

# **Detecting Clusters and Nonlinearity in 3D Dynamic Graphs**

John Fox  
McMaster University

Robert Stine  
University of Pennsylvania

Georges Monette and Nehru Vohra  
York University

Interface 2002

# 1 Introduction

- Three-dimensional dynamic scatterplots have been promoted as
  - geometrical tools for understanding statistical concepts
  - tools for analyzing data
- 3D scatterplots can reveal certain features of data that cannot be apprehended in conventional two-dimensional displays:
  - some kinds of clustering
  - some kinds of nonlinearity (interaction)
- Originally the province of experimental graphical systems, 3D dynamic scatterplots are now found in standard statistical software packages.
- But 3D dynamic scatterplots can be difficult to decode.

- In an experiment on cluster detection, we seek to establish
  - whether data analysts are able to discern clustering in 3D plots
  - if so, whether the probability of detection varies by easily characterized properties of the clusters
  - whether the design of the display is related to cluster detection
- Three further experiments investigate similar issues with respect to the detection of nonlinearity.

# 2 Design of the Experiments

## 2.1 Cluster Detection

- Data Generation
  - Subjects viewed displays of rotatable 3D point clouds. The subjects were asked to report whether they saw one or two clusters of points in each display.
  - All data displays included 100 points, either as a single trivariate-normal cluster or as two trivariate-normal clusters with 50 points each, and identical within-cluster covariance matrices but different centroids
  - Cluster centroids were either displaced *diagonally*, that is along the vector  $(1,1,1)'$ , or *horizontally*, that is parallel to the vector  $(1,1,0)'$ .
  - We maintained the same overall covariance matrix for all diagonally displaced displays (including those at 'zero separation'); a similar procedure was employed for horizontally displaced clusters.

- When the plot rotates around an axis through the center of the data perpendicular to the horizontal plane, the detection of diagonally displaced clusters depends less critically upon the specific orientation of the point-cloud.
- Data were generated at three levels of average separation:
  - (i) no separation — a single cluster of 100 observations
  - (ii) a low level of average separation
  - (iii) a medium level of average separation
- Expected data ellipsoids for the various kinds of clusters employed in the experiment are illustrated in Figure 1.

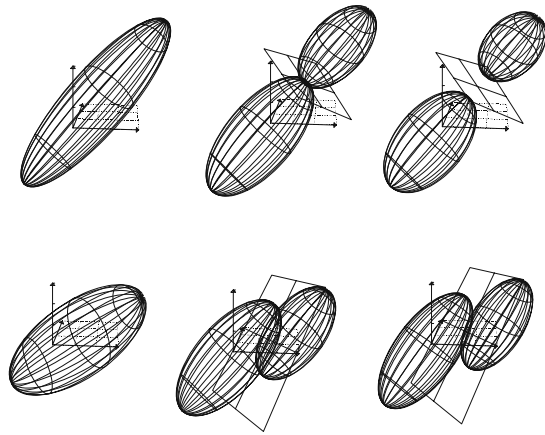


Figure 1. Fifty-percent concentration ellipsoids and planes of separation for diagonally oriented clusters (top) and horizontally oriented clusters (bottom). The level of separation of clusters increases from left to right; the graphs at the far left show single clusters (no separation).

Fox, Stine, Monette, and Vohra

Interface 2002

### • Display Elements

- We concentrated on those aspects that seem particularly relevant to perceiving the relative positions of points in 3D space: perspective, depth-cueing, and motion of the display.
  - \* Perspective was set at three level: none (i.e., orthographic projection), medium, and high.
  - \* Depth-cueing was performed by varying the color-saturation of objects on the screen according to their distance, close points appearing more saturated.
  - \* Motion of the plot was either continuous, by rotation in the horizontal plane, or under the control of the subject.
- Our subjective preference is for a display with depth-cueing and moderate perspective.

Fox, Stine, Monette, and Vohra

Interface 2002

- We find a continuously rotating display easier to look at and adequate for the present task.
  - \* This effect can be achieved by manipulating plot controls, but clumsy use of the controls can also prove disorienting and time-consuming
  - \* More generally, there is no guarantee that a simple automatic strategy such as horizontal rotation will reveal structure in a 3D point cloud.

Fox, Stine, Monette, and Vohra

Interface 2002

### • Software

- The software used to conduct the experiment was written in Lisp-Stat, and is a modified version of the programs described in some detail by Fox and Stine (1998 Interface).
- We modified the low-level handling of 3D plots in Lisp-Stat to incorporate perspective and color-based depth-cueing.

Fox, Stine, Monette, and Vohra

Interface 2002

- Procedures

- Ten volunteer graduate students were recruited for the study. Each subject participated in five sessions, each session lasting approximately one-half hour.
- A session comprised 72 trials (3 levels of separation  $\times$  2 cluster orientations  $\times$  3 levels of perspective  $\times$  2 levels of depth-cueing  $\times$  2 levels of motion, in random order).
- On each trial of the experiment, subjects were asked “to determine whether there are one or two clusters (groups) of points in the data for each graph.” The stimulus graph was visible until the subject responded, to a maximum of 30 seconds.

## 2.2 Detection of Nonlinearity

- Data Generation

- Subjects in three experiments on detecting nonlinearity viewed displays of rotatable 3D point clouds, and were asked whether the response variable in the graph ( $Y$ ) is linearly or nonlinearly related to the two predictors ( $X_1$  and  $X_2$ ).
- All of the datasets employed in the study included 100 randomly generated points.
- Values of the response variable were constructed according to the full-quadratic model

$$Y = X_1 + X_2 + \alpha(X_1^2 + X_2^2 + \sqrt{2}X_1X_2) + \sigma\varepsilon$$

where  $X_1, X_2, \varepsilon \sim \text{NID}(0, 1)$ , and the values of  $\alpha$  and  $\sigma$  were selected to generate different (expected) levels of nonlinearity and correlation.

- We selected three levels of correlation or ‘signal’ (defined as the expected  $R^2$  for the model:  $1/3$ ,  $1/2$ , or  $2/3$ ), and four levels of nonlinearity (defined as the expected proportion of the  $R^2$  for the model due to its nonlinear component:  $0$ ,  $1/3$ ,  $1/2$ , or  $2/3$ , producing 12 fundamental stimulus configurations.

- Figure 2 illustrates the expected response surface corresponding to a relatively high degree of nonlinearity (for which  $\alpha = \sqrt{2/3}$ ).

- The three experiments used the same 12 basic stimulus configurations, but manipulated different aspects of the displays.

- Procedures generally similar to the experiment on clustering.

- **Experiment 1.** General Design: Four variations of spinning 3D scatterplots that are common, and one that is unusual (Figure 3):

- (a) A bare point cloud (except for labelled axes).
- (b) A display with the least-squares regression plane shown as a gray wire-frame.

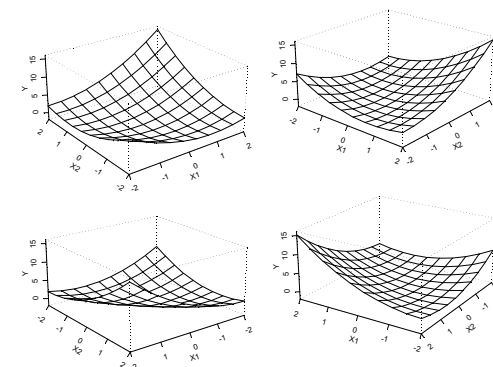


Figure 2. The response surface  $E(Y) = X_1 + X_2 + \sqrt{2/3}(X_1^2 + X_2^2 + \sqrt{2}X_1X_2)$  from several points of view.

- (c) A display that showed the least-squares plane along with color-coded residuals.
  - (d) A display spinning in the regression plane (rather than the predictor plane); the regression plane, which was shown, was oriented horizontally and viewed edge-on as a line.
  - (e) A display that tied the points to the predictor plane with vertical lines (“balloons”).
- **Experiment 2. Motion**
    - (a) Automatic rotation in the predictor plane.
    - (b) Automatic rotation in the regression plane.
    - (c) Subject-manipulated motion.
  - **Experiment 3. Perspective and Depth Cueing**
    - No perspective, “low” perspective, “high” perspective (positioning the viewpoint relatively close to the data).
    - Color-based depth cueing: on or off.

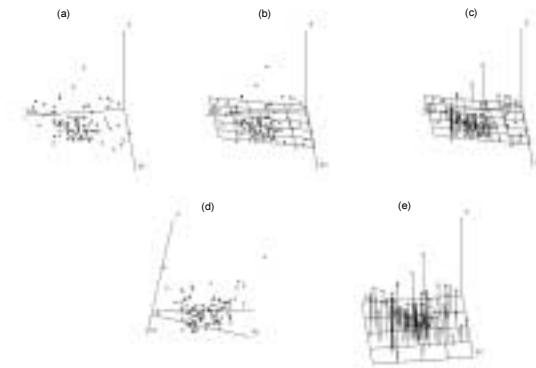


Figure 3. Five 3D dynamic scatterplot displays: (a) bare point cloud; (b) point cloud with regression plane; (c) point cloud with regression plane and residuals; (d) horizontally oriented regression plane, with display spinning in the regression plane; (e) “balloons.” In the experiment, the graphs were in color against a black background.

## 3 Results

### 3.1 Cluster-Detection Experiment

- Figure 4 shows the general relationship between the probability of detecting two clusters and the degree of separation of the clusters, as measured by Hotelling’s  $T^2$ .
- Similar graphs drawn separately for each subject appear in Figure 5.
- In examining the accuracy of subjects’ judgments, we restricted attention to trials in which two clusters were present. The data on accuracy for single-cluster stimuli were uninteresting, because most subjects had a very low level of false positives.
- The significant high-order terms in a fixed-effects logit model fit to the data are graphed in Figure 6.

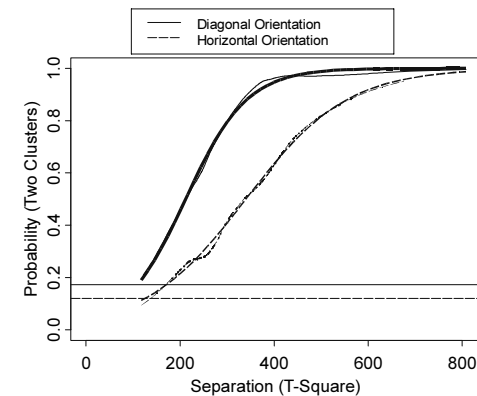


Figure 4. Fitted probability of detecting two clusters by cluster separation ( $T^2$ ) and cluster orientation (diagonal or horizontal). The heavier curves show linear logistic regressions; the lighter curves show locally linear logistic regressions. The two horizontal lines show the proportion of false positives.

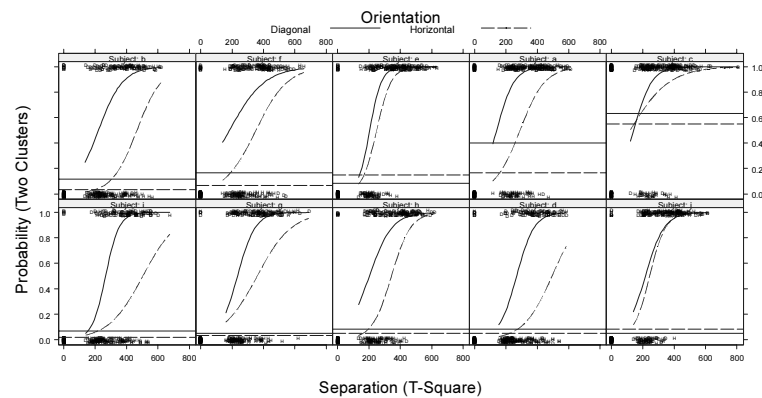


Figure 5. Cluster detection by subject. Two subjects had high levels of false positives.

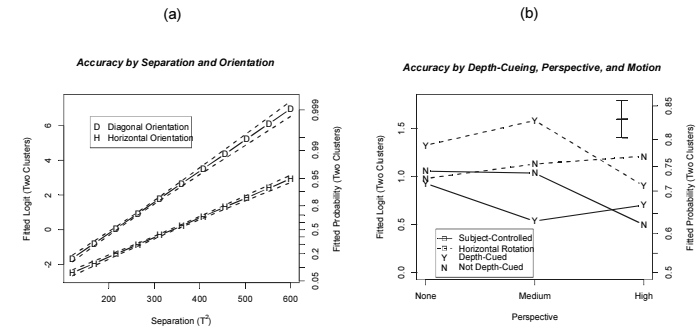


Figure 6. Graphs of effects in a follow-up logit model fit to the data on detecting two clusters. The broken lines in panels (a) give  $\pm$  one standard error around the fit. The bar at the upper right of panel (b) gives twice the average standard error of the fitted values (i.e., an interval of  $\pm$  one standard error). The horizontal axis in panel (a) runs to  $T^2 = 600$ , by which point the fitted probability of detecting diagonally oriented clusters is essentially 1.

- The interaction between separation and orientation, shown in Figure 6 (a), squares with our previous discussion of the data (see Figure 4): Subjects were generally more likely to detect two clusters if they were separated diagonally than if they were separated horizontally; this difference is small when the separation of the clusters is relatively slight and grows (on the logit scale) as separation increases.
- Figure 6 (b) indicates that certain combinations of design characteristics proved more effective than others in producing accurate judgments, but differences among the fitted values in this graph are not very large relative to the standard error of the fits. In most instances, subjects performed better — though to varying degrees — when the display rotated automatically than when they controlled its movement directly. Beyond this point, it is difficult to draw simple conclusions because the results for automatically rotating displays are nearly the mirror of those for subject-controlled displays.

## 3.2 Nonlinearity-Detection Experiments

### • Experiment 1

- We fit a nonparametric logistic regression of subjects' judgments on the degree of nonlinearity and signal in the datasets. The results suggested fitting a parametric logistic regression with linear terms in both predictors together with their product, accounting for 99 percent of the non-null deviance captured by the nonparametric regression. (Generally similar results were also obtained for experiments 2 and 3.)
- The response surface in Figure 7 suggests that at all levels of signal, the probability of reporting nonlinearity rises with the strength of nonlinearity. The relationship is much stronger, however, when the signal is strong. As well, at low levels of nonlinearity, the probability of reporting nonlinearity is inversely related to the strength of the signal, while at high levels of nonlinearity, the relationship is direct.

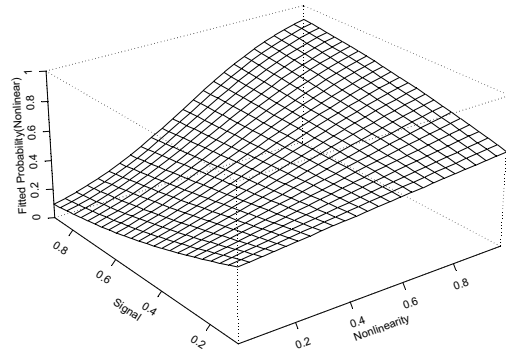


Figure 7. Fitted probability of identifying a dataset as nonlinear as a function of nonlinearity and signal, for Experiment 1. The fit is from a linear logistic regression of subjects' responses on nonlinearity, signal, and their interaction. Note that the 'origin' is the corner closest to the viewer.

Fox, Stine, Monette, and Vohra

Interface 2002

- Then we examined the relationship between each subject's responses and the linear predictor from the previous logistic regression (which we term 'discernibility'), fitting nonparametric and linear logistic regressions to the data. Two subjects tended to report nonlinearity even when discernibility was very low, while another tended to fail to report nonlinearity even when discernibility was high.

Fox, Stine, Monette, and Vohra

Interface 2002

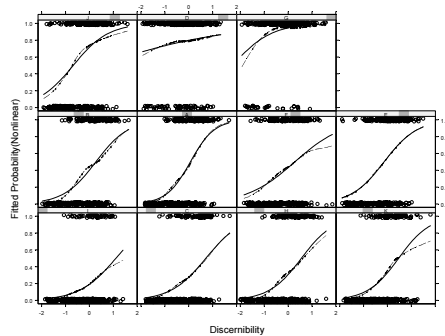


Figure 8. Fitted probability of identifying a dataset as nonlinear as a function of 'discernibility,' plotted by subject, for Experiment 1. The observations (1 = 'yes', 0 = 'no') are jittered slightly to reduce overplotting. The solid line represents a linear logistic regression, the broken line a nonparametric logistic regression. Subjects are ordered (from lower left to upper right) by their overall level of 'yes' responses.

Fox, Stine, Monette, and Vohra

Interface 2002

- We turned next to examining whether subjects' judgments were related to variations in the displays. Results appear in Figures 9 and 10. At low levels of nonlinearity, the probability of reporting nonlinearity was generally similar across the five types of displays, with differences widening as nonlinearity increased. The display that showed a least-squares plane and rotated in the predictor plane (labelled 'P') seemed least effective. Displays of a bare point cloud ('N'), displays employing balloons ('B'), and displays showing residuals along with the regression plane ('R') appeared about equally, and intermediately, effective. The most successful display was the one spinning horizontally in the regression plane ('S').

Fox, Stine, Monette, and Vohra

Interface 2002

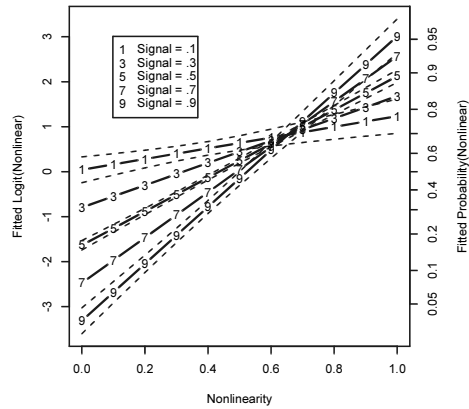


Figure 9. Interaction between signal and nonlinearity in Experiment 1; the broken lines represent  $\pm$  one standard error around the fit, at levels .1, .5, and .9 of signal. Despite plotting only at several levels of signal, both nonlinearity and signal are continuous variables.

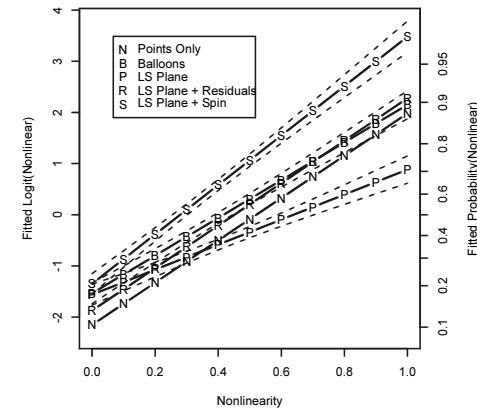


Figure 10. Interaction between type of display and nonlinearity in Experiment 1.

• Experiment 2

- The general response of subjects in the second experiment was very similar to that in the first. We concentrate, therefore, on the novel features of this experiment.
- The three-way interaction among motion, signal, and nonlinearity was nearly statistically significant, and we chose, therefore, simply to graph the fit of an initial model, averaging over subjects. The result appears in Figure 11.
- The interaction between signal and nonlinearity is of the now-familiar form for subject-manipulated motion and for motion in the regressor plane. For a bare point cloud rotating in the predictor plane, however, subjects did not achieve a very high level of detection of nonlinearity, even when the degree of nonlinearity and the signal were quite high; moreover, they did no better at a high level of nonlinearity with a strong signal than with a weak one.

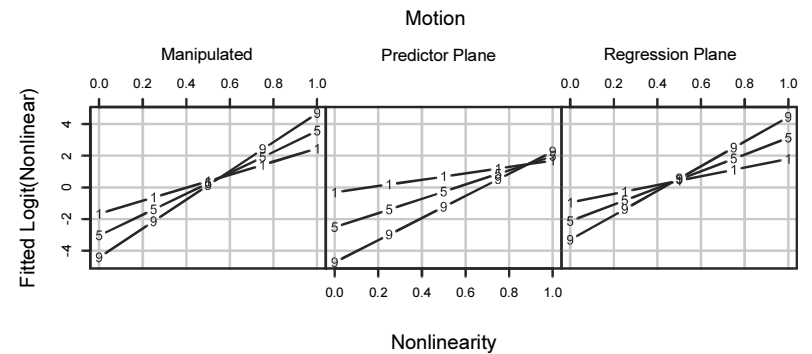


Figure 11. Fitted logit of reporting nonlinearity by type of motion, nonlinearity, and signal in Experiment 2. The lines are labelled by the level of signal: 1:  $R^2 = .1$ ; 5:  $R^2 = .5$ ; and 9:  $R^2 = .9$ . Both nonlinearity and signal are continuous variables.

### ● Experiment 3

- A preliminary fixed-effects logistic-regression model showed a statistically significant four-way interaction among depth-cueing, perspective, signal, and nonlinearity. This interaction is plotted in Figure 12.
- Except in the high-perspective, depth-cueing condition, we see the familiar general interaction between nonlinearity and signal, with detection of nonlinearity more responsive to the degree of nonlinearity when the level of signal is high.
- Depth-cueing and perspective did not enhance subjects' performance in this experiment: They performed at least as well in the no-perspective, no-depth-cueing condition as in any of the others.

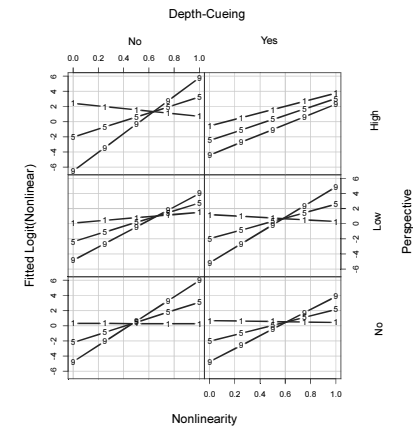


Figure 12. Experiment 3: Fitted logit of reporting nonlinearity by depth-cueing, perspective, nonlinearity, and signal in Experiment 3. The lines are labeled by the level of signal: 1:  $R^2 = .1$ ; 5:  $R^2 = .5$ ; and 9:  $R^2 = .9$ . Both nonlinearity and signal are continuous variables.

## 4 Discussion

- Subjects' ability to discern clusters in the 3D displays employed in this study responded in a reasonable manner to the properties of the data.
- On the other hand, subjects' performance was related relatively weakly to characteristics of the displays.
  - Orthographic dynamic 3D displays without depth-cueing, for example, are intrinsically ambiguous with respect to the direction of rotation, an ambiguity that is resolved by perspective or depth-cueing. Yet the use moderate perspective and depth-cueing was not necessarily advantageous.
  - We expected to see a more general disadvantage accruing to high-perspective displays, since exaggerated perspective tends to push relatively far-away points towards a common vanishing point at the center of the screen.

- We anticipated, but did not observe, an interaction between motion of the display (automatic horizontal rotation versus subject-controlled motion) and cluster orientation: Diagonally displaced clusters are easier to discern than horizontally displaced clusters in a continuous, horizontally rotating display, but this advantage does not hold in general.
- In the experiments on detecting nonlinearity, as in the cluster-detection experiment, most subjects responded in a reasonable manner to characteristics of the data, reporting nonlinearity at a greater rate when the degree of nonlinearity in the data was higher.
- Moreover, this relationship between reporting nonlinearity and the degree of nonlinearity was more pronounced when the signal was strong than when the signal was weak.

- How the detection of nonlinearity was related to display characteristics is less clear, but there were some patterns:
  - More elaborate displays did not perform especially well.
  - Spinning the plot horizontally in the regression plane proved particularly effective in both the first and second experiments.
- Taken together, then, the results of these two studies suggest that dynamic 3D scatterplots can be effective for the kinds of data-analytic tasks for which they hold theoretical promise.
- Moreover, designers of such displays should probably pay greater attention to orienting and rotating the point cloud in effective and revealing ways than to enhancements of the displays, such as displaying regression planes and employing depth-cueing or perspective.