Trellis Display: Modeling Data from Designed Experiments

William S. Cleveland
Statistics Research Department, Bell Labs
Montserrat Fuentes
Department of Statistics, University of Chicago

Abstract

Displaying data by conditioning has surfaced independently in a number of places and for many different types of data. One surfacing is the interaction plots used in the analysis of experimental data. *Trellis display*, an amalgam of ideas for displaying multivariable data, allows conditioning to be carried out in a very general way. Trellis was developed initially in the context of large data sets, but recent experience has demonstrated that it is also useful for modeling data from designed experiments, even small experiments with a moderate number of factors and a limited number of runs. There appear to be two reasons for the success: (1) Trellis is an exceedingly powerful tool for revealing the structure of interactions, even high-order interactions; (2) experiments with limited runs often are accompanied by low noise, because otherwise little would be learned by any technique, so the conditioning of Trellis can yield information even when there are only a very small number of points on each panel. The analysis of variance, pervasively used 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. The addition of Trellis display substantially increases the ability of the data analyst to carry out model identification and diagnostic checking, often resulting in more parsimonious and better fitting models.

1 Introduction

Trellis display is a framework for the display of multivariable data (Becker and Cleveland, 1996a; Becker and Cleveland, 1996b). It was originally developed to assist *model building* in the context of moderate, large, and very large data sets. In this context it has proven to be a powerful system for understanding the structure of multivariable data sets, enabling the analyst to build models that are both parsimonious and that have supportable assumptions. Trellis is useful in the initial stage, when models are first formulated, and during the diagnostic stage, when models are checked to see if they fit the data.

This paper reports the results of an investigation of the use of Trellis display in analyzing data from designed experiments. This area of application of Trellis differs from previous areas in that data sets are often small. One central aspect of Trellis is conditioning — variables are displayed conditional on the values of other variables, possibly many other variables. The question was whether the multiple conditioning was practical and would lead to insight for experiments with several factors but a limited number of runs. 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 consisted of using Trellis in the analysis of data from many experiments, some reported in the literature, and some arising in our own work.

In this paper we report on the investigation by discussing the analyses of three of the data sets we studied. An important aspect of the discussion is verisimilitude; the characteristics of the data sets shown here by the Trellis displays were discovered by these displays. Thus the reader has the opportunity to judge the efficacy of Trellis. We also discuss the analysis of variance, or anova, a tool that is frequently used for building models for experimental data.

Section 2 introduces Trellis using data from an experiment on lead concentration near a major roadway. Section 3 describes the salient aspects of Trellis display. Section 4 returns to the lead concentration data; the published account of the original analysis, which relied heavily on anova to build a model for the data, is discussed in light of what is learned about the data from the Trellis displays of Section 2. Section 5 presents an analysis of data from an experiment on etching the resist layer of computer chip wafers. Section 6 presents an analysis of data from the first of a series of experiments on liquid crystal displays. Section 7 presents and discusses our conclusions:

1. Trellis display typically provides substantial insight into data from designed experiments, even small experiments with a limited number of runs and several factors.

2. The analysis of variance, pervasively used 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.

In the section, we also present some explanations for these conclusions.

2 Trellis Display of Lead Concentration Data

W. F. Hunt (1985) analyzes data from an experiment to determine the spatial variation in lead concentrations at a site next to a major roadway in Ohio 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; emissions, which vary daily and likely have a day-of-the-week effect; and spatial position.

We will use the lead data for three purposes: (1) to convey the salient features of Trellis display; (2) to provide a vehicle for introducing issues about the analysis of variance and its role in analyzing experimental data; (3) to provide an example where an undue reliance on anova can result in an inappropriate model for the data. We will not go beyond these three purposes. In particular, we will not develop a model here that fits the data, in part, in the interest of space, and, in part, because such development is carried out for the two examples in Sections 5 and 6.

We will think of the lead data as being made up 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. We have $3 \times 3 \times 21 - 1 = 188$ 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 $3 \times 21 = 63$ panels arranged into 21 columns and 3 rows. Each panel has a scatterplot 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. The setback changes with the row; for row 1, the bottom row, the setback distance 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 lead concentration tends to decrease as height increases. The decline as a function of height lessens as the setback increases. Figure 2 shows mixed behavior in the dependence of lead on setback. For the lowest height, lead decreases with setback. But for the middle value of height, lead typically first increases with setback and then decreases. For the highest height, 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 automobile tail pipes. The closest of the 9 monitors, the one with the lowest height and the closest setback, 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 at the closest setback.

The arrangement of the panels in Figure 1 allows us to study three collections of patterns, one collection for each row; that is, we can study the 21 patterns of dependence of lead on height for one setback, and then compare the collection with the collection for any other setback. Suppose, however, that we want to study the three patterns for each day, and then compare the 21 collections of daily patterns. This is a more difficult task in Figure 1 because the three panels for each day are arranged vertically in such a way that we have a reduced ability to effortlessly visually assemble the three patterns. In Figure 3, the panels have been rearranged to facilitate the study of the daily patterns. Now the three 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 on most days, the amount of change in lead with height for the closest setback is bigger than that for the other setbacks, and on a few days is dramatically bigger. The most dramatic is Tuesday of week 1; the change in L at the closest position is about 1.25 ug/cm, but is less than 0.02 ug/cm at the other positions.

Figure 3 also shows a substantial amount of correlation. The measurements at each of the 9 positions have time correlation induced by the time correlation in the meteorological factors that have a large influence on the lead concentrations. The autocorrelations for short time spans of one to a few days are positive and large. Furthermore, the 9 values at each time point are highly positively correlated; the 9 values move up and down through time in concert. The meteorological effect causes a far greater variation in the data than the day-of-the-week effect, which appears to have relatively little influence.

Figure 3 also suggests spatial correlation. The changes in concentration with height for the three setbacks also appear highly correlated. To study this phenomenon further we will fit a model to the data with a day effect and with an effect for each combination of position and height, and then graph residuals. That is, if lead concentration on day

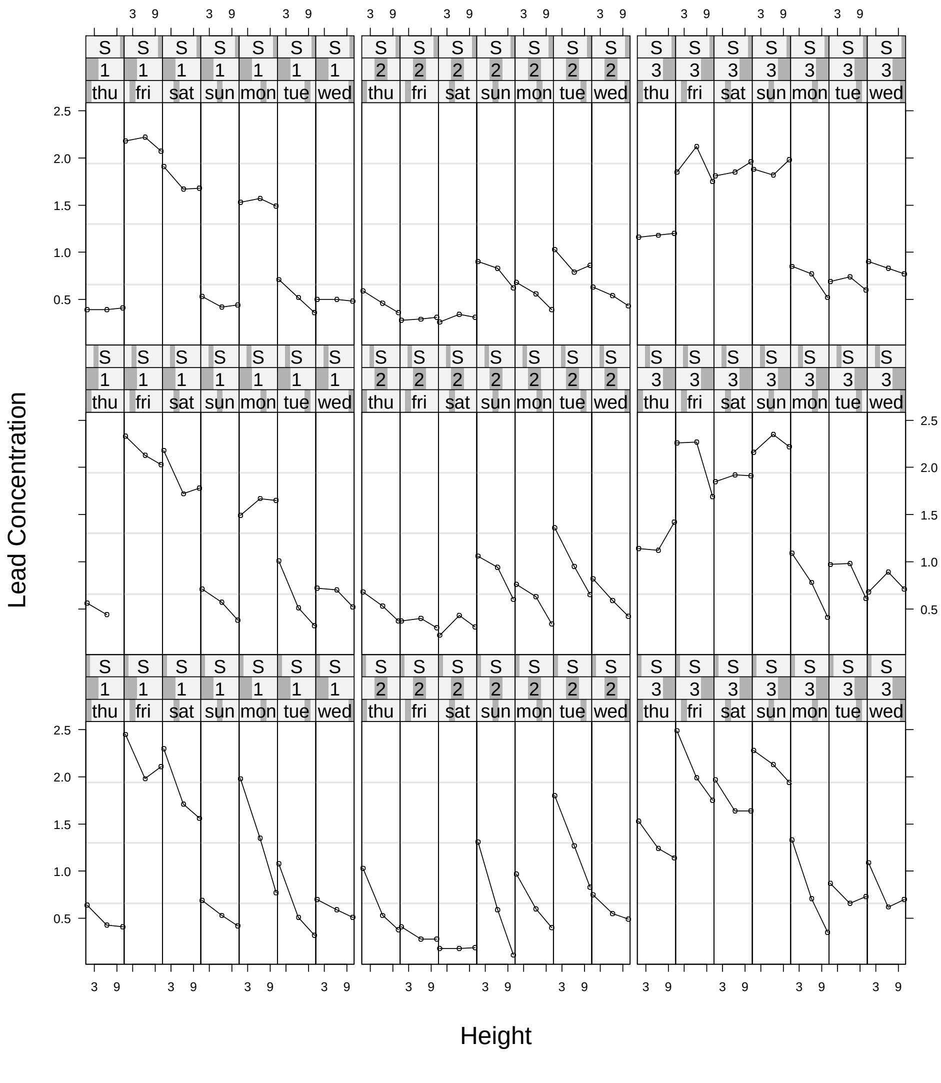

Figure 1: Trellis display of L against H given D, W, and S.

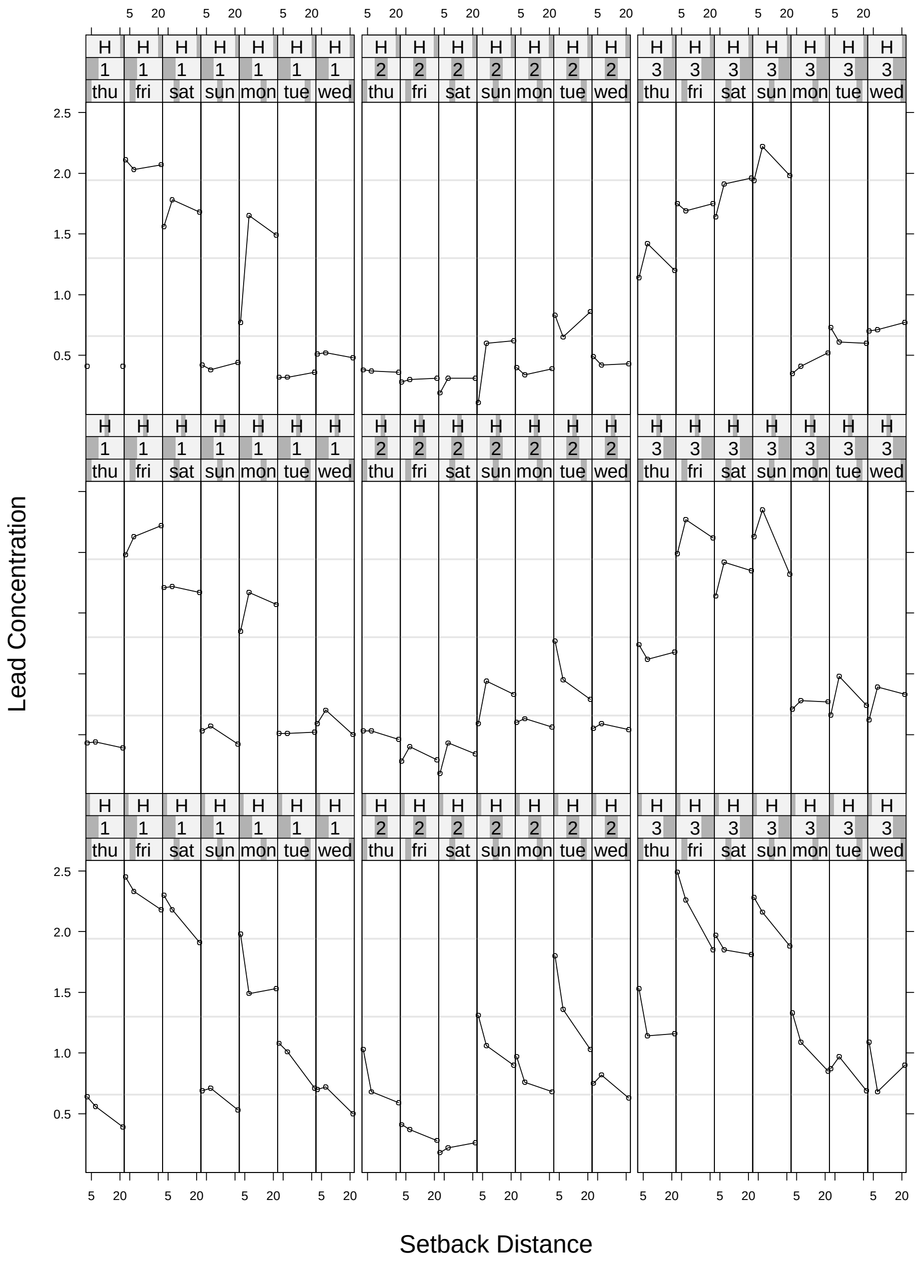

Figure 2: Trellis display of L against S given D, W, and S.

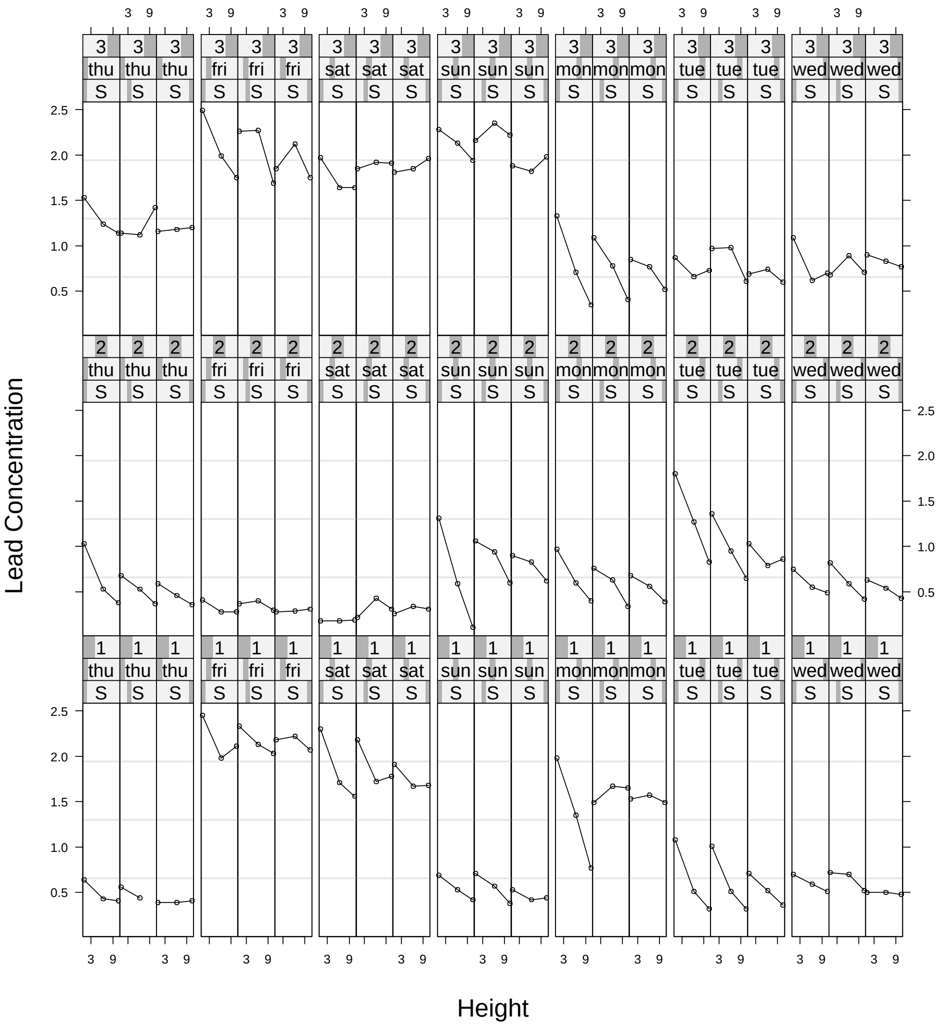

Figure 3: Trellis display of L against H given S, D, and W.

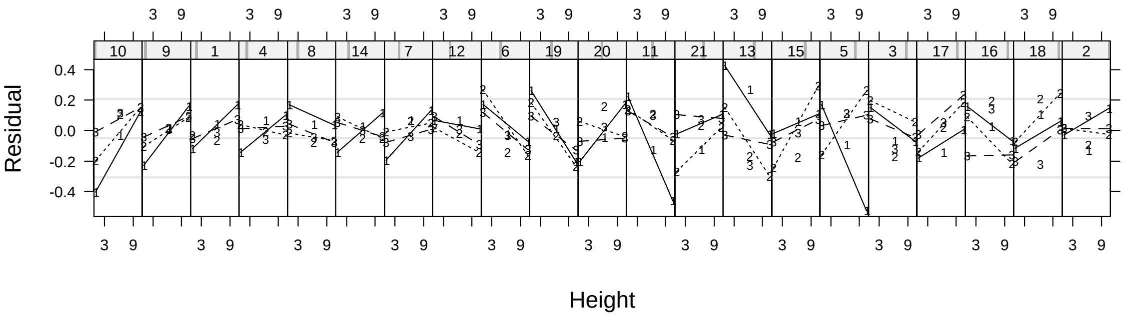

Figure 4: Trellis display of residual L against H given S and day (ordered by mean level).

of the week d in week w for setback s, and height h is ℓ_{dwsh}, then the model is

$$\ell_{dwsh} = \mu + \alpha_{dw} + \beta_{sh} + error$$

where

$$\sum_{d=1}^{7}\sum_{w=1}^{3}\alpha_{dw} = \sum_{s=1}^{3}\sum_{h=1}^{3}\beta_{sh} = 0.$$

In looking at residuals we are adjusting the data for the overall level on each day and the overall level at each of the 9 monitors.

Figure 4 graphs the residuals against height. Each panel has the residuals for one day. The numbers used as plotting symbols show the setback; "1" is the closest and "3" is the furthest. For each setback, the points for the lowest and highest heights are connected by a line segment. The panels in this case are not ordered by time, but rather by the daily means of the 9 monitors. As we go left to right through the panels, the means increase; the day number is shown in the strip label. This ordering allows us to determine whether patterns in the residuals depend on the overall level of response, a common occurrence. Figure 4 shows a strong correlation in the residuals; for example, the slopes of the line segments on each panel are clearly correlated. This means, as we suspected, that the changes with height are correlated, a form of spatial correlation. This adds to the complexity of modeling the data. But the slopes do not appear to increase or decrease monotonically with the daily means, an unfortunate occurrence since with monotone dependence there is a chance of removing such a slope effect by transformation of the data.

3 Trellis Display

Displaying data by conditioning has surfaced in a number of places and for many different types of data (Davies, 1967; Tukey and Tukey, 1981; Tufte, 1983; Snee, 1985; Feiner and Beshers, 1990; Mihalisin, Timlin and Schwegler, 1991; Cleveland, 1993). Trellis display (Becker and Cleveland, 1996a; Becker and Cleveland, 1996b) has features that allow conditioning to be carried out in a general way; we can condition on many variables, and each conditioning variable can be of any type (categorical, ordered categorical, discrete numeric, or continuous numeric). Furthermore, the graphical method used to display the data on each panel is independent of the conditioning, so that any graphical method can be used. The following are some of the ideas that make up the Trellis framework.

Layout

A Trellis display consists of panels laid out into a three-way rectangular array of columns, rows, and pages. In Figures 1 to 3 the layout is 21 columns, 3 rows, and 1 page. In Figure 4, the layout is 21 columns, 1 row, and 1 page.

Traditionally, data visualization methods have been developed with a notion of the display taking a single page (or window on a screen). But for very large databases, the amount of space on a single page is insufficient. The solution in Trellis display is a structured move to multiple pages. In today's computing environments, document viewers allow scanning multiple pages, and we can expect future environments to have vastly improved tools. However, for the small experimental data sets in this paper, we will not need to resort to more than one page.

Panel Variables, Packets, and Conditioning Variables

On each panel of a Trellis display, a subset of the values of *panel variables* are displayed. Each subset, or *packet*, is determined by the values of the *conditioning variables* for the panel. In Figure 3, the panel variables are lead concentration and height, and the conditioning variables are day-of-the-week, week, and setback. Each packet consists of 3 values of lead concentration and 3 values of height for one setback, one day-of-the-week, and one week. In Figure 4, the panel variables are residual lead (encoded on the vertical scale), setback distance (encoded by the numerical symbols), and height (encoded on the horizontal axis); the conditioning variable is the day.

Conditioning Order and Packet Order

Each conditioning variable, whatever the type of the variable, has *levels* that are used to do the conditioning. In Figure 3, the 7 levels for the day-of-the-week variable are the days of the week, the 3 levels for the week variable are the week numbers 1 to 3, and the 3 levels for the setback variable are the 3 unique values of the setbacks. For a continuous numeric variable the levels are intervals; one important method for defining levels for a continuous variable is the equal-count algorithm (Cleveland, 1993).

The data analyst specifies an order for the conditioning variables and an order for the levels of each conditioning variable. In almost all applications, it makes sense to use the natural order of the levels of numeric and ordered categorical variables. For unordered categorical variables, using *main effects ordering* of the levels often greatly increases our ability to compare patterns on different panels.

The packets are given an order by the following rule: the first packet corresponds to the first levels of all conditioning variables; then we proceed through all combinations of the levels of the conditioning variables by letting the levels of the first variable change most quickly, the levels of the second variable changing next most quickly, and so forth. In Figure 1, the order of the variables is day-of-the-week,

week, and setback. In Figure 3, the order of the variables is setback, day-of-the-week, and week.

Panel Order

The Trellis convention for the ordering of the panels is the following: The panel in the lower left is first; we then move from left to right through the first row, then left to right though the second row, and so forth.

Trellising

Trellising is the arrangement of the panels on a Trellis display and the assignment of conditioning variables and panel variables to the panels. Trellising is carried out by the analyst specifying a layout, an order for the conditioning variables, and an order for the levels of each conditioning variable. The packets then go to the panels by matching the packet order and the panel order.

In Figure 3, the conditioning variables and their levels are the same as in Figure 1, and the orderings of the levels are the same, but the order of the conditioning variables is different; in Figure 1, the order of the conditioning variables is day-of-the-week, week, and setback, and in Figure 3, it is setback, day-of-the-week, and week. Figure 1 allows a better appreciation of changes in the pattern of the dependence of concentration on height through time. Figure 3 allows a better appreciation of changes in the pattern of the dependence of concentration on height with setback distance.

Strip Labels

Sometimes, seemingly small matters are in fact the salient ones that enable a system to work. This is the case for the strip labels at the tops of panels in Trellis displays, which convey the levels of the conditioning variables. Early in the development of Trellis display, levels were labelled in the margins of the displays as they are in other attempts at using conditioning in data visualization. Margin labels are exceedingly difficult to automate in software systems, inevitably require tuning by hand by the user, and do not work in some important types of conditioning. Now, in Trellis, the strip labels for a panel form a part of the panel, and so, are carried along wherever the panel goes; thus, trellising is not constrained by the labeling of conditioning variables and their levels. This enables the generality and simplicity of the Trellis conditioning mechanism.

Spacing

In Trellis display, space can be inserted between any two adjacent columns or any two adjacent rows. In Figure 3, the setback variable cycles through its values seven times on each row. To help us see each new cycle, space has been inserted between the panels where one cycle ends and the next begins.

Banking to 45°

On each panel of the Trellis displays shown so far we have judged the pattern of dependence of lead on either height or setback distance. In particular, by judging the orientations of the line segments connecting the successive observations, we extract information about the rate of change of concentration with the variable on the horizontal axis. The aspect ratio, the height of a panel divided by its width, controls the orientations; as the aspect ratio increases, all nonzero orientations tend to vertical, and as the aspect ratio decreases, all nonzero orientations tend to horizontal. Studies in graphical perception have shown that such judgments are most accurate overall when the absolute values of the orientations of the segments are centered on 45° (Cleveland, 1994). This banking to 45° has been used on all of the displays shown so far. For example, in Figure 3, a weighted mean of the absolute values of the orientations of the line segments is 45°; the weights are the lengths of the line segments.

4 Lead Concentration Data

In the source publication for the lead concentration data (W. F. Hunt, 1985), 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 4 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. The significant effects are S, H, SH, W, D, and DW, so these effects provide a modeling of the data. Note that this resulting model is exactly the model used in Section 2 to fit the data for the purpose of plotting residuals in Figure 4.

It is reasonable to begin the analysis of the lead data by hypothesizing the overall model used in the above anova in the hopes that it would provide an adequate description. The time effects are induced by weather conditions which persist for several days at a time. Thus we would expect smoothness in the time effects, that is, serial correlation in the form of persistence. The individual parameters of the above model could account for persistence but would do so without enforcing smoothness. In other words, the

Effect	DF	SS	MS	F	P
S	2	0.25	0.125	5.50	0.00
H	2	2.74	1.369	60.18	0.00
W	2	18.97	9.485	416.86	0.00
D	6	16.02	2.670	117.36	0.00
S × H	4	1.07	0.268	11.78	0.00
W × D	12	36.15	3.012	132.39	0.00
S × W	4	0.02	0.006	0.25	0.91
H × W	4	0.10	0.025	1.08	0.37
Error	152	3.46	0.023		

Table 1: Analysis of variance for lead concentration data.

model is liberal in parameters. A time series model could yield the smoothness. Still, the liberal-parameter model, if it fits the data, is quite simple, and for the purposes at hand, might do adequately.

The problem is that the liberal-parameter model does not fit the data. The reason is that it ignores the spatial correlation revealed by the Trellis displays of Section 2. Figure 4, the plot of residuals, shows that the 6 decreases in concentration with height on a given day are positively correlated. Thus the correlation in the data is much more profound than that described by the liberal-parameter model. By failing to account for the spatial correlation in the data, the sums of squares that are the basis of the above anova table do not have the stated distributions.

We will be satisfied with having made an important point about anova and not continue attempts to model the lead concentration data. The point is that the validity of an anova depends on the validity of the overall model fitted to the data, in the case of the lead data, the liberal-parameter model. If the overall model is not valid, then screening the effects of the model using anova as a way to simplify the model, serves no purpose.

5 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 (which can be smaller than 0.2 μm), and then placing the wafer in a developer solution to remove the exposed areas of the resist.

Nalamasu, Freeny, Reichmanis, Sloane and Thompson (1993) ran a series of experiments to improve the resolution of etched features. Their processing of the wafers involved the following steps:

1. Coating a wafer with a resist solution containing a new photoacid generator, whose amount, or *load*, was varied in the experiment.
2. Using one of two *solvents* in the resist solution.
3. Prebaking the coated wafer at a *temperature* that was varied and for a *duration* that was varied.
4. Exposing the coated wafer to 248 nm light shone through a photo mask.
5. Baking the wafer at a *temperature* that was varied and for a *duration* that was varied.
6. Developing the wafer for 60 seconds in a developer solution.

One response in the experiment, the one we will study here, is the clearing dose, which is the exposure dose in mJ/cm^2 needed to clear a large area of the resist, a cross-shaped region 100 μm by 150 μm. The following are the factors in the experiment:

1. Load of the photoacid generator in the resist (% wt).
2. Solvent (solvents 1 and 2).
3. Temperature of prebake cycle (° C).
4. Duration of prebake cycle (sec).
5. Temperature of bake cycle (° C).
6. Duration of bake cycle (sec).

An I-optimal design was used that had 36 runs in four blocks of nine runs each where the surface was stipulated to be a full quadratic in the numeric variables and to have a solvent main effect and interactions with all numeric variables.

The prebake variables are easily seen to have little or no effect on the clearing dose and will not be considered further in the discussion. Thus the five variables that we will study are clearing dose, C ; solvent, S; temperature, T; duration, D; and load, L.

5.1 Analysis of Variance

Table 5.1 shows an anova. The overall model is quadratic in T, L, and D and has an interaction between S and each of the three linear terms of T, L, and D. Many effects are significant, including a quadratic in T, cross products of all numeric variables, and an $S \times L$ interaction. The abundance of effects means the response surface is complicated, a bit disconcerting since the number of runs is quite small.

To reliably estimate the effects we need more insight into the data than that given by the anova. We need some

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		

Table 2: Analysis of variance for resist data.

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.

5.2 Trellis Display

We will use Trellis display to search for insight into the dependence of the response on the factors. Figure 5 shows intervals that will be used for conditioning on the three numeric variables D, L, and T. The choice of such conditioning intervals is in general an important and complex issue, and the Trellis methodology contains methods to assist in making the choices, but for the resist data, each of the three numeric variables break up quite naturally into three groups.

Figure 6 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 type. 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 labels contain graphical portrayals of the intervals. The strips for each conditioning have a scale but there are no tick marks to indicate the numeric values; the scale value at the left endpoints of the strips is the minimum value of the measurements of the variable, the scale value at the right endpoints is the maximum, and the darkened bars show the intervals on this scale.

Figure 6 reveals something quite important. 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 magnitude of the overall slope of the pattern tends to zero; that is, the magnitude of the first derivative

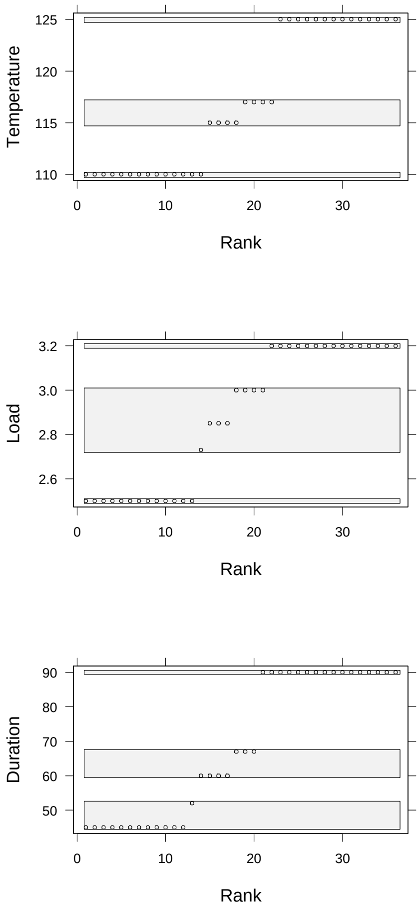

Figure 5: Conditioning intervals for D, L, and T.

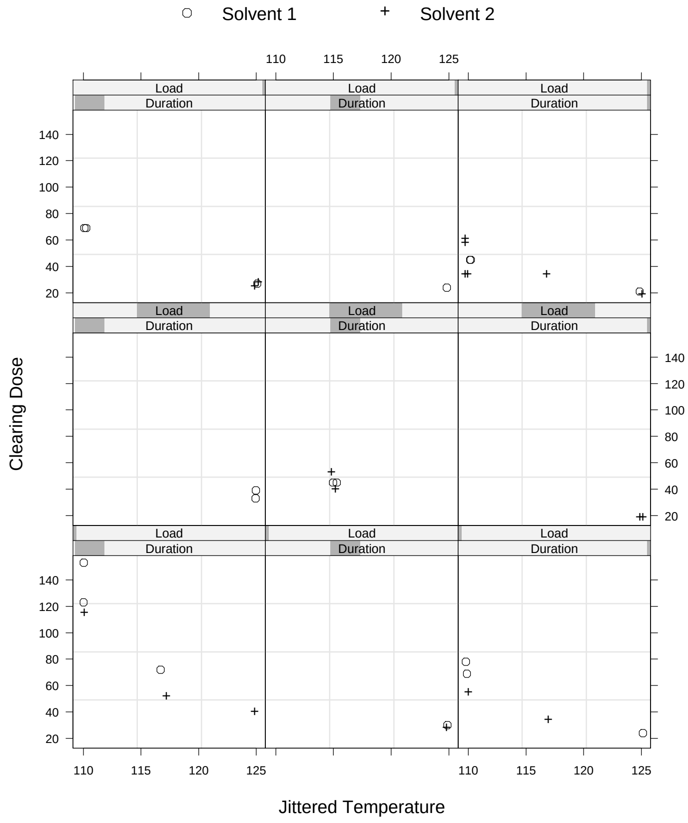

Figure 6: Trellis display of C against T given D, L, and S.

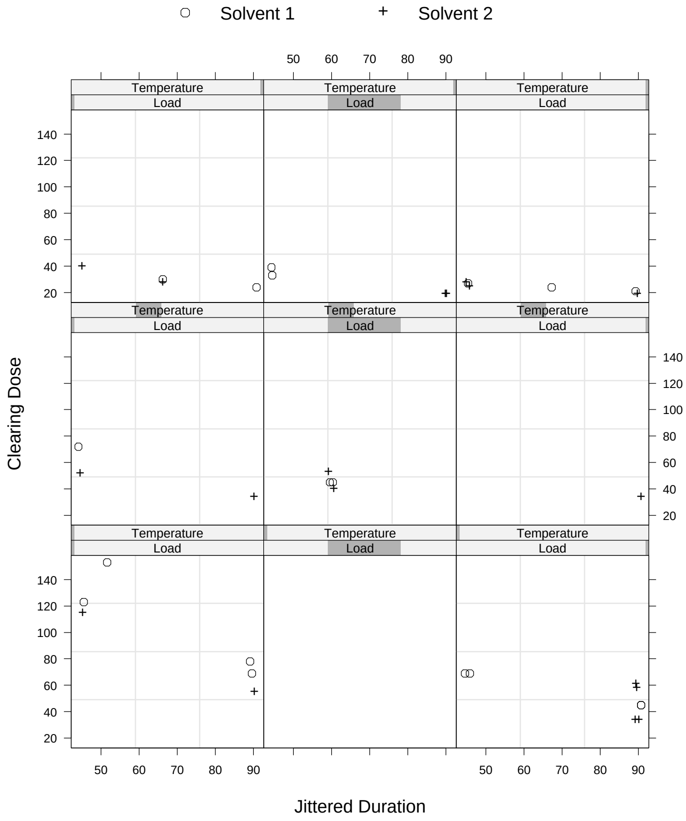

Figure 7: Trellis display of C against D given L, T, and S.

decreases with increasing C. The change in the derivative is quite large indicating a strong interaction. Furthermore, it appears that the curvature of the dependence pattern also decreases as C increases; that is, the magnitude of the second derivative decreases with increasing C. Finally, there is a suggestion that overall, solvent 1 leads to higher values of C than solvent 2, 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 7 and 8 are further Trellis displays of the data. The first is a graph of C against D given L, T and S; the second is a graph 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 derivatives on the level of C and the same dependence of the effect of S on the level of C.

Thus, we have found an important property of the interactions, a regularity that may well lead to a parsimonious model for the data. 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. Another matter that will prove important later is that the first derivatives appear not to go to zero nor the effect of S.

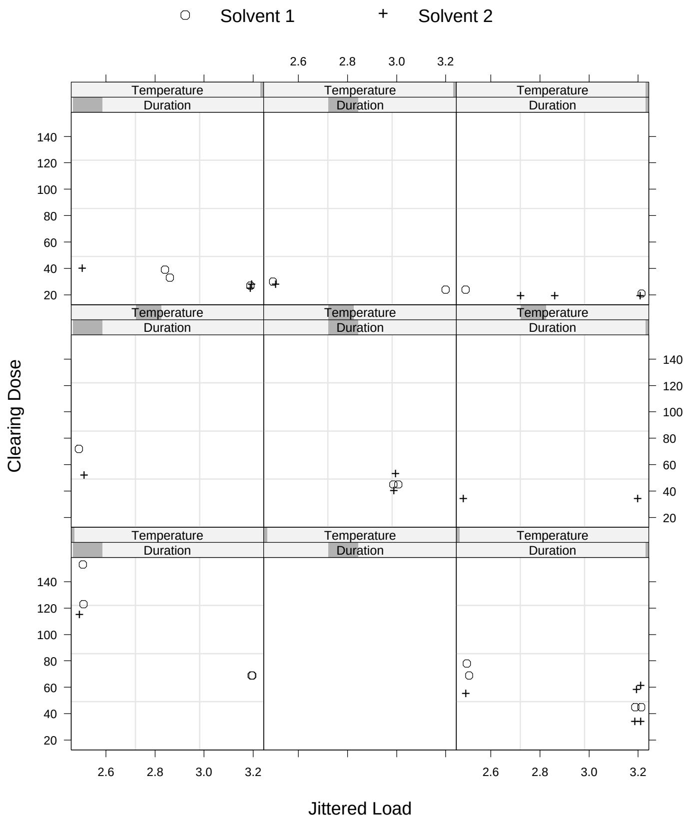

Figure 8: Trellis display of C against L given D, T, and S.

5.3 Exploiting the Regularity to Model the Data

The regularity revealed by the Trellis displays trigger an idea. Suppose that a power transformation of C is additive in the factors and is a linear function of the three numeric factors. That is,

$$\begin{aligned} C &= C(S, L, T, D) \\ &= (\mu + \alpha S + \beta L + \gamma T + \delta D)^{-p} \end{aligned}$$

where C is the clearing dose, L is the load, T is the temperature, and D is the duration, all in their original units, and S is an indicator variable that is -1 for solvent 1, and 1 for solvent 2. 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

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

So the derivative is everywhere negative and 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

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

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 we search for a power transformation of C as a way of removing or at least reducing the nonlinearity and the pervasive interactions among the factors. We repeated the Trellis displays in Figures 9 to 11 for a number of power transformations. The displays for the inverse square root are shown in Figures 9 to 11. 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 5.3 shows an analysis of variance for $100/\sqrt{C}$, carried out in the same manner as in Table 5.1. The new table also suggests an absence of nonlinearity and interaction.

We fitted the linear, additive model and computed two sets of residuals. The first set is on the transformed scale. In addition, we transformed the fitted values back to the original scale, and then subtracted the transformed fit from clearing dose on the original scale.

Trellis displays of the residuals on the transformed scale suggest our additive model has no appreciable lack of fit.

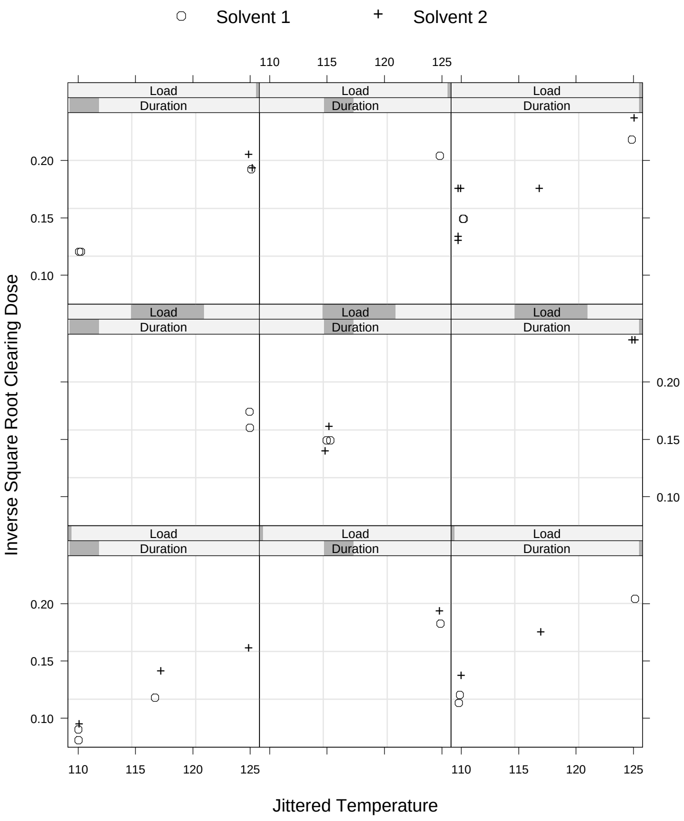

Figure 9: Trellis display of $1/\sqrt{C}$ against T given D, L, and S.

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		

Table 3: Analysis of variance for 100 times the transformed resist data.

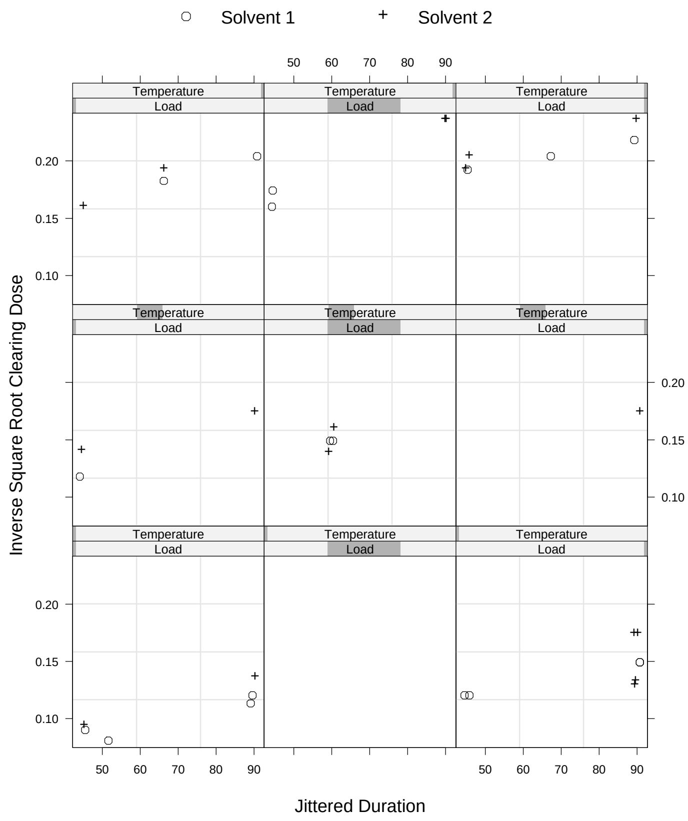

Figure 10: Trellis display of $1/\sqrt{C}$ against D given L, T, and S.

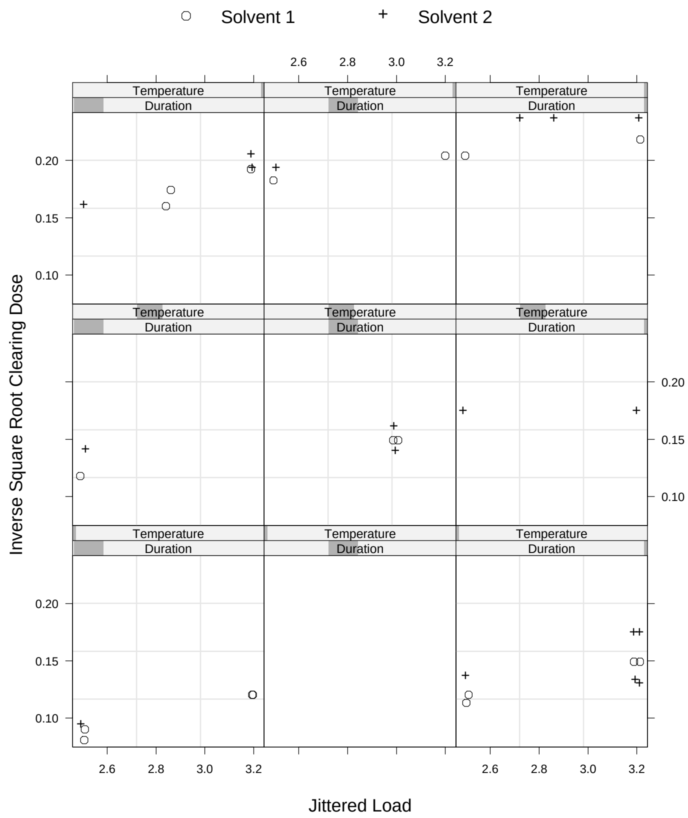

Figure 11: Trellis display of $1/\sqrt{C}$ against L given D, T, and S.

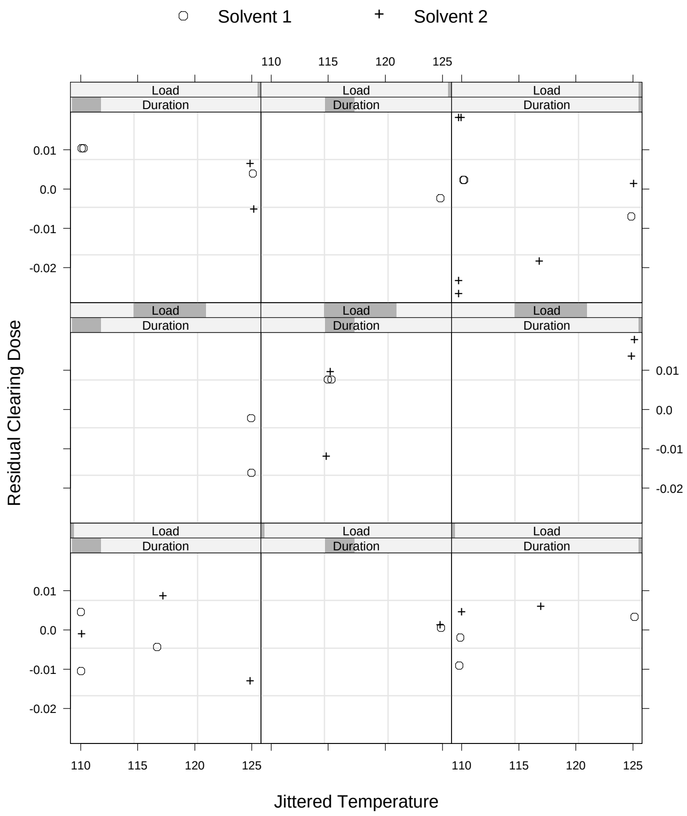

Figure 12: Trellis display of residual $1/\sqrt{C}$ against T given D, L, and S.

One such residual display is shown in Figure 12, a Trellis graph of the residuals against T given D, L, and S.

Figure 13 shows spread-location, or s-l, plots of the residuals and fitted values on both scales (Cleveland, 1993). In each case the square root absolute residuals are plotted against the fitted values and a loess curve superposed with local linear fitting and a nearest neighbor bandwidth parameter of 1; the purpose of the s-l plot is to detect a dependence of the variability of the response on the level of the fitted values. On the original scale there is a dependence, but not on the transformed. Thus the inverse square roots have removed a non-homogeneous variance. On the original scale, the non-homogeneous variance results in an outlier, which is brought into harmony with the other observations by the transformation.

Figure 14 shows normal probability plots of the residuals. On the original scale, the residuals are slightly skewed to the right. On the transformed scale, the residuals are slightly skewed to the left. Neither distribution deviates appreciably from normality.

Figure 15 shows a residual-fit spread plot, or rfs plot, of the residuals and fitted values on the transformed scale (Cleveland, 1993). The right panel is a quantile plot of the residuals and the left panel is a quantile plot of the fitted values minus their mean. The purpose of the display is to compare the amount of variation in the residuals with

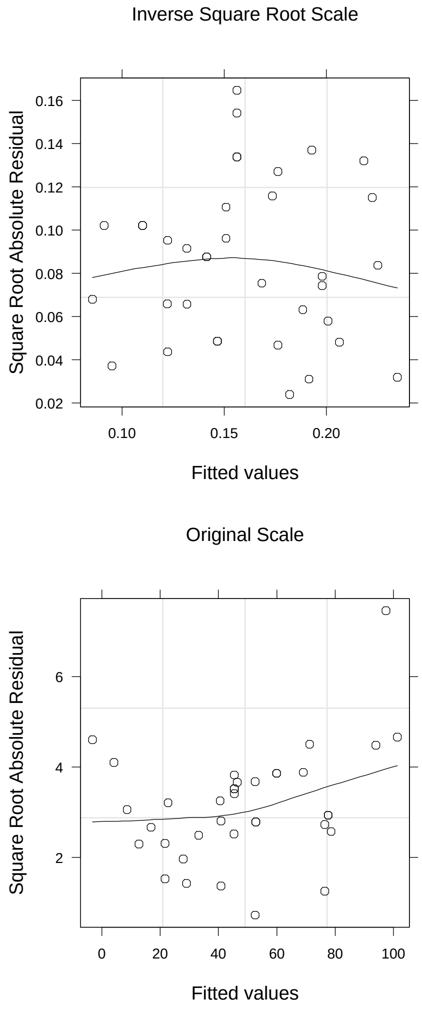

Figure 13: S-L plot.

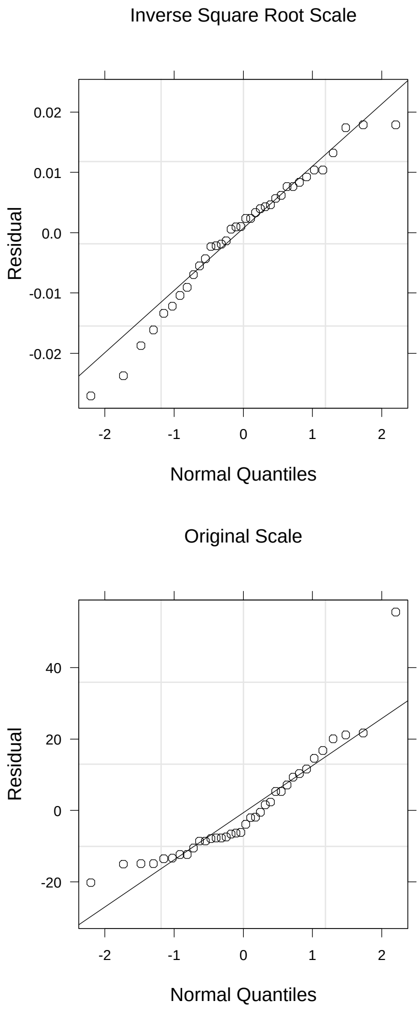

Figure 14: Q-Q plot.

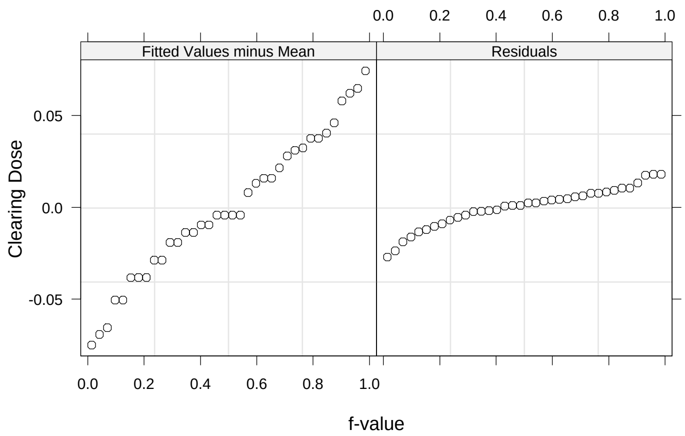

Figure 15: RFS plot.

the amount of variation in the fitted values. For the resist data, our model explains much of the variation in the data.

5.4 Trellis Plots of the Fitted Response Surface

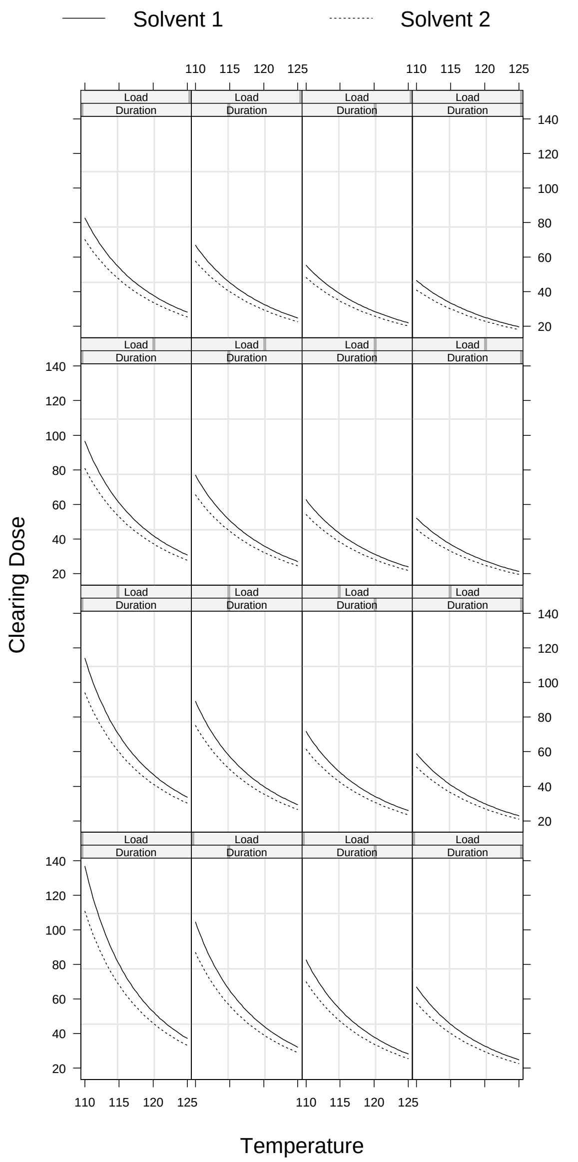

Figure 16: Trellis display of fitted C against T given D, L, and S.

We will study the effects of the factors not on the transformed scale — the scale on which we modeled and did the fitting — but rather on the original scale, to appeal to engineering intuition about clearing dose, which resides on the mJ/cm^2 scale. Again, we went back to the original scale simply by inverse transforming the fit on the transformed scale, although in a more delicate analysis we might want to attempt a more refined back transformation that takes sampling variability into account. The response surface is displayed by Trellis plots in Figures 16 to 18. There is one plot for each numeric factor.

In Figure 16, C is graphed against T given D, L, and S. Suppose that the fitted surface is $\hat{C}(D,T,L,S)$. Consider the lower left panel. The upper curve is graph of

$$\hat{C}(D, T = 110°\text{C}, L = 2.5\% \text{ wt}, S = \text{solvent } 1)$$

against D in the range 45 sec to 90 sec. For the lower curve, S = solvent 2. There are 16 such panels arising from all combinations of 4 equally spaced values of T and 4 equally spaced values of L. The 4 values of T range from 110° C to 125° C, the range of the observations of T; the 4 values of L range from 2.5 % wt to 3.2 % wt, the range of the values of L. Also, the range of evaluation as a function of D, 45 sec to 90 sec, is the range of the observations of D. Thus as a function of the numeric factors, we are studying the response surface over the rectangular solid that just contains the values of the numeric factors in the design space. Figures 17 and 18 are constructed in a similar manner.

Figures 16 to 18 show the nonlinearity and the strong interactions among all factors revealed in our initial Trellis plots of the data, albeit far more incisively here. For example, in Figure 16, 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 three quantities — C, dC/dT, d^2C/dT^2 — all decrease.

6 Liquid Crystal Data

6.1 Polymer-Dispersed Liquid Crystal Displays

Portable electronic computer products need lightweight display devices that can be easily read in both artificial and natural lighting. Reflective displays that are visible in ambient lighting and operate without back lights reduce weight and power requirements. Polymer dispersed liquid crystals are promising materials for reflective displays because they do not require polarizers and appear white in the unswitched state.

Polymer-dispersed liquid crystals (PDLCs) are made by inducing phase separation in a homogeneous mixture of liquid crystal and monomers by polymerizing the mixture using one of a variety of processes such as uv curing. As the monomer polymerizes, the liquid crystal separates into ellipsoidal droplets separated by the polymer. Under normal conditions, the droplets 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.

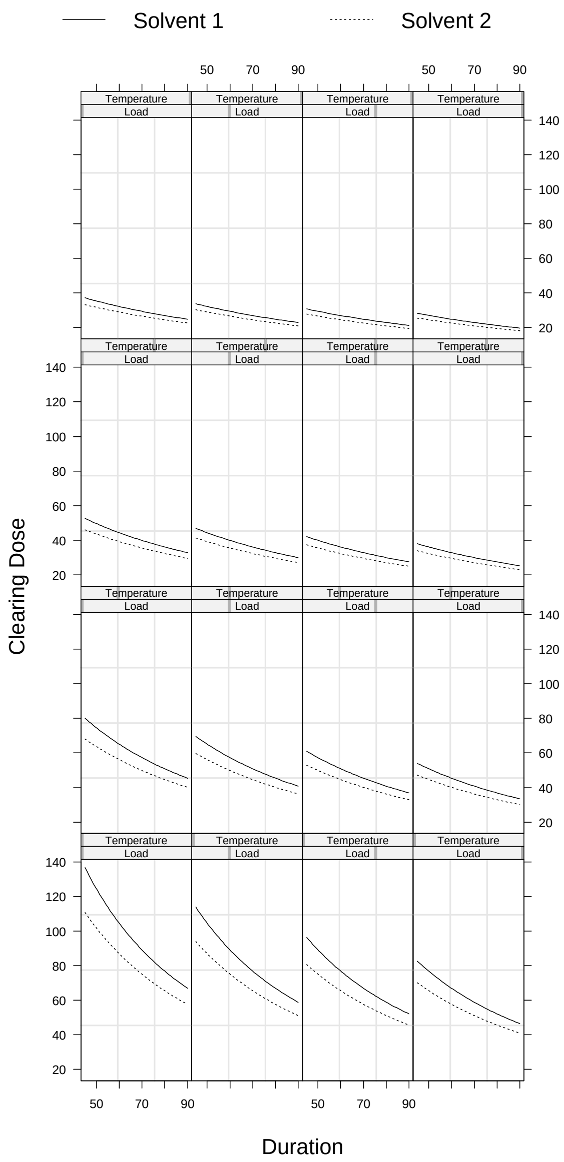

Figure 17: Trellis display of fitted C against D given L, T, and S.

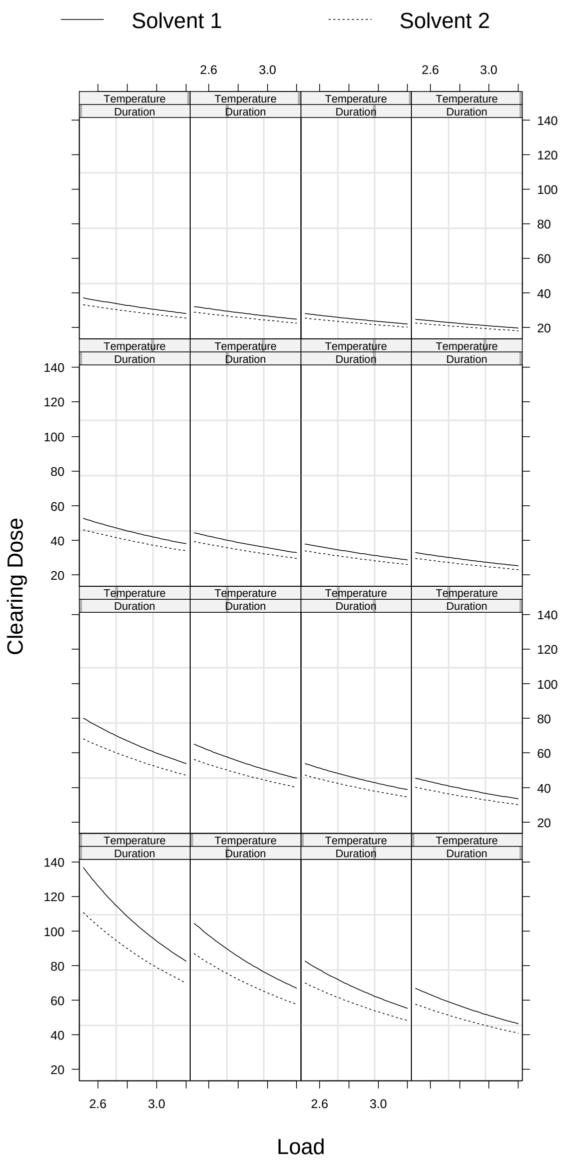

Figure 18: Trellis display of fitted C against L given D, T, and S.

The switching voltage — the voltage necessary to align the droplets — depends on the morphology of the liquid crystal droplets, which in turn depends on a number of factors including the weight percentage of liquid crystal in the mixture, as well as the temperature and uv light intensity during the polymerization.

6.2 The Experiment

LeGrange, Carter, Fuentes, Boo, Freeny, Cleveland and Miller (1997) describe experiments to determine the opto-electrical properties of PDLCs. One series of experiments studied the dependence of switching voltage, V, on the three factors mentioned above — the amount, M, of liquid crystal in the mixture, measured in wt %; the intensity, I, of the light used in the processing, measured in mW/cm^2; and 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.

6.3 Analysis of Variance

Table 6.3 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, 6.2 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 yield an unambiguous model specification. But 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 describes the structure in the 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		

Table 4: Analysis of variance for liquid crystal data.

6.4 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.

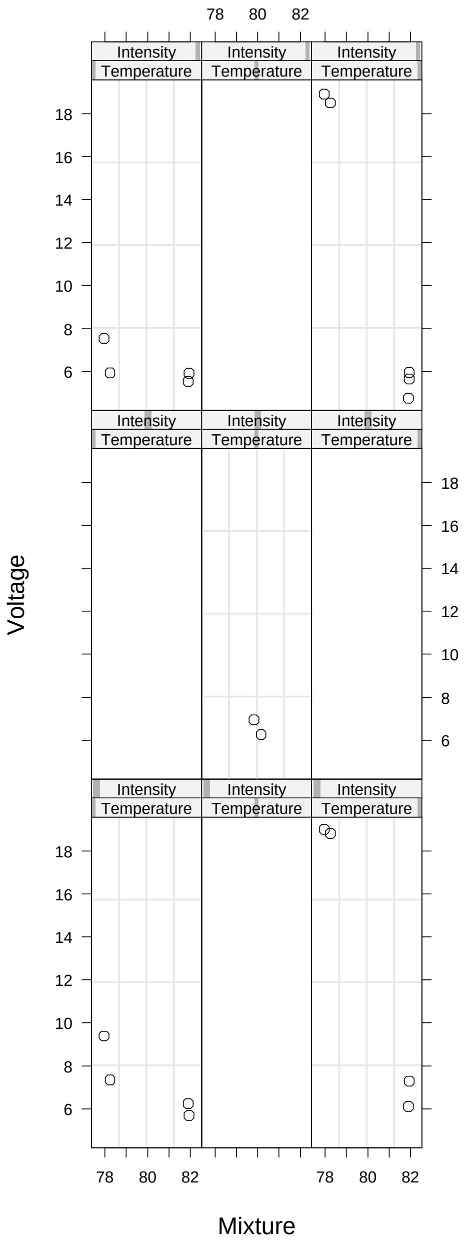

Figure 19: Trellis display of V against M given T and I.

Figure 19 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

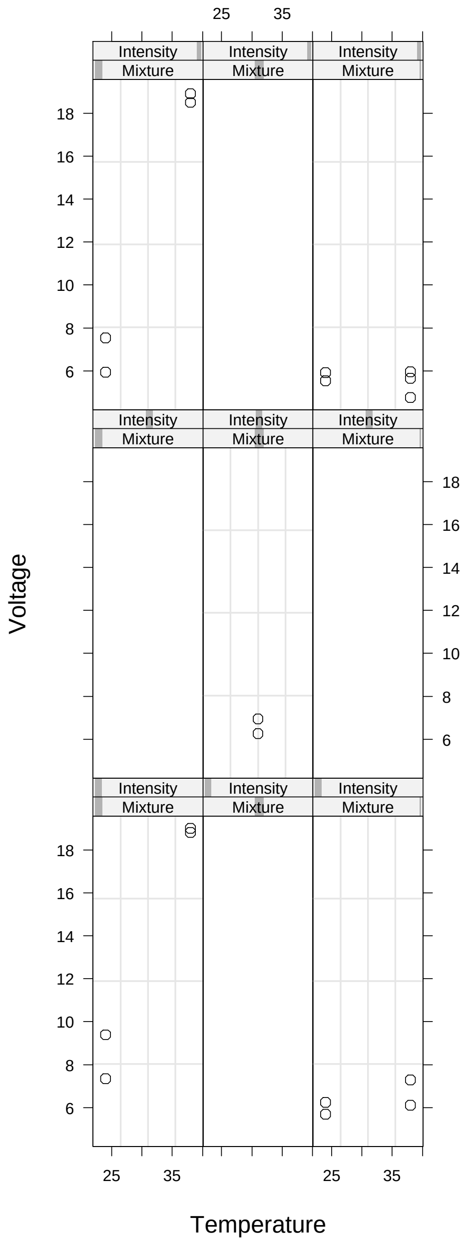

Figure 20: Trellis display of V against T given M and I.

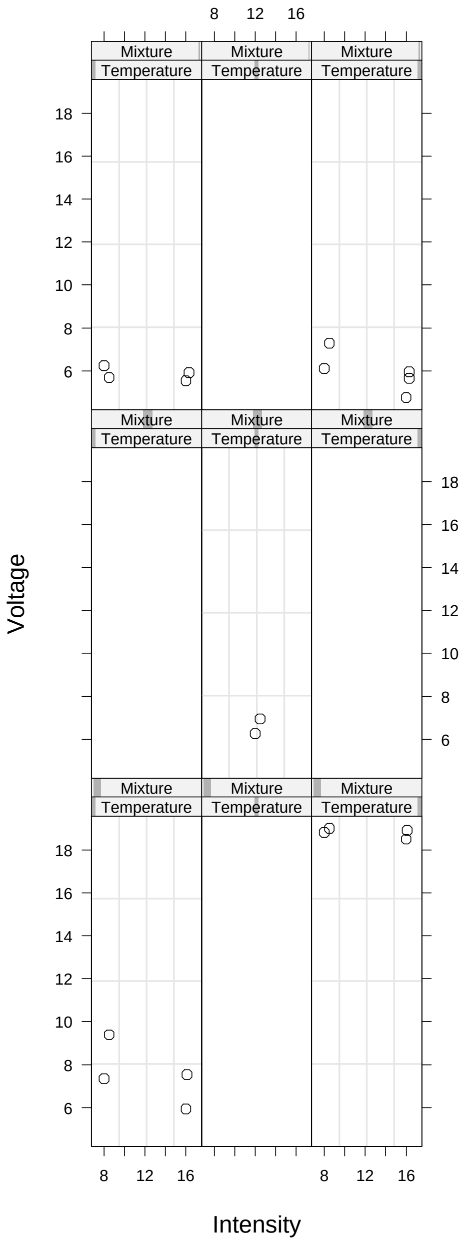

Figure 21: Trellis display of V against I given M and T.

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.

Figure 20 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.

Figure 21 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.

6.5 Modeling the Data

The anova carried out earlier was predicated on a quadratic dependence of V on the factors. 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, voltage is large for M at the lowest level and T at the highest, and much smaller and nearly flat elsewhere. Let us describe the structure we observed in the Trellis displays in terms of Figure 22, a scatterplot of the measurements of M and I with a small amount of uniform random noise 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 22, 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 22 has zero slope.

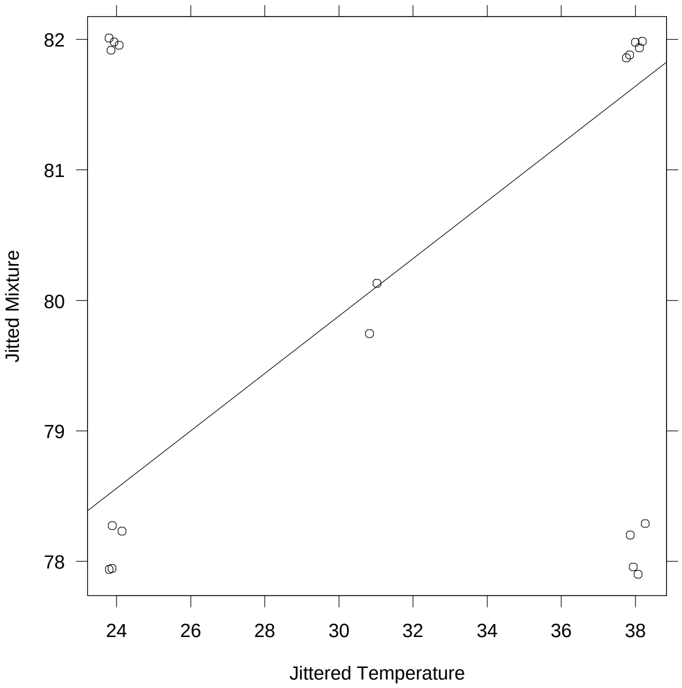

Figure 22: Measurements of M and T. The line is the estimated join line of the spline surface.

Thus the model is

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

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 normally distributed with mean 0 and constant variance σ^2. Thus the parameters will be estimated by nonlinear least-squares.

The line drawn in Figure 22 is the estimated join line,

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

The residual sum of squares is 4.8 V^2, and $\hat{\sigma} = 0.55$ V, far better than the quadratic models. Residual plots suggest there is no significant lack of fit and that the normal distributional assumption for the ϵ_i is reasonable.

Figure 23 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 some of the more subtle behavior in the Trellis displays of the data in Figures 19 to 21. We saw in the first column of Figure 19, where T is at the lowest level, there is a small decrease in V when M goes from the lowest level to the highest. Figure 22 shows why. For T at the lowest level, when M changes from highest to lowest, the points go from the region where the spline fit is low and constant, to the region with a changing V,

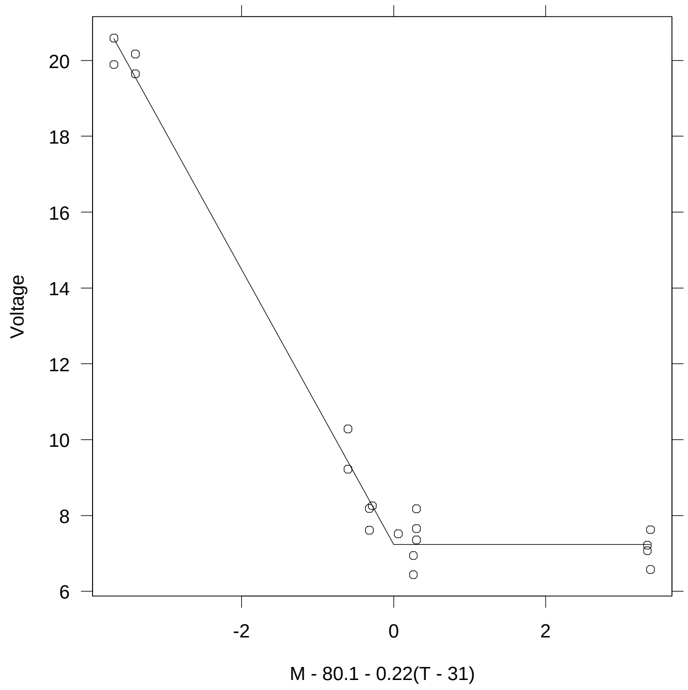

Figure 23: Partial residual plot.

and in the changing region, the points are sufficiently far from the join line that there is a detectable increase in V. By contrast, we saw in Figure 20 that when M is at the highest level, there is little change in V when T increases. Figure 22 shows that for M at the highest level, all points are in the region where the spline fit is constant.

The spline model clearly fits the available data. There are large gaps in the design space where the fit is unsupported, but the evidence for the model is as strong as we can expect from a limited pilot experiment. In fact, the spline model is also supported by physical theory. The phase separation boundary of the mixture of monomer and liquid crystal is a line in the M and T space whose slope is close to that of the join line of the spline. Despite the agreement of theory and the pilot results, more data were taken and the spline model was confirmed (LeGrange et al., 1997).

7 Discussion

Our investigations have led to several conclusions about building models for experimental data. Trellis display is often quite useful for modeling data from designed experiments, even small experiments with a moderate number of factors and a limited number of runs. The analysis of variance, pervasively used 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. The addition of Trellis display substantially increases the ability of the data analyst to carry out model identification and diagnostic checking, often resulting in more parsimonious and better fitting models.

7.1 Why and When is Trellis Display Useful

The design and layout of Trellis provides a flexible mechanism for visualization by conditioning which often does an effective job of revealing the structure of interactions in experimental data, even quite high order interactions. And for experimental data, fathoming interactions is one of the critical modeling tasks.

At the onset of our investigations it seemed likely that Trellis would not be useful for quite small experiments because the conditioning mechanism of Trellis would lead to too few observations in the panels. But, in fact, Trellis did typically prove useful in such cases. The resist data and the liquid crystal data are two examples. The reason for the success is that the variability in the error terms is small compared with the effects; when we condition on several factors, there are just a few points on each panel, but patterns emerge because the amount of noise is small. We believe our experience does not constitute a few unusual, lucky examples. Part of the art of experimentation is using knowledge of the mechanism generating the data to choose factors and a design so that the noise is small; since the art is well practiced by many, experimental data sets with limited noise abound.

The examples in which we found Trellis not particularly useful were those in which conditioning led to one or fewer points on each panel; these were very highly fractionated experiments. This is not a limitation of Trellis display but rather a limitation of such experiments. The data do not contain within in them enough information to validate models. This is quite appropriate in cases where the cost of an experimental run is high, and where our understanding of the mechanism is quite substantial, allowing us to specify a model with a high degree of certainty.

7.2 The Analysis of Variance

The analysis of variance is heavily used as a tool for building models for experimental data. A common approach is to hypothesize a model rather liberal in effects, and then cast out many of them using anova. There are three problems with this. First, it requires a model in the first place, which while seemingly quite general can in fact not contain a case that does justice to the data. Second, it is not an effective approach to discovering a simple model for the data when a simple model provides a good fit. One reason is that simple models often require a re-constituting of the factors, or of the response, for example, by transformation. Third, the liberal use of effects can result in certain of them being judged as insignificant because of the large number of degrees of freedom used in the fitting, but a more parsimonious model shows the effects to be highly significant.

Consider as an example, two-way data with two categorical factors and the same number of replications for each combination of levels of the two factors. A standard model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

for $i = 1$ to $a, j = 1$ to b, where

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_{ij} \gamma_{ij} = 0.$$

One attractive aspect of this model is the orthogonality of the terms; this makes testing, in particular, anova, simpler to interpret. But the importance of simplicity of interpretation of some testing procedure is far less important than the importance of finding models for experimental data that are valid and that are as parsimonious as possible.

The γ_{ij} of the liberal-parameter model are quite general and describe interactions of very different types. One particular interaction that occurs commonly is

$$\gamma_{ij} = \gamma \alpha_i \beta_j$$

This is often referred to as "Tukey's one degree of freedom for non-additivity" (Tukey, 1949). Mandel (1995) discusses the modeling of two-way data by this and similar models.

If Tukey's model is appropriate for a set of two-way data, then it is desirable to use it in fitting the data in place of the liberal-parameter model. Simply appreciating that Tukey's model fits the data leads to substantial insight into the structure of the data. And Tukey's model provides a much more parsimonious quantitative description of the data. Finally, if the magnitude of γ is not large, it can happen that the liberal-parameter fails to detect the interaction, but fitting the Tukey model does so.

In the examples of this paper, anova provided answers to rather specific questions about the data, but anova was far less effective than Trellis in discovering the structure of the data and in making decisions about the form of the model.

The general opinion expressed here about anova is hardly new. For example, Mandel (1995) writes: "Analysis of variance (Anova) is considered by many as the ultimate 'exact' method for analyzing data. It is nothing of the sort."

References

Becker, R. A. and Cleveland, W. S. (1996a). The Design and Control of Trellis Display, *Journal of Computational and Statistical Graphics* **5**: 123–155.

Becker, R. A. and Cleveland, W. S. (1996b). *Trellis Graphics User's Manual*, Mathsoft, Seattle.

Cleveland, W. S. (1993). *Visualizing Data*, Hobart Press, Summit, New Jersey, U.S.A.

Cleveland, W. S. (1994). *The Elements of Graphing Data*, Hobart Press, Summit, New Jersey, U.S.A.

Davies, O. L. (1967). *The Design and Analyisis of Industrial Experiments*, 2nd edn, Hafner, New York.

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

LeGrange, J. D., Carter, S. A., Fuentes, M., Boo, J., Freeny, A. E., Cleveland, W. and Miller, T. M. (1997). The Dependence of the Electro-Optoical Properties of Polymer Dispersed Liquid Crystals on the Photopolymerication Process, *Journal of Applied Physics* **81**: 1–8.

Mandel, J. M. (1995). *Analysis of Two-Way Layouts*, Chapman & Hall, New York.

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

Nalamasu, O., Freeny, A., Reichmanis, E., Sloane, N. J. A. and Thompson, L. F. (1993). 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.

Snee, R. D. (1985). Experimenting with a Large Number of Variables, *in* R. D. Snee (ed.), *Experiments in Industry*, American Society for Quality Control, Milwaukee, Wisconsin, U.S.A., pp. 25–35.

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

Tukey, J. W. (1949). One Degree of Freedom for Non-Additivity, *Biometrics* **1949**: 232–242.

Tukey, P. A. and Tukey, J. W. (1981). Graphical Display of Data Sets in 3 or More Dimensions, *in* V. Barnett (ed.), *Interpreting Multivariate Data*, Wiley, Chichester, U. K., pp. 189–275.

W. F. Hunt, J. (1985). Experimental Design in Air Quality Management, *in* R. D. Snee, L. B. Hare and J. R. Trout (eds), *Experiments in Industry*, American Society for Quality Control, Milwaukee, pp. 89–98.