricky.

But essentially, what they're after is trying to get a model for the whole surface. And then once they have the whole surface, then calculate an overall metric as if I had an infinite number of measurement points representing the whole surface. So I want an integrated measure, an integrated statistic i, that captures the total deviation from a target or total deviation from some nominal value.

And so what they're doing here is calculating a normalized integral of deviation of this interpolated surface, gr is their thin plate spline, from some target value, and then just integrating that over the surface. Total deviation from some target. Presumably the target would be just one value. I guess you could extend to whatever target spatial distribution you wanted. And then they're normalizing it by the total volume of the surface. So it's sort of like a standard deviation over mu except what it is is total deviation over total volume.

Now, this is just total deviation. And if I was just summing up deviations as a measure of variation just in some randomly sampled data set not having to do with space, you might be worried about that. Why would you be worried about just summing up total deviations? Positive deviations, negative deviations, they'd cancel out. So if I was plus and minus huge amounts but about the same amount plus and minus, I might fool myself and think they might cancel out and come to 0. And I'd say, oh, everything's great. And they're really completely different.

And that's why we often use things like squares or absolute value to try to account for or penalize for both plus and minus deviations. And so in the Davis paper, one can go in and apply some other generalized loss transformation to that deviation. So you can do a sum of squared deviations that looks a lot more like a standard deviation. But again, the idea here is they want to integrate that over the whole surface. Represent all of the surface, not just the few data points that you had, but do that using the interpolated function.

Now, there's an approximation to this integral that they mention, which is if I have my lost transformation, I have my actual data points. So these are my measured data points. And I have the deviation of that measured point and then my h might be the square or the absolute value. If I simply sum those for my data points, I get something close, again, to standard deviation, a sum of squares.

But there's one big difference. And that's right here. The c sub j. And essentially the c sub j is a weight that corresponds to if I was doing the spatial integration, I would be integrating over these thin little patches around each of my measurement points. It wouldn't be just that measurement point. It would be representing the surface near that measurement point.

And the c sub j is a weight that says, how much area on my surface does that measurement point represent? How much vote does that one point have? And that can be very different in these two sampling plans. So what's neat is that if you weight appropriately for area that the point represents, you actually get fairly close to something like the integration coming from a thin plate spline.

OK. So there's a couple of plots in Davis's paper where they're taking that typical radial sampling pattern. Here's a perfectly circularly symmetric pattern, 15% systematic variation, 2% random superimposed on it. And then what they're doing is looking at an SNR. This is basically just sigma over mu of just the data points. And they're showing what estimates over 300 different runs of this model with different amounts of noise each time, just different instantiations of the random 2% noise, what do they calculate for sigma over mu? So each time they do 300 runs, out here they do it taking 73 different measurements around the wafer with a typical radial spatial pattern. This is the problematic one we saw earlier.

And notice there is a spread, as you would expect, because the noise is in there. But look what happens as you go to smaller numbers of measurements. First off, the spread or the variance in your estimate of sigma over mu increases. That makes sense, right? I have less data. I'm going to have more variance in how I'm estimating standard deviation. We already know that estimating standard deviation in general requires a lot of data. It's really a tough thing. So it's getting worse.

But even worse than that, especially for, say, five measurements or even 13 measurements, a little bit there, is the average value coming from repeated evaluations is not equal to the true underlying average? In other words, there's bias. The estimator of sigma over mu is biased. It's not a very good estimator. And as you have smaller numbers of parameters, that unequal weighting of the surface is biasing or fooling you. Even on average you're going to on average be wrong.

What they're showing is if they actually do their whole thin plate spline I stat statistic, they argue both the variance is smaller, but more importantly is it's unbiased. Even when you get down to small numbers of measurements, you're at least on average right. And that's just these two plots that are simply boiled down into this comparison that shows as a function of the number of measurements how good the statistic is. So the spread of the statistic over the mean of the statistic. And showing that the variance of the statistic itself is getting smaller with the integration statistic. So you're basically doing a better job of estimating with their proposed integration statistic.

Now, it's not perfect. And in fact, conceptually, if I only have a small number of sampling points, I can still be prone to where those sampling points happen to lie if I have a complex surface. And so they're also showing an asymmetric underlying non-uniformity, this crazy shape down here in the lower left. And imagine I'm only sampling with five data points. Well, if those five data points land in one orientation or if they land then slightly offset in angle, I am spatially sampling different parts of that surface. So what happens then?

And what they basically show is, yes, their I stat does have some sensitivity. But because you're trying to interpolate the rest of the surface, you do a better job. You're not completely dependent just on that local measurement point. Both approaches are still highly sensitive to angle, but the pure sigma over mu is much more sensitive. You get about 20% smaller variation with their I stat.

OK. So what are the key lessons out of this part of the case so far, this paper? I think the key lessons are watch out for your sampling plan. And especially if any of you go into the semiconductor industry or I guess any industry dealing with round substrates, be careful, because you will run into this circular sampling plan again and again. And if you're calculating some metric that is equally weighting each of these points where each of these points is, in fact, points off on the edge of the wafer are representing a bigger area. But they're equally weighted with smaller sampling points near the center of the wafer. And this is a problem that I've seen come up again and again and again that people, I think, are not as aware of as they ought to be in the industry.

It's easy to fix. It's easy to fix. One fix that I recommend, that I think is the cleanest, is simply do a more uniform sampling. That rectangular sampling has a great benefit. How much area does each point represent? An equal area. It's not dependent on an r theta kind of calculation. Second is if you do have some non-uniform spatially sampling plan, you can kind of fix it by doing a weighted metric or a weighted regression. Yeah?

AUDIENCE: Normally, you have a different kind of sampling plan like that for, say, the [INAUDIBLE] terms. But it would be four ranges of the [INAUDIBLE] and four ranges of the wafer. And then we would applicate hundreds of times over the wafer. So in that case, does that factor then, because you're--

DUANE BONING: So Nalish's point is often, say, if you're sampling within each chip, so here's our big wafer with lots and lots of these chips, you might sample spatially the same ring oscillator in different corners of the chip and then-- since there are multiple chips. I think, again, you've got to be a little bit alert to what it is you're doing with that data and how you're calculating it. But that chip by chip kind of sampling plan has the advantage that chips are generally rectangular. And so maybe by accident or by free, you're usually getting a more equally representative sampling plan that lets you fit or calculate metrics a little bit more fairly.

AUDIENCE: [INAUDIBLE] you had mentioned the sensitivity to the amount of area cover. But a lot of times you have more of a sensitivity to factor, depending on if SRAM is nearby or a populated logic is nearby.

DUANE BONING: Yeah. So this is kind of good for wafer level modeling. Nalish is also pointing out in many cases, maybe this is an SRAM and this is logic over here, and you're sensitive to other, what I would call, finer grained systematic sources of variation, like layout pattern density, proximity to other structures, other perturbing effects. And those are extremely interesting and fascinating to look at.

And I think the basic strategy is if you know what those are as factors and can represent them as factors, you can actually do some of the kind of ANOVA and model fitting with those as factors in addition to or separating from spatial xy dependencies. So it's usually better if you know what those factors are not just trying to let x and y be the stand in for them but actually explicitly include those things so you know what causes what. And that's a very interesting area. Any other questions on this spatial sampling? I think it's kind of cool but relatively intuitive once you think about it.

So the next thing I want to do is say, OK, now I know how to be careful in building a spatial response surface model. But what if I'm worried about how that surface changes as a function of process conditions. What I'd like to do is have both process and spatial dependencies in it. I'm going to show first off an approach that basically builds a two layered response surface model. This is the Mozumder and Loewenstein paper.

And then I'll come back to it and flip it around that says in the first case, the first approach is basically to build the model of space and then take each of the coefficients in your spatial model, your a3 times the xy term, and now model that a3 term as a function of your process conditions. Got it? The second approach does the opposite. It says, I'm going to build a model of this particular site location as a function of process conditions. And then I can fit a spatial model once I have that. So it's two completely separate or orthogonal ways of looking at the problem.

So the first one here is "Model for Semiconductor Process Optimization Using Functional Representations of Spatial Variations in Selectivity." And they make some of the same kinds of observations that we've talked about. But what they're doing is basically trying to say, OK, I have to be careful in how I calculate my non-uniformity metric, but I want to calculate that and then understand how that changes as the function of different process conditions. And in particular, they're looking at a silicon nitride wet etch looking-- or silicon nitride plasma etch, excuse me. And looking at things like the etch traits of the silicon nitride, etch rates of the silicon dioxide as a function of things like gas flows process conditions.

Now, they're actually worried about two key parameters or three key parameters. One is the etch rate. Faster rates are better. Uniformity is better. But they also worry about the relative etch rate between nitride and oxide and they want high selectivity. They want to be able to etch through the oxide but stop on the underlying nitride or vice versa, I guess. This is a nitride etch.

OK, so what they do here, they take a typical spatial map, 19 measurements. The paper says two concentric hexagons. But if I look at it, they look like two concentric octagons, because it's eight points. So I don't know where that came from. But it's basically these octagons plus the center point. There may have been replicates to get to the 19 measurements. Not entirely clear. But then their process conditions are temperature, microwave power, pressure in the chamber, as well as nitrogen and hydrogen flows. So it's sort of a five factor DOE that they're going to be exploring.

And what they note is that the removal rates do have a spatial dependence as a function of these conditions. Things like gas flow rate means more of the gas is coming in contact with the center or the edge of the wafer perhaps at different temperatures, different process conditions. So they do have an etch rate that varies spatially across the wafer, and they would like to model that.

Now, what they're going to do is say if I know my rate spatially across the wafer, I am going to then calculate non-uniformity as a derived parameter. I'm not just going to sum up my data points. I am actually going to try to use kind of the model and do a ratio of standard deviation to mean. But it is derived from the model.

And so their basic idea is to do this two level model or two layered model of the etch rates. So what they're going to do is say the etch rate is a function of both spatial terms, sort of the same thing we saw before, xy, xy, x squared, y squared. But also of processed terms. So what they would like to be able to do is plug in and say, OK, if I know my hydrogen flow rate and I want to look at some particular xy location, I would like to know at this process condition and at this spatial location what is my etch rate.

So what they have to do is basically fit this two layered model that has both dependencies on the spatial terms and on the process terms in it. And another way of looking at this is essentially what they're doing is, as I mentioned earlier, they are building a spatial model where each coefficient, sort of an a1, is also then a function of the process conditions.

So here's an example of a regression surface that comes out. This is the spatial regression. Again, as a function of xy and so on for the removal rate or etch rate of silicon nitride. And then these coefficients themselves, they go in and they build a model of those coefficients as a function of the process. And when you do that, you can get some pretty good fits. Look at these. These are square root of in the range 0.93, 0.90 is the worst, 0.99 for each of the significance for each of those terms. In other words, this is the spatial model for that c sub 2n whatever. C sub 2x. I guess c sub 2y yn. But the goodness of fit, then, of doing that spatial coordinate model is a function of the process. It's a pretty good fit.

And then what they do is say, OK, I'm going to do polynomial regression, as we saw before. They actually find kind of a nasty thing. To get good fits, they had to go to cubic fits. That may be. Just a couple of other points of interest here. They use the Latin hypercube sampling rather than a central composite sampling. This is a really cute sampling plan. Did I describe this Latin hypercube sampling plan before? I think I did.

But basically this was saying, OK, if I have a space, some x1 parameter and some x2 parameter, and I only have five data points I'm going to allow myself to sample, what you do is you basically divide both your coordinates or dimensions up into five spaces, five equal spaces. And then you take a sample point and make sure that every one of your five sample points uniquely follows in one row and one column. So you might get something like this. What am I missing? Maybe that.

And what's cool there is if you project down onto either x1 coordinate, now I've got sort of five data points spread throughout that space. Or if I project down onto the x2, I similarly have five coordinates spread throughout that space. You have to be a little bit careful, because that algorithm could have accidentally given you a spatial correlation in your sampling.

So for example, those points follow my rule of one row, one column. But that has an unintentional correlation in it. So there's a part of the Latin hypercube sampling algorithm that swaps rows and columns to avoid that. So that's what they do there. Actually, it's good that you're aware of Latin hypercube sampling, because especially for constrained cases where you know how many data points you can sample, it's actually quite an interesting approach.

OK. Then what they go in and do is build out of these fundamental rate functions that they fit, they built derived functions. These are derived. It says if I know the removal rate or etch rate at different locations, now I can fold that together to calculate an aggregate uniformity kind of using some of these ideas we were talking about before. I'm not simply calculating from my raw data uniformity number and then building a separate model, a separate response surface model, a separate polynomial model of uniformity.

Instead I'm getting close to the physics as I can and build this great model and then recognize that uniformity is simply a calculated function across that underlying rate function. And they happen to be using something like a rate sigma divided by rate mean. So they're integrating that over their space. By the way, this little n simply indicates that they're calculating separately or building models separately for the oxide rate and the nitride etch rate. They both vary spatially. So they built these two models.

They do the same thing for this selectivity. Again, this is the ratio of etch rate of nitride to silicon dioxide. And that also is a derived model. And notice that there's something really important about these derived models and why they help. May be less obvious, although we've already talked about it here with the uniformity metric. Maybe they'll focus first here on the selectivity. Is selectivity a linear function? No. It's kind of nasty.

It's a ratio of two other rates. Why would that necessarily be linear? This is kind of a complicated functional form that if you were trying to directly model selectivity, it might be very difficult to capture by chance that complicated functional form. And even worse, think about the computations that go in the calculation of standard deviation. Remember standard deviation? It's the square root of the sum of some x sub i minus the mean squared divided by 1 over n minus 1.

So you got squaring. You got a square root in there. That's a very nonlinear operation. Why would a linear model be very good at capturing that? In fact, if you tried to do it directly, you might get a non-uniformity that needed not only cubic terms, but who knows, fourth order terms, fifth order terms, which would be quite complicated to actually fit it. And what they're showing in the paper is an example where a relatively simple rate function, they're doing it here also as a function of, say, theta and r, not just, say, xy, but that transformation is pretty simple.

If you then feed that into the uniformity calculation, the expansions of squares and square roots, you get kind of almost, if you will, for free a functional form that is quite complicated in capturing the true functional dependencies of uniformity. That's the important idea here is you have a simple, direct, underlying model of rate. And then if you've got a complicated functional dependence, like non-uniformity metric on that, you calculate that from the simpler underlying model. I like that. I think that that's nice. That's a key idea in this paper.

Then they go on and do, once they've got these models, they do multiple objective optimization trying to minimize or maximize selectivity, minimize their non-uniformity value, maximize the rate. They've got some constraints. They go in and kind of do a multiple objective. Come up with some optimum values and show how much improvement they get.

But the key idea, I think, was that contribution of this notion that you want to start with something that's simpler model. Then if you've got a complicated optimization objective, use the simple model, build the complicated objective or complicated derived function, and then use that. So any questions on that? Does that seem pretty clear?

What I'm going to show next is the last piece of this puzzle, which is-- yes, before I do that?

AUDIENCE: [INAUDIBLE]

DUANE BONING: Yes.

AUDIENCE: It's different to [INAUDIBLE] when you compare rectangle to circle. So the circle has the minimal area to perimeter ratio, while the rectangle has the maximum. So it's most likely when we have a rectangle we most likely have to have a uniform sampling. Is it correct to look at it that way?

DUANE BONING: I don't see a relationship between the perimeter to area.

AUDIENCE: For a given perimeter.

DUANE BONING: And the reason I'm not sure that works is if I have-- here's my wafer. In the rectangular case, it's easy to sort of tessellate the space and cover the space with non-overlapping regions that each data point represents. In order to do that with the circular sampling plan, each one of these is not a circle. The way you have to sort of do it is each one of these is this arc to cover the space with non-overlapping.

Now, who knows? If you allow yourself sort of overlapping coverage, maybe what you're talking about makes sense. But I don't think it's really a perimeter that each area-- I think it really is area, just purely area that it's representing. But I haven't thought about the perimeter issue. It's kind of interesting.

And by the way, this is the kind of calculation you would do back in spatial modeling to figure out as a function of your xy coordinate for your measurement point how much area that ought to represent. You're basically dividing it up and you have to kind of do those calculations.

OK. The last thing I want to show is this paper by Guo and Sachs. And again, I remind you Elliot Sachs is a professor here in mechanical engineering who two lives ago did a lot of work in semiconductor process control. His life after that was 3D printing. And his current life is now solar energy. So he's made some wonderful contributions and some big moves. And I actually like this paper quite a lot.

The basic idea in this, it's looking at this issue, again, of spatial uniformity, how one does modeling, and then uses that for optimization and control of uniformity. And basically, here he builds on, I think, he builds on this notion that we learned from the Mozumder and Loewenstein paper, which is build simple underlying models and then combine them and use them to solve the larger problems. But keep it as simple as you possibly can.

And he suggests, together with Andy Guo, he suggests that we should flip around what Loewenstein and Mozumder did and actually build models for each spatial location first. Don't fit the spatial model first and then take that coefficient and try to build a model of that coefficient as a function of process. Instead just look at that one spatial location and build a model of that spatial location's response as a function of the process conditions.

And now I can combine that to other spatial derived things, like fitting a spatial model or not necessarily even fitting a spatial model, but now I can basically use spatial information, not lose it. I still know the left side or the right side is higher or lower. What happens when you calculate standard deviation over mu? You lose information of left side or right side was higher or lower. You just boil it down to they were different. He says keep the sites. Keep the sites. Use that information to drive towards an improvement. Keep the simple model, build those, and then you can combine them.

So let's look and see how that works. Here's the basic idea. Let's say I was measuring three different locations on the wafer, y1, y2, y3. And I have two equipment settings. So two different process parameters. What you're going to do here is build a response surface model and a DOE.

And there's two different approaches one can take. One can take what I'll called the single response surface, the classical approach that says I'm looking at different combinations of my input parameters. I measure my three points. I've got three data points I can then calculate for that process condition, that run number one.

I can calculate the standard deviation across those three data points, divide by the mean of those three data points. That gives me what the uniformity spatially was for that run. I can then repeat that for all of my different process conditions. And then I can build a single response surface at the end, SRS, Single Response Surface, that is the response surface of the response of non-uniformity as a function of my process conditions. That's a classic approach, and you will still run into that in many papers. We didn't see it in the Loewenstein paper. They used a slightly different approach. But this is a classic approach.

And what he says is no, don't do that. Do not do that. Instead use the same design of experiments, same data, but do something that I think is a little bit smarter, which says take your different site. This is a site one. I'm always measuring the left side. Maybe this is the center. This is the right side of the wafer. And build a response surface for that site response, maybe it's the thickness or the etch rate or whatever, as a function of the process conditions.

Do the same for all of the sites that you've got. Build separate multiple response surfaces. And now you know how each site responds for whatever process conditions you want. And now you can combine them if you want a uniformity metric to see which is better. But you can also do smarter things like try to balance them. Get the left side back up to match the right side. Do other sorts of things.

OK, let's see how that works. His basic point here is very similar to the Mozumder and Loewenstein, which is the single response surface is a very tough job to directly model in one response surface, this highly nonlinear sigma over mu. And in fact, if I try to build a second order model, you often need a second order model as a function of the process conditions, which means lots of sampling in your multiple levels in your process conditions, at least three, maybe more. And you get a very complicated model.

What he also shows is that very often with the multiple response surface model, you can have much simpler lower order models for each site. And in fact, what he will do is show if I just build a linear model for each site as a function of my process conditions, meaning I actually need fewer levels in my DOE. I can actually get simple linear models for my individual sites.

And then when I functionally calculate the mu, the mean of those three, that's a simple formula, and standard deviation over mu is a simple formula. But embedded in it is the right nonlinear functional form for sigma over mu. And so you get this non-linearity out of very simple linear models. And that's the key idea. Similar to Mozumder and Loewenstein. Build the simple model and then use it and combine it in whatever complicated functional form you have.

So he argues that you can actually get a smaller number of data. Not just use the same data in a smarter way, but actually sample less in your process conditions in many cases. So you get a savings in the DOE.

Another very nice advantage that's articulated in the paper, I'm not going to go into here, is that each of those models of the process is a nice, simple linear model. And you can use simple linear models with the cycle to cycle run by run control that Dave Hardt talked about a couple of lectures ago. You can basically take that as a model, adapt that model rapidly to changing process drifts or conditions in your equipment, use that updated model in sort of a PI or PID kind of fashion, and use that to improve the selection or pick the selection of the right process condition for the next wafer run. It's a lot easier to use this kind of a model for cycle to cycle control.

He also has some other very nice points. He does a bit of analysis in the paper saying which one is less susceptible to noise. It turns out that the site models kind of average out noise a little bit better than putting everything into signal to noise.

And then the last point is a really important one, which is if you get this really complicated functional cubic model that tells you how non-uniformity changes as a function of your process condition, do you as the process engineer have any idea really intuitively what will happen if you change a knob or what's happening spatially across your wafer? His argument is you've lost a lot of that information. If you actually have your site models, you can build that full spatial map, see how each site is changing as you change your process condition, and actually see-- build some process knowledge, see what's going on much closer to the process.

So here's a couple of examples. This relates to, again, the point that the complexity of the uniformity metric has a lot in it and you actually lose information. So here's the point where if I measure the thickness at the left site, thickness at the right site, so red is my left and blue here are my right measurement points. Let's say I have simple linear responses for how those two sides change as a function of the input parameter. They really do change linearly.

Notice what happens, in one case, the left side is much thicker than the right and in the other process condition, the right side is much thicker than the left. If you actually just directly do a single response surface kind of idea or do our calculation of sigma over mu, you can easily get a little bit fooled here. They're both equal in terms of a uniformity value, but you've lost track just looking at these values by themselves. You don't know what's going on on the wafer. So looking back here, you can see what's going on in the wafer.

The second point is if one wants to actually go in and do a medium input and interpolate if I have these site models, I could actually project and guess and say, I think that if I were to run at the middle, I would have very good uniformity. The thickness would be the same in both cases. And if I calculate a value based on these underlying models, I can actually project what happens at the intermediate values. This would suggest for a control problem or an optimization try the center point. Whereas if all I had were these two data points, I would say they're the same. I can't really improve things much. Yeah?

AUDIENCE: Left or right with respect to [INAUDIBLE] if you have a wafer and you have a knot and then [INAUDIBLE] wafer is in a certain way with respect to the notch, but you still have a reticle. And that can be with any direction with respect to the notch. So any information that you get, you would never be sure for the friendships how--

DUANE BONING: Yeah, I'm not projecting down to chip level. I'm thinking here wafer level. And I think you always know-- you can always refer to something with the wafer notch or the wafer flat. So I think that's-- right and left, these are conceptual here, but I think they can be very well defined. If you are also worried about repeated chip scale things coming from lithography fields, that's an extra layer of spatial concerns.

OK. And then this is essentially the implications for control are pretty much what I was just describing. If you just use sort of the high and the low, and this is your intended input, you can now predict a little bit better what your actual output would be with these simplified cases.

So in the paper, they do some examples actually using some experimental data that they generated. This was for the low pressure chemical vapor deposition of polysilicon. And it has spatial uniformity in it not just on the wafer. Actually what they're doing here is spatial uniformity across the tube. So as a function of wafer position, these are big multi-wafer tubes that might actually process 24 or 48 or 96 wafers all at once. And there's a gas flow. There's an injector for the gas, and the gas flows. We've got a center injector, source injector. The gas flows are somewhat non-uniform. And you may, in fact, have systematic spatial dependencies as a function of temperatures, gas flows, and so on.

And what they basically went in and did is compared single response surface models to multiple response surface models. Pictured here is basically showing this complicated dependence of signal to noise ratio, the single response surface model, but the simple dependence of deposition rates on these two parameters. So down here, this is basically saying my site here is wafer number 26. I want to know what the average thickness on wafer 26 is as a function of these two flow rates.

And it's a nice, simple dependence on the flow rates. If I look at a different site, a little bit further down the tube at wafer number 124, I also get a simple model. But if I now combine these into a sigma over mu, it has a very complicated shape. Built out of these simple dependencies, simple responses spatially, they combine to a very complicated non-uniformity signature.

Then what they did is say, OK, I'm going to compare SRS to MRS. And I'm, furthermore, going to tie one arm behind my back when I do MRS. The arm I'm going to tie behind my back is I'm going to let SRS do three level DOEs. Two parameters, flow rate one and flow rate two. But I'm doing a full factorial three level. I do nine different process experiments. And I fit the model for non-uniformity for that.

For MRS, I'm just going to let myself pick the high and the low. I'm just going to do less than half the number of experiments and build my site models and then see what kind of predicted signal to noise ratio I will get. And I'm going to look now at also how noise factors into this and see, do I do better with SRS or do I do better with MRS?

And this is a picture that shows, hey, in these two cases, they're kind of suggesting different non-uniformity signatures. But what happens now if I repeat that with different amounts of injected noise? And it's basically showing that the MRS in four repeats of this kind of experiment with different amounts of injected noise pretty much gives the same surface each time.

Look what happens over here in the SRS case. Three out of the four times, they're kind of the same bowl shaped. But one case, you actually get this weird hyperbolic surface. More susceptible, this is just kind of an example not proving anything, but just showing an example that actually the models you get with the SRS can actually vary and be much more sensitive to noise even though you got more measurements, more spatial sample or more levels of your DOE.

And then they went in and did sort of a more thorough example, building lots of models, and then using that model to drive or select what an optimum point is. So maybe you were trying to maximize the removal rate, find the optimum point in each of those cases and see what the spread is with the SRS and MRS in the predicted optimal points.

And you can see over here on the right the basic conclusion that they come to is based on the single response surface model and the multiple response surface model. They're spread in the two in terms of the optimum. But the single response surface model is basically biased. It drives you to an optimum point that is actually not close to the measured optimum point on average. They basically say single response surface modeling is dangerous. Not only is it inefficient, but it can drive you to make errors. So that's a very interesting observation.

So I think I'm going to wrap it up here so we have a little bit of time to talk about some of the projects, especially with Singapore people while we've got the video. But I think these are really neat papers, really neat ideas. Basically first idea is spatial sampling, how you sample, and what area each of those sampling points represents matters. So you can be smart about that, using either weighting or uniform sampling. And then there's some neat ideas about combined process and spatial modeling.

I basically really like the multiple response surface modeling approach. I think the Mozumder and Loewenstein was a nice stepping stone on the way to recognize that you'd like to build simple models and then derive on them. I like the Guo and Sachs because it carries it one step further. It says build simple models of your response at each of your spatial locations and then combine those or use those as you see fit to either boil down and come up with an overall uniformity or make other kinds of control or optimization decisions.

So any questions on that before we switch over to talking about some projects? Think about it. Maybe some of your data is actually a spatially sampled for one of your projects. It'd be cool. Cool to add some spatial modeling in there. How finely refined can you probe that monkey brain?

OK, that's it. I know some of the folks here in Cambridge, I've met with one group. I think Dave has talked to Matt and you guys are thinking. I talked also-- will want to try to set up a meeting. I think I sent email pretty much to everybody saying I think the projects look good. Here's some ideas you might think about. The purpose of the meeting is not to get approval.

Consider this now. You're approved. Go, run. You've only got a week. Less than a week. Go full boar. The purpose of the meeting is to answer questions, help out if you've got additional inquiries. Feel free to contact me, Dave Hardt, Hayden, at any time. But I would like, if we haven't had a chance to talk yet, just try to set something up.

AUDIENCE: [INAUDIBLE]

DUANE BONING: Well, one and a half hours less than one week from today will be the-- so what I will also be doing is making assignments for what groups will present on Tuesday for the next week and what groups will present on Thursday of next week. So it'll be kind of random luck of the draw. But I'll let you know by the end of this week when your time will be. And then everybody's reports are due on Friday of next week, the end of class. OK?

I didn't mean to panic you. You've got almost at least a full week for the presentation. And then in fact, one idea is if you get some quick questions or feedback from the class, from us, during the presentation, it gives you a quick chance to maybe add a little bit or do a quick additional analysis before the report is due. So I actually do recommend that you take that opportunity to fold in some feedback from the presentation. Do not think that you're locked in completely at the presentation. That's most of the way there, but it does give a chance for a little bit of additional work in the last couple of days. OK? Thanks.

So let's see. Singapore folks. First I think I have-- let me get out of this. Somebody sent me slides. Let me talk with that group first. Yeah. So David, Stephen, Jenny. These were the slides you guys sent. So somebody want to-- somebody talk me through these. I can push buttons if you want or you can come up. It's probably easier to do this than try to switch control to Singapore.

AUDIENCE: OK, so our project try to apply the [INAUDIBLE] method to supply chain case study. So we can click to next slide. Next slide. I think we can skip this. Also this will be basically the model. We can consider the supplier and Company X and can consider the customers.

So following other parameters that we will consider, [INAUDIBLE] service level that your supplier promised to you will be the buffer size in the company between the machines. And I have TDF and TDR. And p, the production rate of the flow line. LT, the transportation lead time from company to a customer. LC is the lead time that the company promised to the customer. And mu and sigma for the demand.

DUANE BONING: So quick question. Are all of these sort of input parameters or are some of these output parameters?

AUDIENCE: These are the input parameters I think. All of these parameters are the DOE parameters.

DUANE BONING: OK. So some will be sort of set that essentially define the setup. And some will be DOE. That's what you're saying?

AUDIENCE: Yeah.

DUANE BONING: OK, OK.

AUDIENCE: And next slide. It's about the problem that is for the company. Sometimes the service level will out of the range. And also using the data for the service level is unlike other parameters. The data is quite less. So we need to detect the shift as soon as possible. And after that, you may want to improve some parameters to improve the service level. So the objective-- yeah.

DUANE BONING: So you've got sort of two branches here in part. One is essentially SPC, right? So that's what you're saying? And the other is optimization.

AUDIENCE: Yeah, that's right.

DUANE BONING: OK. Got it.

AUDIENCE: OK. And you click again. This objective. To find the first step is apply the SPC. And you find out what are the main effects in the second stage that you apply the RSM to do optimization to give advice for the company that how can you distribute the budget to adjust some of the parameters to get the best result for improving the service level.

DUANE BONING: OK. So one thought here, just going back. Let's see. Back to this slide. Some of these parameters might be discrete, which is interesting, I think. So it looks like buffer size is probably discrete. Is that the plan?

AUDIENCE: Yeah, yeah.

DUANE BONING: Is that the only discrete parameter?

AUDIENCE: Yeah, I think that's the only one.

DUANE BONING: Because that might be interesting in terms of the DOE. Does the discreetness matter? You have only choice A, B, C, and it might be hard. Here it might still be monotonic, meaning there really is a progression. So you could have discrete levels, and they might have three levels that are ordered in some meaningful way. So that'll be an interesting-- that might be an interesting twist both on the DOE, but more interestingly, for response surface modeling and optimization. Because what do you do if your optimization says pick a buffer size of 1.2?

AUDIENCE: Oh, I see.

DUANE BONING: So I think that's an interesting twist you'll have to deal with. So you might actually have to sort of come up with some, on that point, some strategies for rounding off and maybe assessing both the high and low round off points to see which one's actually better. That might be an example strategy. OK, keep going.

AUDIENCE: OK, thank you. OK. That's the three stages that we summarize. First stage is the advanced SPC. The second stage the DOE. And the last stage is response of this model.

DUANE BONING: When you say stage, does it have to be in that order?

AUDIENCE: I think yeah. First because you try to plot the control chart for the service level. Then you find the problem. Then you start to focus on the effects. That's stage two. And after that you do the optimization. That's what we think should be order.

DUANE BONING: Because one thought maybe to think about is whether switching that order might make sense. I mean, in one scenario, I think you're right. And in one actual sort of real life scenario, you may have SPC charts already up. And then you go out of control. You see you're out of spec somewhere. And then if I understand your approach here, you're saying to figure out what the problem was, you want to do a DOE and solve the problem.

That might be an interesting scenario, but often the other way around is do a DOE up front so you actually know what you should be monitoring. You actually understand larger the basic relationships, what's significant, what you should be monitoring, where there's a likely actual effect. And then help use the DOE to set up the SPC.

Or maybe the sequence-- maybe what I would do, maybe it goes the other way around. Stage four, revise SPC. Maybe that would be a good compromise. Because I think once you've built a DOE, now you've got a lot more information about what the-- you've got a more solid model, a more functional model. And you can use that to perhaps set better limits on SPC and decide what parameters are most important. So I like that.

AUDIENCE: That's right, we can revise the control chart parameters. Yeah, that's right. And I think we'll use CellSim. It's an Excel based simulation software to build a model, supply chain model, and used Minitab to follow data analysis. OK?

DUANE BONING: OK. Great. So this is all going to be-- do you actually have some company data or are you going to pretty much mostly use the CellSim model?

AUDIENCE: Next slide. You can change to next slide.

DUANE BONING: Ah, good, yeah.

AUDIENCE: Yeah. So for the service level, I think we will get from the simulation. But the input from the company. Because we can't get the service level data from the company.

DUANE BONING: OK. So if I understand what you're saying is you might actually try to model kind of a real line with some of the parameters that tell you how that line is structured and then use that to do simulations and generate service level?

AUDIENCE: Yeah, that's right.

DUANE BONING: I see.

AUDIENCE: So the input from company, yeah.

DUANE BONING: OK. But you have no output data from the company? It's all going to be simulation?

AUDIENCE: Yeah, that's right.

DUANE BONING: OK, OK. That certainly gives the advantage I talked about today of synthetic data. I mean, it's nice that you've got a realistic scenario based on a company scenario for setting up your simulation. But you have lots of latitude in your synthetic data to actually-- I would encourage you to consider exploring intentional perturbations. You can introduce different faults into your line and see what happens both in the SPC and in the model. You've got a lot of flexibility because you're doing the simulation.

And you can play with and say what happens if this piece of equipment goes down or the mean time to failure triples. And I didn't know that. How does things improve? Or mean time to failure becomes 1/10 what it was. I'm failing much more often. So I think you can construct very nicely a lot of different scenarios based on your knowledge from sort of factory modeling that might be relevant ones to study with a DOE approach. So I think that's rich. That gives you a lot of opportunity. So that's good.

AUDIENCE: OK. I think that's the end of the slides.

DUANE BONING: Yeah. Hayden, do you have-- you want to-- here, why don't you?

GUEST SPEAKER: We're just thinking more about additional things you might explore. So let me just make sure I understand how the Excel model works. I mean, it sounds like you actually have a function that determines y as a function of all these input variables. And you know what that function is.

DUANE BONING: Well, it's very complicated.

GUEST SPEAKER: But you know what it is. So if you're doing DOE, you're trying to build a model, but you already have the model. So I'm not sure the timeline. Basically, you're going to be trying to get to the model--

AUDIENCE: Actually you don't have to the exact model. The service level relates to all of the parameters. But you cannot derive a model.

GUEST SPEAKER: You can't derive a-- so how is this exact model? Is it a lookup table or something?

AUDIENCE: We will simulate the supply chain and we run the simulation to get a production lead time, a production cycle time. That is influenced by the buffer size, by the MTTF, MTTR. And we will check for the production lead time plus transportation lead time. If this is less than the lead time that you promised your customer, then that means you satisfy the customer. And if the production lead time plus the transportation lead time larger than LC, then it will account for one [INAUDIBLE]. That's what we generate the service level.

GUEST SPEAKER: OK, I see. So you want some approximate way of predicting what it's going to be. OK, fair enough.

DUANE BONING: I think Hayden's point is if you have, in general, very often you have very complicated models that may take a long time to actually run any one point. And so I think that's lurking in your mind. What you would like to do is build simplified approximate derived models off of a complicated long simulation time sophisticated CellSim model. Is that right?

GUEST SPEAKER: OK. Great.

DUANE BONING: I mean, even though CellSim may actually be fairly fast. In what I've seen in many of these factory simulations, running the factory simulation itself may take hours or days. And that's where you might actually be building a simplified approximate model.

AUDIENCE: I have one question here. I'm not very sure about the demand data. But seems like the CellSim can only have a normal distribution, an exponential distribution input. So if the normal distribution is not very appropriate for the demand, what can we do?

DUANE BONING: Good question. I think if it can at least do two different distributions, one interesting thing would be to see how sensitive the simulation is to whether it's-- you said a normal distribution and the other was a Poisson? Or what was the other one?

AUDIENCE: It's a exponential.

DUANE BONING: Exponential. I think usually an exponential is what's often assumed for these kinds of arrival processes. But it would be interesting, actually, to see how much-- that might end up being a nice question for you to ask you and try with your data and your simulation. How do things differ as a function of not just the parameters of the distribution but the type of distribution? It might not matter at all. That might be crucial. That might be the most important thing. I don't know. Good. Any other questions on this one?

AUDIENCE: No, we're fine.

DUANE BONING: Great. Go for it. It sounds interesting. I like how this integrates your other classes and your other learning too.

AUDIENCE: That's the inspiration from another class.

DUANE BONING: Exactly. OK. All right. So let's see. I had talked with-- is there-- who wants to go next? Other people have not sent slides, but let's see. Who else have-- I sent email to I think everybody kind of late last night or early morning for you guys. They're all mixing together. Who wants to talk about one of their projects next? Because we still have about 20 minutes on the video time.

Just talk me through your written thing and some of the comments I sent. How about the-- let's see. I'm looking at the one for the iron sole plate. Who's kind of a lead person on that? By the way, I liked this one, because I got to go and tour the Phillips factory when we were there and see some of these things being made. So I'm like, oh, this is cool. This is great.

AUDIENCE: Well, so actually one we have sent to you. The other part we have in mind is that the company is going to run some intermediate points. But it turns out to be otherwise. So what we have is that it's a five parameter. So they are trying to build a prototype. And they're trying to find what is the output. So that's the last one that actually blow out [INAUDIBLE]. And then yeah, so what we have is actually they actually run a full DOE. So that's what we have.

DUANE BONING: Is that data already available?

AUDIENCE: Yes it is. So then what we have is a full DOE. So that's to the power of 5. So that's 32 runs. And then each run, there is three radicates.

DUANE BONING: OK, good. Yes. Are all of those parameters basically continuous parameters? I on the diagram you mentioned different diameters and heights. Are all of those continuous parameters? Are any of them sort of discrete choices? I know they only did two levels of each. So in some sense, they are discrete. But if you were building a response surface model, would you be able to actually have all of them be continuous parameters? Or would there be a mix of some arrangement of holes and some diameter that was continuously varying?

AUDIENCE: I think it's continuous parameters. And then I think maybe some of the inputs are not so measurable. So maybe they try to-- since some of the inputs look at the noise input. And then they may use the [INAUDIBLE] to do the robust design. And maybe it is a potential area, because right now there is still some problem with the data connection. And then we are trying to figure the problem, all those days. So actually the full area has not been exactly designed yet. So that is one of our problem. Yeah.

DUANE BONING: OK. But you think you'll have data?

AUDIENCE: Yeah, we have several data. But we do ANOVA, and then we find out that our value is not exactly so high. It's just a little larger than 0.5. So I think maybe the data is not complete. So maybe we need more data.

DUANE BONING: OK. Well, if you get more data, that's great. But it's not just the r squared, but I assume your ANOVA does show some significance to a model? OK. Good, good. Then I think you've got stuff to work with. Now, in the email--

AUDIENCE: Professor?

DUANE BONING: Yes?

AUDIENCE: I actually think that it may not be very significant from the model. Yeah, I do think that it's not that significant. And I actually talked to the company to see why they are not trying to find more points, since they are not really able to find a good fit. But the response that I got from them is that to run the experiment, they actually need to build a prototype, which is very expensive.

And as a result, they are not planning to run that. Actually, I thought that is a good idea maybe to run a midpoint. I was not able to convince the company for doing that. Yeah, so that's a problem they're facing. That's a lack of data for this study.

So at the same time, we have also [INAUDIBLE] exploring some other alternatives. So the [INAUDIBLE] we got is actually-- so in the parent case, it's a product design optimization. So then the next one that we're looking into is a process optimization. So in the process optimization, basically it's a process line where there is actually different sole plates that are actually they're produced in a different period of time. And then they also manage it differently. So that's also like the [INAUDIBLE] influence.

Like for example, we are also trying to see, for example, whether we can see there's any correlation between the defects. The defects are [INAUDIBLE] or things like this. So we're trying to see whether we can find correlation between different layers or even spatially. Because it's arranged in terms of rows and columns, I think that's five rows and seven columns. So in each carrier, that's 35 units. So then we thought that probably we can do a nested variance analysis on that.

So one aspect I'm looking into is because we also have the data for the different stations. So one thing that we can look into is which station is the main contributor. So that's one. So secondly is between the shift of the timing. So for example, every hour, so an hour and maybe the different shift, maybe one shift an hour, things like this. So then we can also do a nested variance based on the shift timing.

DUANE BONING: Right. This sounds very interesting. I don't think we have anybody else doing a nested variance kind of analysis. We have several DOEs. So in the overall class, I think this would be very interesting to have this different project you're talking about. Now, for that one you've already got the data. It sounds like it's good. They have lots of manufacturing data in that case. Is that correct?

AUDIENCE: Yeah. There are [INAUDIBLE] data. But then the person who ran the data, they actually are not so careful to take note of every parameter. So then initially I actually talked to the company, the person, I asked whether we can get all the data and then we can combine all the data together and then we'll be able to get a very rich set of data. We can do all the analysis. But it seems that the data that we get is missing in pieces.

For example, they are doing hypothesis testing kind of thing. So basically, they are doing a p-test. So it's basically a hypothesis testing where they actually check each parameter to see whether it's significant or not significant. And so while they are doing that, they do not consider any other parameters. So although the process is controlled by many inputs, but at any single, one time, they only care about a particular set of a parameter. So as a result, we are not able to combine all the data together as one big cell of data and do as much analysis as we would like to have. But we thought that that would be an interesting project to go about.

And at the same time, I also think that we are trying to continue to explore a product thing and be a process thing. So as to get a comparison-- because we thought that this real life situation, we also like to see, is the point of view for a product DOE and process DOE-- how is it different. Also we understand that there are some limitations for the product DOE, because it pretty difficult for the company to run, to make a new prototype each time compared to a process DOE where they can change a setting rather easily. Yeah, so there is a plan.

DUANE BONING: I think this sounds like a very good plan. Because I wanted to come back to the product DOE, because I think-- so first off, I think with your team, you've got a good team of four people. So being able to look at both of these problems is great. And that includes a case study and examples of real life limitations that are there in both situations and recommendations you could make in both cases to the company that might help them get more value out of the activities that they've got to be able to make better decisions in the future.

So going back to the product one with the DOE, the two to the fifth DOE, I think it would be very interesting to somehow include, maybe you just make it up, but include what more you could do if you had a center point. And show if the center point looks like this, look how different your decision would be in design optimization. Versus if your center point looks like this, then the model looks completely different, and this is what you should do.

In other words, show how important that center point might be to their thinking. Could be a nice part of this that actually might help make the case persuasively to the company that in their future it might be very valuable and worth the expense of building just one more prototype. If they're already building 32 different prototypes with some missing data perhaps, one more would tell them so much more potentially. So I think that could be very interesting. And I'm glad, I like the plan of trying to do some of both. Do I have that right?

AUDIENCE: Yeah. So I actually think that when they made the prototype, maybe they choose the high and low because it's easier for them to make the prototype rather than making a center point. So I do understand that a center point has the advantage of being able to fit models and things like that. In the technical aspect, maybe they are limited by machines or things like these. Maybe they have a template where using a template they are able to get a plus or a minus. But they are not able to get the midpoint. So I'm not sure-- perhaps, is it possible that we discuss points like this?

DUANE BONING: Yes. Absolutely. Those are very realistic limitations. You might find that a couple of parameters can have center points. Other ones can't. And so it's a mixed center point. And so that's a good issue. Absolutely. All of these kind of realistic constraints that drive-- that help limit or drive them to the decisions that they made but also exploring is there some room to do some more? So yes, that's a very good point.

AUDIENCE: OK, thank you.

DUANE BONING: Good. Well, this sounds exciting. Get the data as quickly as you can even with all of its ugliness and get going on that. And then if issues come up, send us email and we can see what we can do to help.

So I think I just have a short period of time now. Let's see. I saw you in the back while the camera was focused there, Priyanka, I saw you in the back there. So you guys project. [INAUDIBLE] Stanley. Thoughts on your project? First off, I really like the ECM aspect, because I've got a student working on electropolishing or electrochemical mechanical polishing for semiconductor stuff. So the student has been learning a little bit about electrochemistry. So I saw this and I thought it was really interesting.

AUDIENCE: The story is sad but true. It actually happened in my undergrad university. And I got rescued on a boat, and we went to rescue the machine, and it was bolted on a concrete slab and it was under 10 feet of water.

DUANE BONING: Do you have pictures?

AUDIENCE: Yeah, I do.

DUANE BONING: Oh good. Excellent. Now, in terms of-- I think I sent a little bit of information back in the email, because it looked like, if I'm remembering, you wanted to do a couple of things to see was the equipment restored and operating back to its original state. And I understood--

AUDIENCE: That's correct.

DUANE BONING: Yeah, I understood that part. And then the other interesting piece was building an analytic model. And using that, it sounded like-- well, tell me what you would like to do with the analytic model.

AUDIENCE: So initially, the experiment was done to get data to substantiate an analytical model. So the model was based on the Faraday's laws of electrolysis. But that related only a few of the parameters. So we were basically trying to incorporate the flow rate term into that equation. But the focus kind of got shifted because the machine stopped running midway. And we spent a very long time trying to get it back on its feet. And the readings that we took after that, so I can sort the readings on the basis of date.

We were four students working on that project, and we took a large set of readings. So I can sort it by date and find out what's before the flood and what's after the flood and to find out if really a shift had occurred after that. When we were using it, we never treated it as a separate set of data. We combined and indiscriminately averaged out the whole thing. So that's something that in hindsight we probably should have looked at.

DUANE BONING: OK. I think that's sort of the core of the idea for your group. And I like that. What I'm thinking about is if there's one or two additional ideas that you might be able to explore given the fact that you've also got this analytic model. So it's going back a little bit to this idea of synthetic data or looking at your analytic model and maybe even doing things like a sensitivity analysis on the model.

Saying based on the model, these are the parameters I think we would be most sensitive to in a DOE. Some of those kinds of things. I think it might be neat to just brainstorm or think about things you might actually use the analytic model, since you've got that also. Even though you only know some constants from physics and some you were trying to fit to your data. And that whole piece sounds very interesting.

AUDIENCE: OK. The data we have basically consists of varying three input parameters. Feed rate, flow rate, and voltage. And the output is the MRR. So the way we measured MRR was we took the difference in the weight of the slab.

DUANE BONING: Yes, that sounds good.

AUDIENCE: And the time for the machining. And so we have the MRR for about-- we have about 50 to 60 readings that we've taken.

DUANE BONING: Good. Yeah, that was my other question in the email was how much data you had. Because it sounded like each run is fairly long. Yeah, so you've got lots of students working on this for a long time.

AUDIENCE: Yeah, we have four students working on it almost around the clock. So quite a lot of readings.

DUANE BONING: OK, that sounds very interesting. And it will be an interesting story to hear as well.

AUDIENCE: Yeah, I'll try to put in some pictures.

DUANE BONING: So if other questions come up, let me know, especially as you think a little bit about-- your group thinks of ways to use the analytic model. OK.

AUDIENCE: A problem with the data is that the flow rate, we couldn't vary the flow rate on digital control. So there's this whole set of six valves that we had to manipulate to set a flow. And so all the other factors are set at fixed levels. But the flow rate is kind of an average value that it kind of oscillated about. So I'm not very sure how to deal.

DUANE BONING: Yeah. Well first off, that's an interesting question we haven't talked a lot about in class. So it'll be interesting to raise that and talk about it. That basically, there's a spread on your input. There's variation on your input. And some of the methods we've talked about might include just recognizing that and recognizing that that might propagate through the data. Because often we pretend in all of our DOEs that when I pick the input, it's rock solid and it's never rock solid. Usually it's very well, tightly controlled. But often there is variation in that.

And the basic approach, that might be a place where looking at your analytic model might actually give you a sense of how big a perturbation variations on the input might produce in the output. So you could actually get an estimate of roughly how important it is to control more accurately to the inputs. Things like that. So I think that's a very interesting real life problem. And that's just one suggestion.

AUDIENCE: There's another issue. The flow rate was the cause of all our troubles throughout. Because we couldn't find a proper correlation analytically for the flow rate, because the voltage and the other parameters could be easily incorporated into the analytic model just using an extension of Faraday's laws. But the flow rate was a little difficult to capture.

DUANE BONING: So it's not in the model.

AUDIENCE: It's not really in the model.

DUANE BONING: I see. OK.

AUDIENCE: And the way we calculated flow rate, the way we've measured our flow rate is we've got a maximum value, a minimum value, and a most common central value. That's the way we've measured the flow rate.

DUANE BONING: I think it'll be interesting to see how you dealt with that or recommendations on how to deal with that. That's good. I think these real life challenges are a good thing. Thanks. I think that one sounds good.

Now let's see. I'm trying to see. Have we covered all the projects? Or I've left somebody out, right?

AUDIENCE: [INAUDIBLE]

DUANE BONING: Yeah. I actually have to run. Let me see. Which one are you?

AUDIENCE: We're the team working on the injection molding data.

DUANE BONING: Ah, yes, OK.

AUDIENCE: Die casting, sorry. We're working on die casting process. And we got our data from a paper.

DUANE BONING: Ah, yes. So we had some good email on that. Right?

AUDIENCE: Right, right.

DUANE BONING: Yeah. I think you guys are in very good shape. There were some suggestions from Hayden. And I sent an email also.

AUDIENCE: That's right.

DUANE BONING: So did you have follow on questions? I think you guys-- that's good data and you can do a lot more with it. So any additional thoughts or questions?

AUDIENCE: So we have a question concerning inputs. There are three inputs. Two of them they have three levels. And the other one has only two levels. Does it matter when you do-- when you do the ANOVA, ISM, all these analysis? Because the levels are different for inputs.

DUANE BONING: The quick answer is no. But it should be interesting for you to show why that still is good and that's no problem. Just think about it. I mean, in terms of some model dependencies, if you have three levels, you might be able to do quadratic in that model parameter. But if you only have two levels in others, you'll only have linear terms for that. So that's fine. That's OK.

AUDIENCE: OK.

DUANE BONING: OK. So if questions come up as you're working on that one, send us more email. Because Hayden had lots of ideas. There may be more than you can do. But I think there are some very interesting ideas, because there's replicate data. Variance might be interesting to model. Several neat things you can do with that.

AUDIENCE: OK. We'll follow with emails.

DUANE BONING: OK. Good. And we can also again on Thursday if there's more questions that come up. After class we can talk briefly. Because by then you should hopefully be well into the project.

AUDIENCE: Thank you.

DUANE BONING: Thank you. See you guys later.