Trellis Display for Modeling Data from Designed Experiments

Montserrat Fuentes[1], **Bowei Xi**[2,*], **William S. Cleveland**[3]

[1]*Department of Statistics, North Carolina State University*

[2,3]*Department of Statistics, Purdue University*

Abstract: Visualizing data by graphing a response against certain factors, and conditioning on other factors, has arisen independently in many different types of applications. One case is the interaction plots used in the analysis of data from designed experiments; these plots show conditional dependence based on the output of methods and models applied to the data. Trellis display, a framework for the visualization of multivariable data, allows conditioning to be readily carried out in a general way. It was developed initially in the context of data sets with a moderate or large number of observations to support the conditioning. This article demonstrates through examples that trellis is also a highly useful visualization framework for designed experiments with a small number of runs. This includes highly fractionated designs: the number of runs is just a fraction of the number that would result if all combinations of the factors were run. Trellis succeeds because it allows the visualization of conditional dependence, not based only on the output of models and methods, but also based on the raw data directly. This provides incisive assessments of patterns of dependence in the data that can guide the choice of the final methods and models used in the analysis, contributing substantially to the understanding of the data.

Keywords: Data visualization, statistics, machine learning, fractional factorial design, interaction plot.

1. INTRODUCTION

Trellis display is a framework for the visualization of multivariable data [1, 2, 3, 4]. One capability is a mechanism to study the dependence of a response variable on predictive factor variables. It does this through a mechanism for visualizing the dependence of the response on one set of factors, given values of other factors. The visual design allows assessment of how the conditional dependence changes with the given values. This can be done for both the raw data and for the numeric and categorical output of mathematical methods and models applied to the data. The former is particularly valuable, allowing an understanding of the patterns of dependence in the data free of assumptions about the patterns. Figure 1 is a trellis display that shows the dependence of a response on one factor given three others using the raw data. The data and the display will be described in Section 2.

The success of trellis display in the visualization of multivariable data has led to implementations in a number of software systems including S-Plus [5], R [6], and Tableau [7].

Trellis display was originally developed in the context of moderate, large, and very large data sets, and has been widely used in this context. This article reports the results of an investigation of the use of trellis in analyzing data from designed experiments that result in small data sets. While the sizes of data sets have grown dramatically in many domains, controlled experiments with a small number of runs are still commonplace in the many settings where a single run is very costly [8]. It is often the case that designed experiments are highly fractionated: values of each factor are chosen, but the experiment is run on just a small fraction of the number of possible combinations of the values of the factors.

The question in our investigation was whether success using trellis conditioning methods and visual methods would be inhibited by the limited number of runs and fractional experimental design. It seemed quite possible to us that for such data the number of observations in each subset resulting from a multiple conditioning would often be too small for patterns to be seen. Our investigation, over a long period, has consisted of using trellis in the analysis of data from many experiments reported in the literature, some arising in our own work.

Displaying data by conditioning has surfaced independently in a number of places and for many different types of data [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. The experimental design literature contains a long history of data visualiza-

[*]*Correspondence to:* Bowei Xi (xbw@purdue.edu)

tion [20, 21, 22, 23, 24]. Included in this design literature is a widely used method of visualizing conditional dependence by the *interaction plot* [13, 14, 25, 26, 27, 28]. This visualization tool shows interaction effects based on the output of methods and models applied to the data. Its use in practice has chiefly been to show two-factor interactions, but in principle, higher-order interactions can be shown [13, 14]. The trellis mechanism described here is much broader, allowing, as described above, conditional dependence to be studied for the raw data as well.

Results of this article are conveyed through describing analyses of three data sets from designed experiments in Sections 2 to 4. The data sets are representative of what we have seen more generally. Section 5 is a discussion of results.

In the course of the discussion of trellis conditioning in Sections 2 to 4, another important matter for all data visualization is investigated in the context of designed experiments. Methods of data analysis can be divided into two categories [29]: (1) mathematical methods and models in which formulas are computed to produce numeric and categorical output; (2) visualization methods whose output is visual displays, either of the raw data or of the output of mathematical methods and models. The analysis of variance (ANOVA), used pervasively in the analysis of experimental data, is a mathematical method for answering specific questions about terms in a model for the data, and thus is a model building tool. We discuss the relative power of ANOVA and the trellis visual methods.

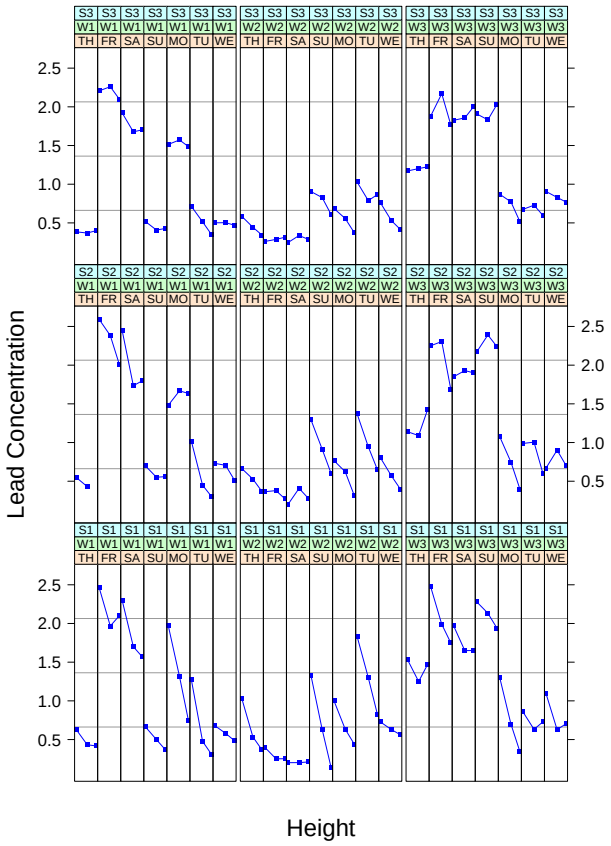

Figure 1: Trellis display of L against H given D, W, and S.

2. TRELLIS DISPLAY OF LEAD CONCENTRATION DATA

Lead concentrations at a site next to a major roadway in Ohio were measured and analyzed in an experiment to determine their spatial variation [30]. The concentrations were measured at 9 positions on one side of the roadway. There were three setback distances from the roadway: 2.8 m, 7.1 m, and 21.4 m. There were three heights: 1.1 m, 6.3 m, and 10.5 m. The 9 positions, each height combined with each setback distance, form a 3 by 3 vertical spatial grid. Measurements were made at the nine positions for 21 consecutive days. Each measurement is an accumulation of lead over a period of 24 hours. Thus the data consist of 21 daily lead measurements at each of the 9 positions; one observation is missing. For such data we would expect the lead concentrations to be affected by a host of factors: meteorological conditions; traffic, which has a day-of-the-week effect; and spatial position.

The lead data consist of five variables: (1) lead concentration, L; (2) setback distance, S; (3) height, H; (4) day-of-the-week, D; and (5) week number, W. D and W describe time — that is, the day — but do so in a way that allows for a day-of-the-week effect. There are $3\times3\times21-1=188$ measurements of each of the five variables.

Figure 1 is a trellis display of L against H given D, W, and S. The display consists of 3×21 = 63 panels arranged into 21 columns and 3 rows. Each panel has a scatter-plot of L against H given D, W, and S. The strip labels at the top of each panel indicate the values of the three conditioning variables. S changes with the row; for row 1, the bottom row, S is smallest, and then increases as we go up the rows. As we go left to right through the columns of each row, we go in order through the days. In a similar manner, Figure 2 is a trellis display of L against S given D, W, and H.

Figure 1 shows that L tends to decrease as H increases. The decline as a function of H lessens as S increases. In other words there is a spatial effect with an interaction between H and S, which is not surprising.

Figure 2 shows the spatial effect in a different way. There is mixed behavior in the dependence of lead on S. For the smallest value of H, L decreases with S. But for the middle value of H, L typically first increases with S and then decreases. For the largest H, lead occasionally has the increase-decrease pattern for about 1/3 of the days, most of them days with large concentrations, and is relatively stable for the remaining days. This behavior is consistent with air transport mechanisms. Lead is emitted at ground level from automo-

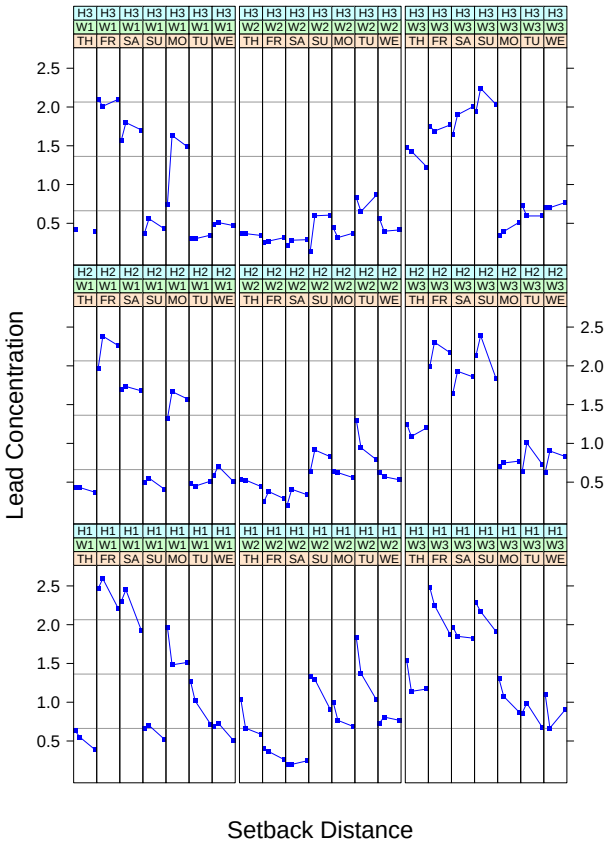

Figure 2: Trellis display of L against S given D, W, and H.

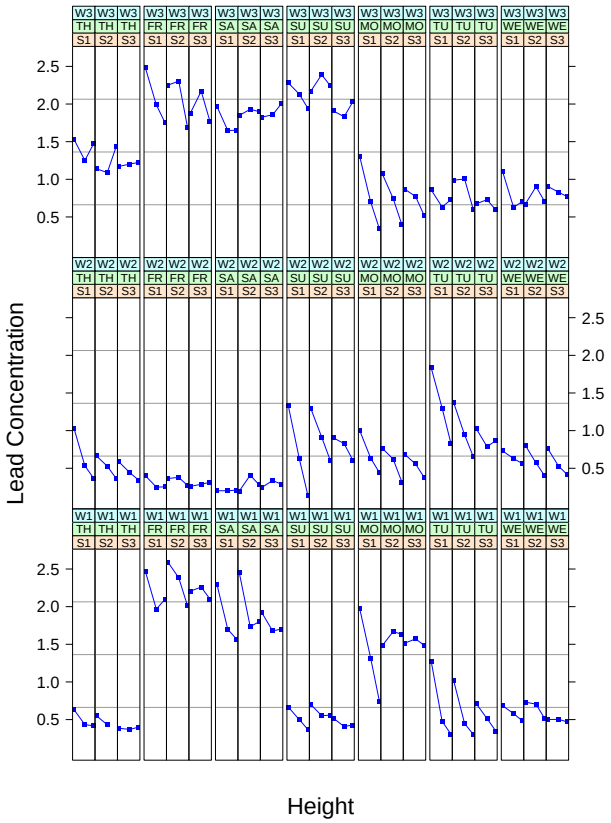

Figure 3: Trellis display of L against H given S, D, and W.

bile tail pipes. The closest of the 9 monitors, the one with the smallest values of H and S, has the largest concentrations because it is close to the pollution source. From the source, the lead is carried laterally by the wind, spreading upward as it moves. This plume-like behavior can cause the concentrations to be relatively small at the higher monitors with the closest setback.

The arrangement of the panels in Figure 1 allows study of 3 collections of patterns, one collection for each row. This provides a comparison of the patterns of dependence of L on H as S changes. Suppose, however, that we want to study the 3 patterns for each day, and then compare the 21 collections of daily patterns. This is a more difficult task in Figure 1 because the 3 panels for each day are arranged vertically in such a way that we have a reduced ability to effortlessly visually assemble the 3 patterns. In Figure 3, the panels have been rearranged to facilitate the study of the daily patterns. Now the 3 panels for each day are juxtaposed horizontally, and each row is now the data for one week. The panels in the bottom row are week 1, the panels in the middle row are week 2, and the panels in the top row are week 3.

Figure 3 shows that the within variation of the 9 measurements for each day is much smaller than the variation across days. The cause is changing weather conditions which have a substantial effect on the concentrations. Rain and high wind speeds reduce concentrations and low wind speeds increase concentrations; the 9 measurements on a given day are affected in the same way by the weather. Weather conditions are correlated through time; fronts move in and persist for a few days. This is visible in the concentrations; the figure shows that collectively, low or high concentrations persist across days. By contrast, D does not appear to be salient in that there does not appear to be a systematic day of the week effect in the data whose magnitude is more than minor compared with the weather effect. The conclusion is that there is a strong time correlation in the concentrations, across days, which is not surprising.

Figure 3 also suggests spatial correlation, likely induced by the weather effects interacting with spatial position; the plume has different shapes depending on the meteorological conditions. For each of the three setbacks on one day, there are two differences in L with height: L for H1 minus L for H2, and L for H2 minus L for H3. There are 6 such differences for each day. The 6 differences appear positively correlated. When one difference is larger than average, the others tend to be larger as well; a similar statement holds for smaller than average. A robust estimate of the correlation matrix [31] of the six dif-

ferences is shown in Table 1. There are indeed high positive correlations as expected from our observations of Figure 3.

Table 1: Robust estimtes of correlation coefficients of differences.

	S1:1-2	S1:2-3	S2:1-2	S2:2-3	S3:1-2	S3:2-3
S1:1-2	*	0.61	0.60	0.28	0.75	0.30
S1:2-3	0.61	*	0.75	0.51	0.91	0.73
S2:1-2	0.60	0.75	*	0.16	0.89	0.34
S2:2-3	0.28	0.51	0.16	*	0.26	0.85
S3:1-2	0.75	0.91	0.89	0.26	*	0.52
S3:2-3	0.30	0.73	0.34	0.85	0.52	*

In the source publication for the lead concentration data [30], ANOVA was used as a model building tool. The author states: " One potential problem is that the lead concentration data may be serially correlated and this could interfere with the assumption of independently identically distributed errors. This problem was minimized by introducing the effects of day [day-of-the-week], week, and their interaction to isolate the variations due to the effects of time and hence serial correlation." Table 2 is an ANOVA for the same effects fitted by the author. The missing value has been estimated by maximum likelihood, but is treated as not missing for the purposes of carrying out the ANOVA.

Table 2: Analysis of variance for lead concentration data.

Effect	DF	SS	MS	F	P
S	2	0.30	0.148	5.35	0.00
H	2	2.97	1.486	53.70	0.00
D	6	16.64	2.773	100.22	0.00
W	2	19.22	9.611	347.41	0.00
S × H	4	1.00	0.251	9.06	0.00
W × D	12	38.13	3.177	114.85	0.00
S × W	4	0.07	0.018	0.65	0.63
H × W	4	0.18	0.045	1.63	0.17
Error	152	4.21	0.028		

The significant effects shown in Table 2 are S, H, SH, W, D, and DW, so these effects provide a modeling of the data. The model is quite simple. Let L_{DWSH} be the lead concentration for day of the week D, week W, setback S, and height H. Then the model is

$$L_{DWSH} = \mu + \alpha_{DW} + \beta_{SH} + error$$

where

$$\sum_{D=1}^{7}\sum_{W=1}^{3}\alpha_{DW} = \sum_{S=1}^{3}\sum_{H=1}^{3}\beta_{SH} = 0.$$

In other words we model time effects with 21 coefficients that sum to 0, and we model spatial variation with 9 coefficients that sum to zero.

Figure 4 is the classical interaction plot discussed in Section 1 for the H by S interaction, the spatial effect. The values

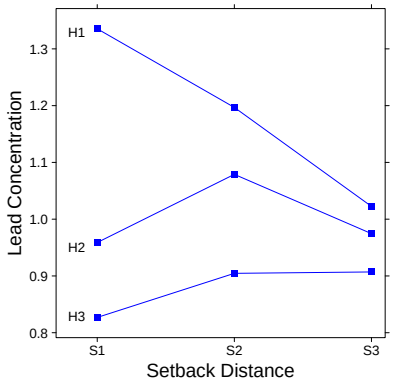

Figure 4: Interaction plot for L by spatial location ($H \times S$).

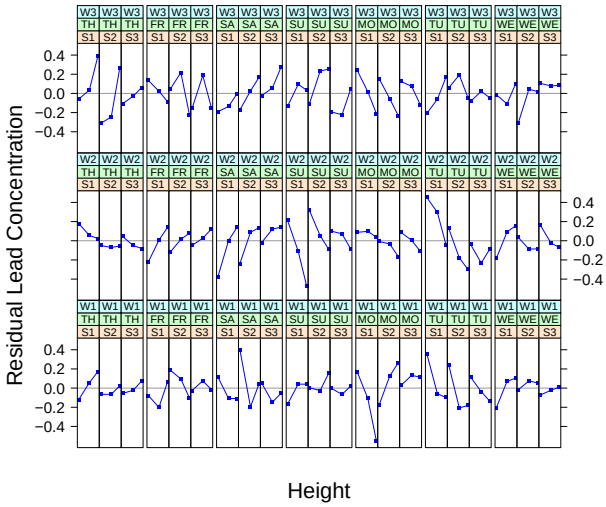

Figure 5: Trellis display of residual L against H given S, D, and W.

plotted are $1+\hat{\mu}+\hat{\beta}_{SH}$ for S = 1 to 3 and H = 1 to 3, where $\hat{\mu}$ and $\hat{\beta}_{SH}$ are the least squares estimates. We see a summarization of the effect that was observed from Figure 2. These are the effects in the data as seen by the ANOVA and the resulting model. Note that the interaction effects span a range of about 0.5 on the lead concentration scale.

The residuals from the model, $\hat{R}_{DWSH} = L_{DWSH} - \hat{\mu} - \hat{\alpha}_{DW} - \hat{\beta}_{SH}$, are the remaining variation in L_{DWSH} after subtracting the least squares model fit. The $\hat{R}_{DWSH}$ are the variation not explained by the model. Figure 5 graphs the $\hat{R}_{DWSH}$ in the same way that the raw data are graphed in Figure 3. We can see clearly the correlation revealed in Figure 3 and Table 1. Now however we can judge the magnitude of the correlation compared with the model fit. On many of the panels the values span 0.5 or more, very significant compared with the range of the effect, also 0.5. This means that the change in the spatial effect with the meteorological conditions is quite substantial.

The trellis plots of the lead concentrations show us that the simple ANOVA model of the data misses an important effect with a large magnitude. Furthermore, the effect is quite complex, an interaction between meteorological conditions and the spatial pattern — a changing plume. Unfortunately, the current data are not sufficient to estimate this effect. Given the salience of the effect we must conclude that the experiment has not succeeded in its goal, which is an understanding of spatial variation. Success would require detailed information about the meteorology, or a large enough number of days to provide a representative sample of meteorological conditions.

3. MODELING DATA FROM A RESIST EXPERIMENT

Computer chips are manufactured by creating them on wafers, circular or near circular silicon disks that are coated and processed by hundreds of steps. Then the wafers are cut up to produce the individual chips. One manufacturing process is etching: coating a wafer with a resist solution, exposing the resist to light to create the chip features, and then placing the wafer in a developer solution to remove the exposed areas of the resist.

In an experiment run to improve the resolution, processing of the wafers involved the following measured steps [32]: (1) coat a wafer with a resist solution containing a new photoacid generator, whose amount, or *load*, was varied in the experiment; (2) use one of two *solvents* in the resist solution; (3) expose the coated wafer to 248 nm light shone through a photo mask; (4) bake the wafer at a *temperature* that was varied and for a *duration* that was varied; (5) develop the wafer for 60 seconds in a developer solution. The response in the experiment is the clearing dose, C, which is the exposure dose in mJ/cm^2 needed to clear the resist in a cross-shaped region 100 μm by 150 μm. The following are the factors in the experiment: (1) L: load of the photoacid generator in the resist (% wt); (2) S: solvent, with value -1 for solvent 1 and value 1 for solvent 2; (3) T: temperature of bake cycle (° C); (4) D: duration of bake cycle (sec). The fractional experiment design consisted of 36 runs in four blocks of nine runs chosen to optimize estimation of a response surface for L specified to be a full quadratic in the numeric variables and to have a solvent main effect and interactions with the linear terms of the numeric variables.

3.1. Analysis of Variance

Table 3 shows an ANOVA. The conjectured model, described above, is that used for the choice of the design. The quadratic term for T is significant but not for the other two numeric variables. The interaction of S with the numeric variables is significant only for L. The results are unintuitive. It is possible the design is not sufficient to estimate the effects. If we are to reliably estimate the effects we need more insight into the data than that given by the ANOVA. We need some good luck in the form of a simple model explaining the data, and we need methods that allow us to perceive the simpler structure if it exists.

Table 3: Analysis of variance for resist data.

Effect	DF	SS	MS	F	P
S	1	2193.36	2193.36	20.35	0.00
T	1	13323.44	13323.44	123.61	0.00
L	1	4977.83	4977.83	46.18	0.00
D	1	4054.75	4054.75	37.62	0.00
T^2	1	1091.47	1091.47	10.13	0.00
L^2	1	69.90	69.90	0.65	0.43
D^2	1	29.37	29.37	0.27	0.61
D × T	1	1004.99	1004.99	9.32	0.01
T × L	1	1455.48	1455.48	13.50	0.00
D × L	1	1048.62	1048.62	9.73	0.00
S × T	1	52.46	52.46	0.49	0.49
S × L	1	554.89	554.89	5.15	0.03
S × D	1	66.41	66.41	0.62	0.44
Error	22	2371.33	107.79		

3.2. Trellis Display of the Raw Data

We will use trellis display of the raw data to search for insight into the dependence of the response on the factors. Figure 6 shows intervals that will be used for conditioning on the three numeric variables D, L, and T. Each set of 3 conditioning intervals consists of low, medium, and high. Low and high in all

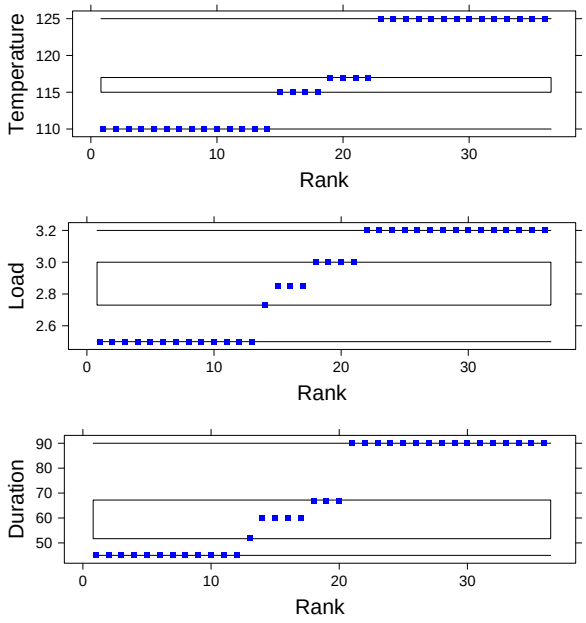

Figure 6: Conditioning intervals for D, L, and T.

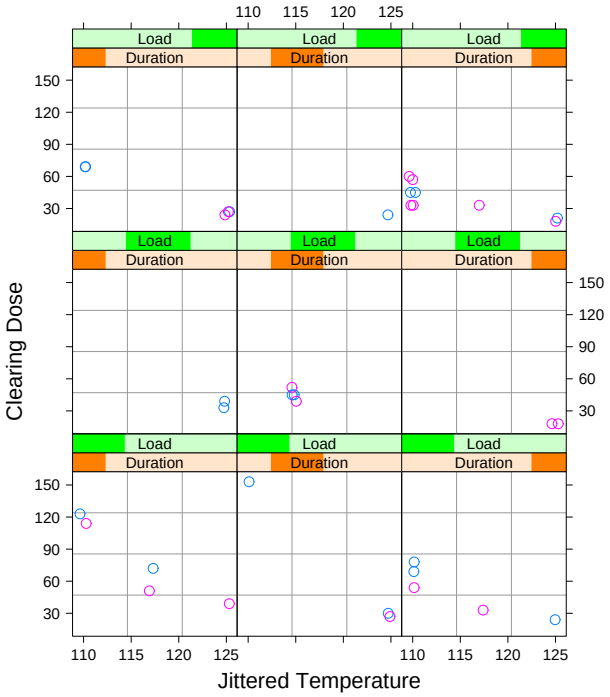

Figure 7: C against T given D, L, and S.

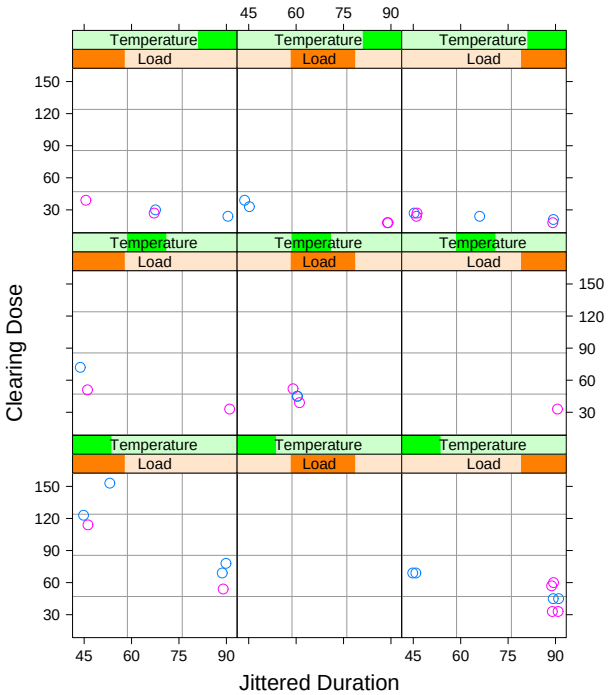

Figure 8: C against D given L, T, and S.

cases consist of a single value, an interval of length 0. Conditioning on S is simple; there are two conditioning categories, solvent 1 (-1) and solvent 2 (1).

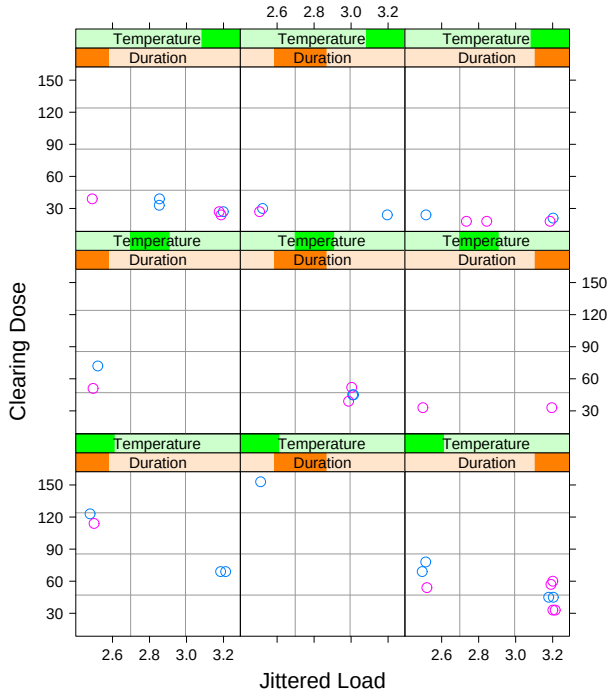

Figure 9: C against L given D, T, and S.

Figure 7 is a trellis display of C against T given D, L, and S. Each panel shows the values of C and T for those runs with D in one of its intervals and L in one of its intervals; and on the panel S is encoded by the symbol color; solvent 1 is magenta and for solvent 2 is cyan. To avoid exact overlap of some data points, a small amount of random uniform noise has been added to the values of T. The intervals of D are the same for all panels in the same column; as we go from left to right through the columns, the intervals increase. The intervals of L are the same for all panels in the same row; as we go from bottom to top through the rows, the intervals increase. The strip label for each panel contains a graphical portrayal of the conditioning interval. The strip has a scale but there are no tick marks to indicate the numeric values of the interval; the scale value at the left endpoint of a strip label is the minimum value of the measurements of the conditioning variable, the scale value at the right endpoint is the maximum, and the darkened bar shows the interval.

Figure 7 reveals important patterns. As either D or L increases for a fixed interval of the other, the pattern of values of C as a function of T shifts downward. In addition, the absolute value of the slope of the pattern tends to zero; so the absolute first derivative decreases with increasing C. The change in the derivative is quite large indicating a strong interaction. Furthermore, the positive second derivative also decreases as C increases; so the second derivative decreases with increasing C. Finally, there is a suggestion that overall, solvent 2 (cyan) leads to very slightly higher values of C than solvent 1 (ma-

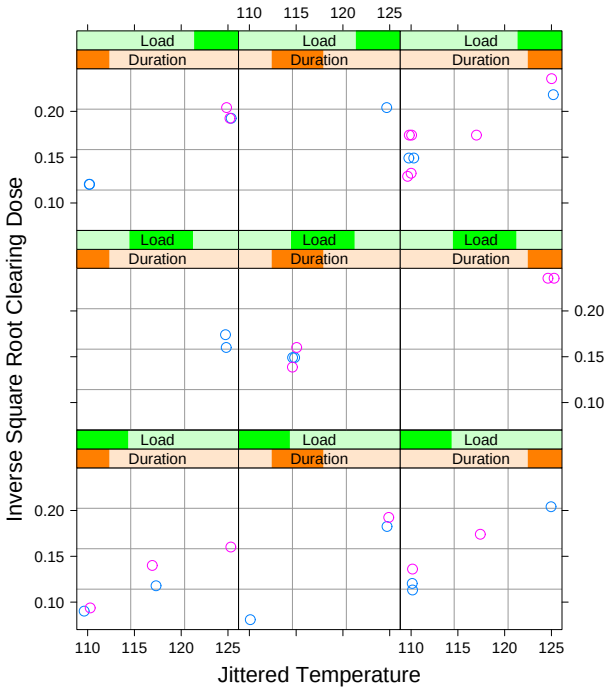

Figure 10: $1/\sqrt{C}$ against T given D, L, and S.

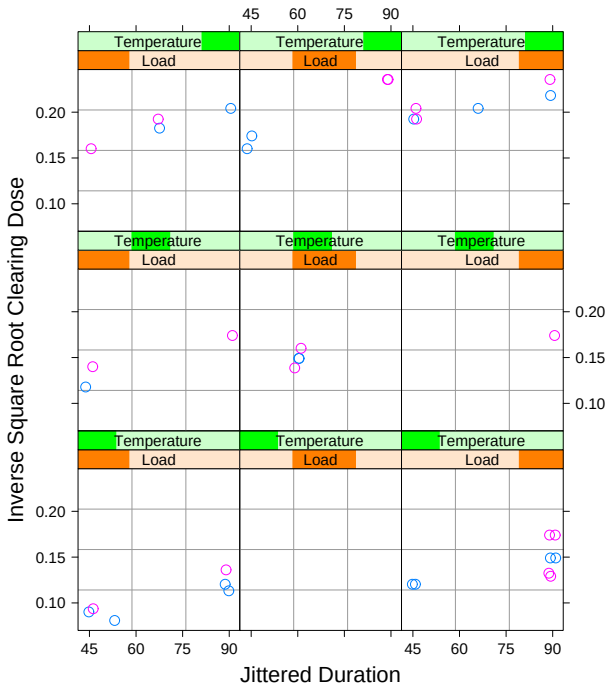

Figure 11: $1/\sqrt{C}$ against D given L, T, and S.

genta), but in addition, the solvent effect decreases as either D or L increases, which means that S is interacting with D, L, and T as well.

Figures 8 and 9 are further trellis displays of the data. The first is a display of C against D given L, T and S; the second is a display of C against L given D, T, and S. As a function of each of these two other numeric factors, C exhibits the same dependence of the first and second derivatives on the level of C, and the same dependence of the effect of S on the level of C, as seen in the previous two figures.

The complexity of the dependence of C on the factors is quite bad news. Unless we can find a simplification, more runs are needed because the design cannot support the number of interactions, at least to order 3, seen in the data.

3.3. Exploiting An Observed Regularity

There is one regularity revealed in the trellis displays that might be exploitable. The magnitudes of the first and second derivatives of C with respect to each numeric factor decrease monotonically toward zero as the level of C decreases. And the effect of S also appears to decrease as C decreases. Also, the first derivatives appear not to go exactly to zero, but are still positive. The same is true of the effect of S.

Suppose a power transformation of C is additive in the factors and is a linear function of the three numeric factors. That is,

$$C = C(S,L,T,D)$$

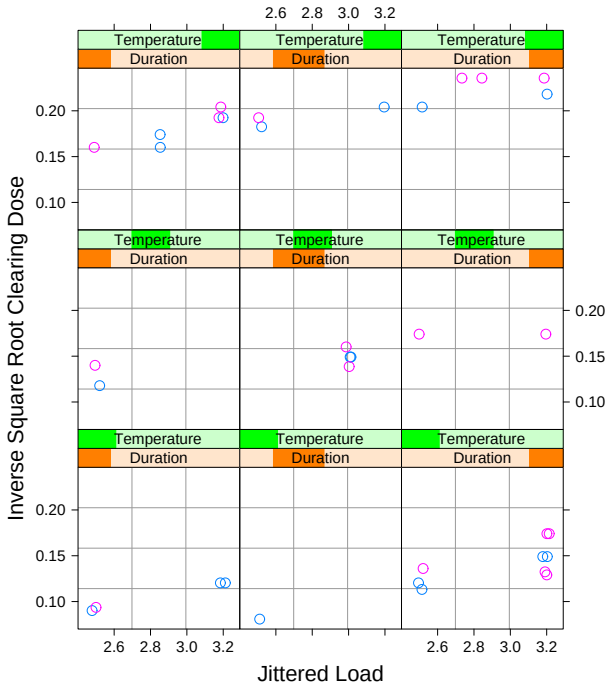

Figure 12: $1/\sqrt{C}$ against L given D, T, and S.

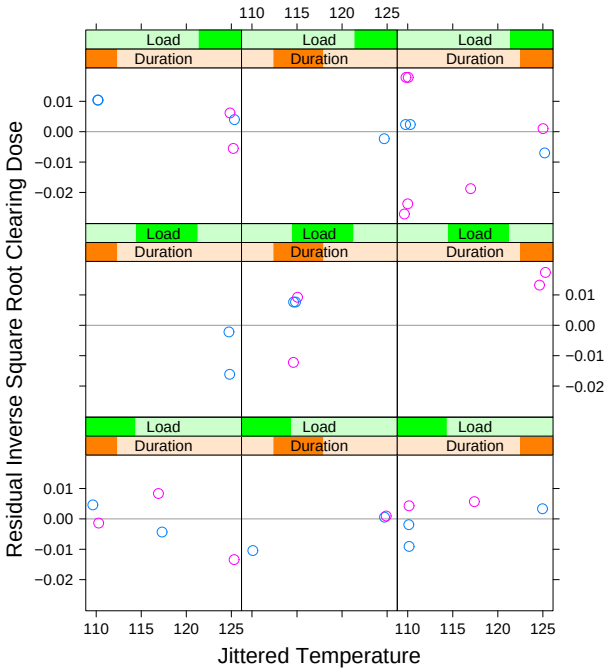

Figure 13: Residual $1/\sqrt{C}$ against T given D, L, and S.

$$= (\mu + \alpha S + \beta L + \gamma T + \delta D)^{-p}.$$

Suppose also that $p > 0$. Finally, suppose that the coefficients α, β, γ, and δ are all positive. Then the derivative of C with respect to any one of the numeric variables, say L, is

$$\frac{dC}{dL} = \frac{-\beta p}{(\mu + \alpha S + \beta L + \gamma T + \delta D)^{p+1}} = \frac{-\beta p}{C(S,L,T,D)^{p+1}}.$$

So the derivative is everywhere negative and its absolute value decreases with C, exactly the behavior of the resist data revealed by the trellis plots. Also, the second derivative of C with respect to any one of the numeric variables, say L, is

$$\frac{d^2C}{dL^2} = \frac{\beta^2 p(p+1)}{(\mu + \alpha S + \beta L + \gamma T + \delta D)^{p+2}} = \frac{\beta^2 p(p+1)}{C(S,L,T,D)^{p+2}}.$$

So the second derivative is everywhere positive and decreases with C, exactly the behavior of the resist data revealed by the trellis plots. Finally, it is easy to see that an analogous result holds for S.

The transformation analysis suggests that a power transformation of C might remove or reduce the non-linearity and the pervasive interactions among the factors. We repeated the trellis displays in Figures 10 to 12 for a number of power transformations. The displays for the inverse square root are shown in Figures 10 to 12. The units for the response are now cm/$\sqrt{\text{mJ}}$. The plots suggests that the dependence of $1/\sqrt{C}$ on the factors is additive, and linear in the numeric factors. Table 4 shows an analysis of variance for $100/\sqrt{C}$, carried out in the same manner as in Table 3. The new table also suggests an absence of non-linearity and interaction. Nature has been exceedingly good to us. A simple power transformation of C has resulted in a very simple model.

Table 4: Analysis of variance for $100/\sqrt{C}$.

Effect	DF	SS	MS	F	P
S	1	41.70	41.70	28.46	0.00
T	1	357.89	357.89	244.21	0.00
L	1	57.88	57.88	39.50	0.00
D	1	88.58	88.58	60.45	0.00
T^2	1	0.05	0.05	0.03	0.86
L^2	1	0.28	0.28	0.19	0.67
D^2	1	0.05	0.05	0.03	0.85
D × T	1	3.70	3.70	2.52	0.13
T × L	1	0.37	0.37	0.25	0.62
D × L	1	4.33	4.33	2.96	0.10
S × T	1	0.38	0.38	0.26	0.61
S × L	1	1.50	1.50	1.02	0.32
S × D	1	0.10	0.10	0.07	0.79
Error	22	32.24	1.47		

Trellis displays of the residuals on the transformed scale suggest our additive model has no appreciable lack of fit. One such residual display is shown in Figure 13, a trellis graph of the residuals against T given D, L, and S. A normal probability plot of the residuals shows that their distribution is well approximated by the normal. A spread-location, or s-l, plot [10] shows that the variance does not change with the fitted values. The estimate of the standard deviation using the residuals is 0.0121, a very small number since the range of $C^{-0.5}$ is close to 0.30; the model explains much of the variation in the data.

3.4. Trellis Display of the Fitted Response Surface: Higher Order Interaction Plots

Trellis display is also very effective for studying models fitted to data from designed experiments, providing a convenient and effective expansion of the interaction plot to higher order interactions. This is illustrated in the three trellis displays of Figure 14. The simple model for $C^{-0.5}$ just fitted has been transformed back to the original scale by the inverse square, allowing study on the scale of mJ/cm^2 to appeal to engineering intuition about clearing dose. There is one trellis display for each numeric factor.

For example, in the left display of Figure 14, C is graphed against T given D, L, and S. There are 16 panels on the display. There are 4 equally spaced values of D and 4 equally spaced values of L, each ranging from the minimum to the maximum value in the data. C is valued for all 16 combina-

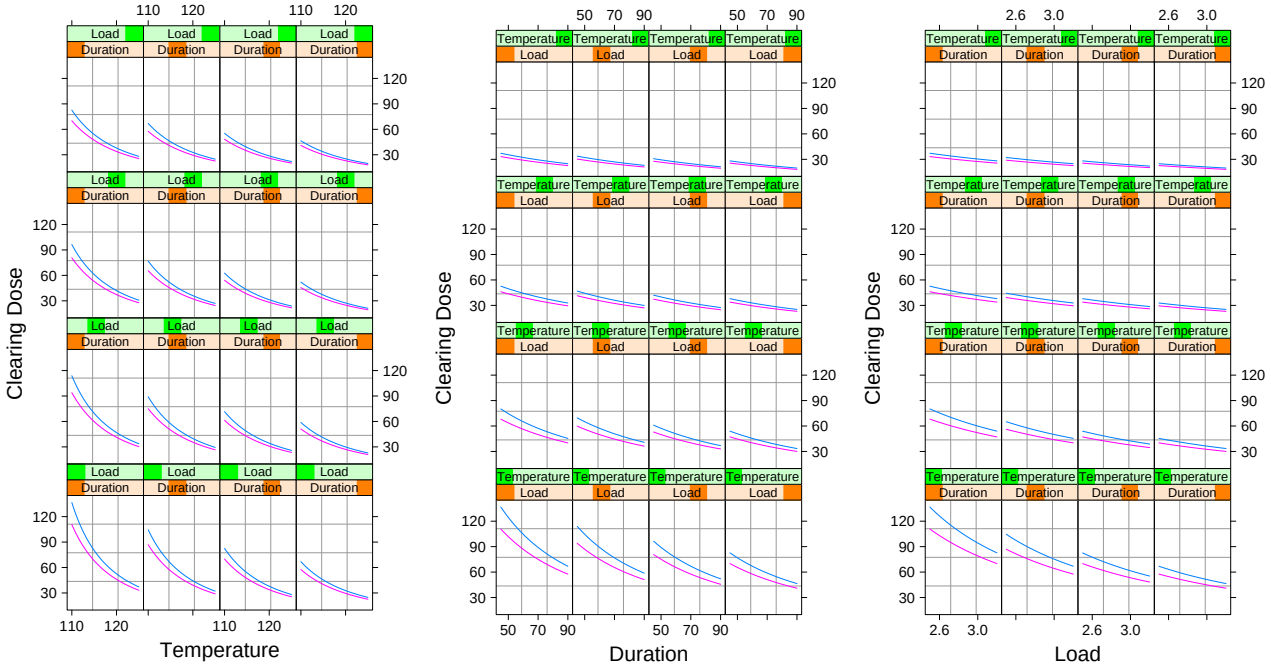

Figure 14: Left: Fitted C against T given D, L, and S; Middle: Fitted C against D given L, T, and S; Right: Fitted C against L given D, T, and S.

tions of these two factors at 50 values of T for each of the two values of S. This results in two curves on each panel, one for solvent 1 (magenta) and one for solvent 2 (cyan). The other two displays have similar evaluations.

Figure 14 shows the non-linearity and the strong interactions among all factors revealed in our initial trellis plots of the data, albeit far more incisively here. For example, in the left trellis display of Figure 14, we see clearly that as the conditioning value of D increases for fixed L, or as the conditioning value of L increases for fixed D, the magnitudes of the three quantities — C, dC/dT, d^2C/dT^2 — all decrease. We can also see that solvent 2 (cyan) results in a larger C than solvent 1 (magenta). From this display we are able to assess the complex properties, which possess complex interactions of the factors. This shows why modeling on the original scale of C, mJ/cm^2, is so challenging.

4. LIQUID CRYSTAL DATA

4.1. Polymer-Dispersed Liquid Crystal Displays

Reflective displays that are visible in ambient lighting and operate without back lights reduce weight and power requirements. Polymer dispersed liquid crystals (PDLCs) are promising materials for these reflective displays. Under normal conditions, the droplets of a liquid crystal are randomly oriented, and the material is white because light is scattered. But when a voltage is applied to a section of the liquid crystal, the droplets align, scattering is reduced, and the section becomes transparent. If the background behind the material is black, applying a voltage makes the section go from white to black.

4.2. The Experiment

The switching voltage is the voltage necessary to align the droplets. One series of experiments [33] studied the dependence of switching voltage, V, on three factors: (1) the amount, M, of liquid crystal in the mixture, measured in wt %; (2) the intensity, I, of the light used in the processing, measured in mW/cm^2; (3) the temperature, T, of the mixture during processing, measured in °C.

We will describe here the modeling of the data from the pilot experiment that began the series. In the pilot, each triple of values of the three factors was close to one of nine design locations — the corners and center of a cube whose edges are parallel to the factor axes. Eight of the design locations had two runs and one had three, so there were 19 runs in all.

4.3. Analysis of Variance

Table 5 shows an ANOVA for an overall model that is quadratic in the variables. If we drop the terms that are insignificant, the residual sum of squares, 6.22 V^2, remains nearly the same and $\hat{\sigma}$ = 0.74 V. But if in this reduced model we drop T^2 and replace it with M^2, the residual sum of squares also remains nearly the same. Thus this ANOVA does not

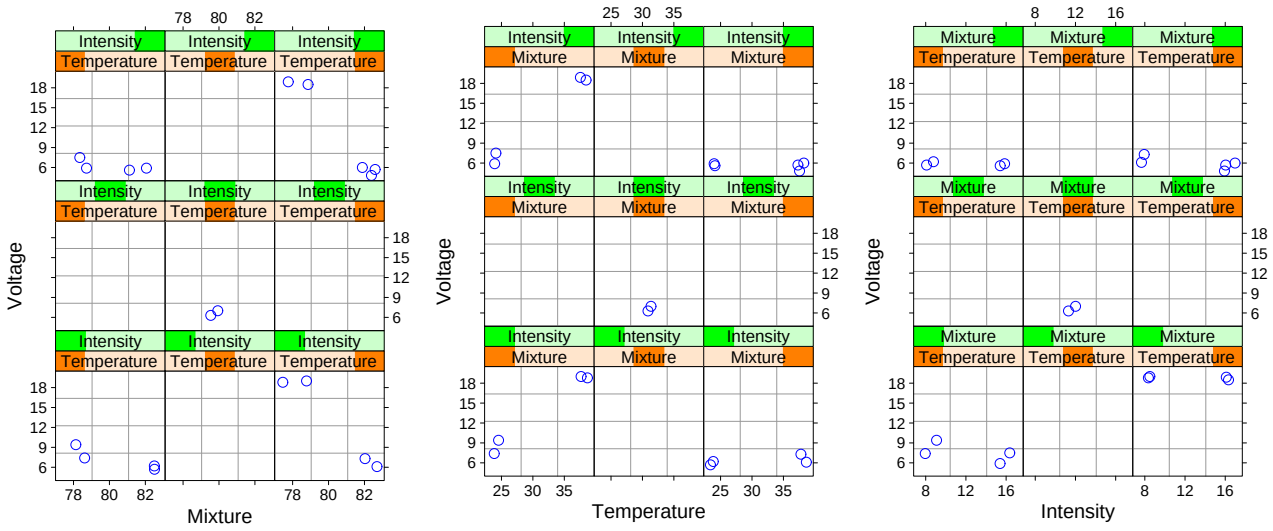

Figure 15: Left: V against M given T and I; Middle: V against T given M and I; Right: V against I given M and T.

Table 5: Analysis of variance for liquid crystal data.

Effect	DF	SS	MS	F	P
I	1	6.49	6.49	9.40	0.01
M	1	219.23	219.23	317.38	0.00
T	1	126.43	126.43	183.04	0.00
I^2	1	0.23	0.23	0.34	0.58
M^2	1	0.64	0.64	0.93	0.36
T^2	1	14.49	14.49	20.98	0.00
M × T	1	126.39	126.39	182.97	0.00
I × T	1	0.02	0.02	0.03	0.87
I × M	1	0.03	0.03	0.04	0.85
Error	9	6.22	0.69		

yield an unambiguous model specification. We should not take this to mean that there is an irresolvable ambiguity in the data because, our ANOVA rests on the unsubstantiated hypothesis that the overall quadratic model adequately describes the structure in the data.

4.4. Trellis Displays of the Raw Data

We will use trellis display to search for insight into the dependence of the response on the factors. Each factor in the experiment — I, M, and T — has low values, medium values, and high values. We will condition on each factor using three intervals that divide its values into low, medium, and high. The number of combinations of the three sets of three intervals is 27. However, the design only covers 9 of them, so we can expect to see gaps in the trellis displays.

The left trellis display of Figure 15 graphs V against M given T and I. The values of T go from low to medium to high as we go from left to right through the columns. The values of I go from low to medium to high as we go from bottom to top through the rows. For the highest level of T, there is a large decrease in V with M; for the lowest level, there is a small decrease. Furthermore, for the middle level of T, the values of V are close to what they are for the lowest level of T. But the changes in V with M do not appear to depend on the level of I. Thus there appears to be a strong interaction between T and M, but no interaction between I and M.

The middle trellis display of Figure 15 graphs V against T given M and I. There is more information about the T and M interaction. For M at the lowest level, V increases by a large amount with T. But for M at the highest level, there does not appear to be an effect due to T. And for M at the middle level, V has values close to what they are for M at the highest level. Finally, there appears to be no appreciable interaction between I and T.

The right trellis display of Figure 15 graphs V against I given M and T. As I increases, V decreases. The sizes of the decreases vary but there is no consistent pattern to their variation and the magnitude of the variation is not large compared with the variation of the replicated points, so there appears to be little or no interaction between I and the other two factors.

4.5. Modeling the Data

The ANOVA carried out earlier was predicated on a quadratic dependence of V on the factors: M, T, and I. But the structure of the data revealed by the trellis displays calls into question the appropriateness of a quadratic model. The reason is the radical change in slope. As a function of M and T, V is large for M at the lowest level and T at the highest. V is much smaller and nearly flat elsewhere.

Let us describe the structure we observed in the trellis displays in terms of Figure 16, a scatter-plot of the measurements of M and I with a small amount of uniform random noise

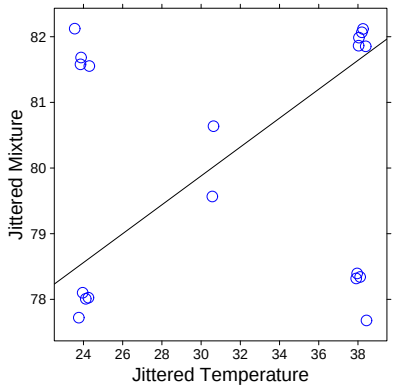

Figure 16: Measurements of M and T. The line is the estimated join line of the spline surface.

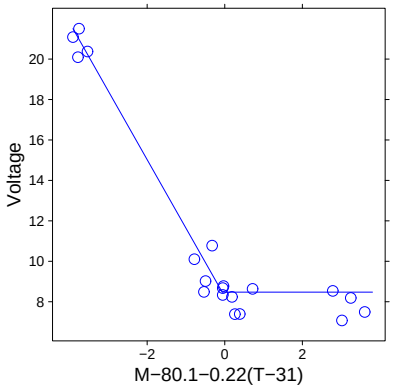

Figure 17: Partial residual plot graphs $V_i - \hat{\gamma} I_i$ against $M_i - \hat{\alpha} - \beta T_i$.

added to break up the overlap of plotting symbols. (The line on the plot will be explained later.) At the points in the lower right corner, V is high and much lower everywhere else. In going from the points in the upper left to the lower left, there is a small increase in V. In going from the upper left to the upper right, V is constant. At the two center points, V is between where it is for the upper points (left and right) and where it is in the lower left.

A simple model explains the structure revealed by the trellis displays: (1) linear in I; (2) a continuous piecewise linear spline in T and M consisting of two half planes that join along a line in the T and M space that, in Figure 16, is close to the center points and the points in the lower left and upper right; (3) the half plane covering the upper left in Figure 16 has zero slope. Thus the model is

$$V_i = \mu + \gamma I_i + \delta (M_i - \alpha - \beta T_i)^- + \varepsilon_i$$

where x^- is x if $x < 0$ and is 0 otherwise. The join line is

$$M - \alpha - \beta T = 0.$$

We will begin with an assumption that the ε_i are independent and normally distributed with mean 0 and constant variance σ^2. Thus the parameters will be estimated by nonlinear least-squares. The residual sum of squares is 4.8 V^2, and $\hat{\sigma} = 0.55$ V, better than the quadratic models. Residual plots suggest there is no significant lack of fit and that the above assumptions about the error terms ε_i are reasonable.

The line drawn in Figure 16 is the estimated join line,

$$M - \hat{\alpha} - \hat{\beta} T = 0.$$

Figure 17 is a partial residual plot that shows the spline fit. $V_i - \hat{\gamma} I_i$ is graphed against $M_i - \hat{\alpha} - \beta T_i$. The fitted function tracks the data. The spline fit explains the subtle behavior in the trellis displays of the data in Figure 15.

5. DISCUSSION

We investigated the use of the conditioning methods of trellis display for experimental data with a small number of runs including highly fractionated designs. Such experiments arise in settings where each run is very expensive, or as in some computer experiments, each run takes a long time. The investigations, which covered a large number of data sets over time, have led to several conclusions.

Trellis display is almost always quite useful for modeling data from these experiments, and commonly produces major changes in the analysis, modeling, and results due to the discovery of patterns in the data not suspected before the data were collected. The finding is true even for highly fractionated experiments. The pattern discovery is a result of the scope of the trellis conditioning methods. They are applicable not just to fitted models, but also to the raw data. This allows incisive assessments of patterns in the data that can lead to substantial improvements to models initially entertained, or to a conclusion that an experiment failed. Both were demonstrated in the examples described here.

We also found that ANOVA, used pervasively in the analysis of experimental data, is a powerful tool for answering specific questions about models for data, but a poor tool for guiding the overall modeling process.

One might find it remarkable that conditioning methods can often succeed for highly fractionated experiments. Conditioning for such designs often results in just a few points on the panel of a trellis display, potentially making it hard to assess dependence because of variability in the error term. However, we found patterns often did emerge as demonstrated in our examples. The reason appears to be that for success, highly fractionated designs require an engineering practice that keeps error variability small. Such designs, by their very nature, cannot succeed in cases with large error variability that require a large aggregation of runs to see a signal. The very practices that make such experiments succeed allow trellis methods to

succeed.

Acknowledgments: This work was supported in part by National Science Foundation Award CCF-0937123, Scalable Visualization and Model Building; and in part by National Science Foundation Award DMS-0532217, Data Mining, Statistical Learning, and Data Visualization for Complex Data.

References

[1] R. A. Becker, W. S. Cleveland, and M. J. Shyu. The Design and Control of Trellis Display. *Journal of Computational and Statistical Graphics*, 5:123–155, 1996.

[2] P. Murrell. *R Graphics*. Chapman & Hall, New York, 2006.

[3] J. C Pinheiro and D. M. Bates. *Mixed Effects Models in S and S-Plus*. Springer-Verlag, New York, NY, 2000.

[4] D. Sarkar. *Lattice: Multivariate Data Visualization with R*. Springer, New York, NY, 2008.

[5] A. Krause and M. Olson. *The Basics of S and S-Plus*. Springer, New York, 2000.

[6] J. M. Chambers. *Software for Data Analysis: Programming with R*. Springer, New York, 2008.

[7] S. McDaniel. *Rapid Graphs with Tableau Software: Create Intuitive, Actionable Insights in Just 15 Days*. CreateSpace, United States, 2009.

[8] G. E. P. Box, J. S. Hunter, and W. G. Hunter. *Statistics for Experimenters: Design, Innovation, and Discovery*. Wiley, Chichester, U. K., 2005.

[9] R. A. Becker and W. S. Cleveland. Brushing scatterplots. *Technometrics*, 29:127–142, 1987. Reprinted in *Dynamic Graphics for Data Analysis*, edited by W. S. Cleveland and M. E. McGill, Chapman and Hall, New York, 1988.

[10] W. S. Cleveland. *Visualizing Data*. Hobart Press, Chicago, 1993.

[11] O. L. Davies. *The Design and Analyisis of Industrial Experiments*. Hafner, New York, 2nd edition, 1967.

[12] S. Feiner and C. Beshers. Worlds within Worlds: Metaphors for Exploring n-Dimensional Worlds. *Proceedings of UIST '90 (ACM Symp. on User Interface Software)*, pages 76–83, 1990.

[13] A. E. Freeny and J. M. Landwehr. Graphical Analysis for a Large Designed Experiment. *Technometrics*, 37: 1–14, 1995.

[14] R. B. Gramacy and H. K. H. Lee. Adaptive Design and Analysis of Supercomputer Experiments. *Technometrics*, 51:130–145, 2009.

[15] W. A. Larsen and S. J. McCleary. The Use of Partial Residual Plots in Regression Analysis. *Technometrics*, 14:781–790, 1972.

[16] T. Mihalisin, J. Timlin, and J. Schwegler. Visualizing Multivariate Functions, Data, and Distributions. *Computer Graphics and Its Applications*, 11: 28–35, 1991.

[17] R. D. Snee. Experimenting with a Large Number of Variables. In R. D. Snee, editor, *Experiments in Industry*, pages 25–35. American Society for Quality Control, Milwaukee, Wisconsin, U.S.A., 1985.

[18] E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut, U.S.A., 1983.

[19] P. A. Tukey and J. W. Tukey. Graphical Display of Data Sets in 3 or More Dimensions. In V. Barnett, editor, *Interpreting Multivariate Data*, pages 189–275. Wiley, Chichester, U. K., 1981.

[20] R. R. Barton. *Graphical Methods for the Design of Experiments*. Springer, New York, NY, 1999.

[21] J. A. Cornell and L. Ott. The Use of Gradients to Aid in the Interpretation of Mixture Response Surfaces. *Technometrics*, 17:409–424, 1975.

[22] C. Daniel. *Applications of Statistics to Industrial Experimentation*. Wiley, New York, 1976.

[23] R. D. Snee. Graphical Display of Results of Three-Treatment Randomized Block Experiments. *Journal of the Royal Statistical Society. Series C*, 34:71–77, 1985.

[24] W. J. Youden. Graphical Diagnosis of Interlaboratory Test Results. *Industrial Quality Control*, 15:24–28, 1959.

[25] K. Hinkelmann and O. Kempthorne. *Design and Analysis of Experiments: Introduction to Experimental Design*. Wiley, Chichester, U. K., 1994.

[26] Y. Hung, V. R. Joseph, and S. N. Melkote. Design and Analysis of Computer Experiments With Branching and Nested Factors. *Technometrics*, 51:366–376, 2009.

[27] A. C. Shoemaker, K.-L. Tsui, and C. F. J. Wu. Economical Experimentation Methods for Robust Design. *Technometrics*, 33:415–427, 1991.

[28] A. E. Vine, S. M. Lewis, A. M. Dean, and D. Brunson. A Critical Assessment of Two-Stage Group Screening Through Industrial Experimentation. *Technometrics*, 50: 15–25, 2008.

[29] NIST and SEMATECH. Engineering Statistics Handbook, www.itl.nist.gov/div898/handbook/eda/section1/eda15.htm, Verified 2010.

[30] Jr. W. F. Hunt. Experimental Design in Air Quality Management. In R. D. Snee, L. B. Hare, and J. R. Trout, editors, *Experiments in Industry*, pages 89–98. American Society for Quality Control, Milwaukee, 1985.

[31] P. J. Rousseeuw and K. van Driessen. A Fast Algorithm for The Minimum Covariance Determinant Estimator. *Technometrics*, 41:212–223, 1999.

[32] O. Nalamasu, A. Freeny, E. Reichmanis, N. J. A. Sloane, and L. F. Thompson. Optimization of Resist Formulation and Processing with Disulfone Photo Acid Generators Using Design of Experiments. Technical report, Bell Laboratories, Murray Hill, New Jersey, U.S.A., 1993.

[33] J. D. LeGrange, S. A. Carter, M. Fuentes, J. Boo, A. E. Freeny, W. Cleveland, and T. M. Miller. The Dependence of the Electro-Optical Properties of Polymer Dispersed Liquid Crystals on the Photopolymerication Process. *Journal of Applied Physics*, 81:5984–5991, 1997.