DT# 49463 QA: NA
CB 10/12/06

One Line or Two? Perspectives on Piecewise Regression.

Robert P. Ewing, Iowa State Univ, Dept of Agronomy, Ames, IA 50011-1010 and David W. Meek, National Soil Tilth Lab., 2150 Pammel Dr., Ames, IA 50011-4420. .

Sometimes we are faced with data that could reasonably be represented either as a single line, or as two or more line segments. How do we identify the best breakpoint(s), and decide how many segments are "really" present? Most of us are taught to distrust piecewise regression, because it can be easily abused. The best method for identifying the breakpoint varies according to specifics of the data; for example, the minimum sum of squares method excels for "well-behaved" data. In some cases, hidden Markov methods are more likely to succeed than are more "obvious" methods. Likewise, the most appropriate method for deciding between one or two lines depends on your expectations and understanding of the data: an unexpected break requires more justification than an expected one, and some decision criteria (e.g., the Akaike Information Criterion) are less strict than others (e.g., the Bayesian Information Criterion). This presentation will review some options and make specific, practical recommendations.