

## COPLOTS, NONPARAMETRIC REGRESSION, AND CONDITIONALLY PARAMETRIC FITS

BY WILLIAM S. CLEVELAND

*AT & T Bell Laboratories*

Conditionally parametric regression surfaces provide parsimonious fits in the common situation where the effects due to some factors are small but significant, and the effects of other factors are large, complicated, and require nonparametric fitting. One approach to nonparametric regression is local regression, specifically the local fitting of linear and quadratic polynomials of the factors. Recent work of Fan and of Hastie and Loader has shown that local regression is superior to kernel estimation and modified kernel estimation, two methods that have had extensive theoretical investigation and that work poorly in practice. Local regression can be modified in a simple way to produce conditionally parametric fits. The coplot is a graphical method that is particularly helpful for carrying out regression studies, in particular, for determining factors that can be taken to be conditionally parametric.

**1. Introduction.** Graphical methods are critical tools for analyzing and modeling multivariate data. For example, coplots are particularly useful for regression studies, both parametric and nonparametric (Cleveland (1993 to appear)). Coplots are the first topic of the paper.

Many multivariate data sets are better fitted by nonparametric regression surfaces than parametric ones because the latter lack the flexibility to track all but very simple patterns. One method of nonparametric regression is *local fitting*. The method is an old idea first used by time series analysts to smooth their data (Macauley (1931)). It was brought to more general regression studies in the 1970s (Cleveland (1979), Stone (1977)). Making this computer-intensive method a practical one required that a number of computational methods be developed (Cleveland and Grosse (1991)), but this work proved successful, and local regression is now regularly used in regression studies (Cleveland et al. (1991)). There are a multitude of other nonparametric regression methods. One is splines (Reinsch (1967), Silverman (1985), Wahba (1978)). But for two or more factors, splines have resisted solutions to nasty computational problems, and at this writing are still an  $n^3$  operation. Another

---

AMS 1980 Subject Classifications: 62-09, 62G07.

Key words and phrases: Graphics, loess, smoothing.