Comparing Methods for Multivariate Nonparametric Regression

David L. Banks* Robert T. Olszewski†
Roy A. Maxion†

January 1999
CMU-CS-99-102

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

*Bureau of Transportation Statistics, Department of Transportation

†School of Computer Science, Carnegie Mellon University

This research was sponsored by the National Science Foundation under Grant #IRI-9224544.

Keywords: multivariate nonparametric regression, linear regression, stepwise linear regression, additive models, AM, projection pursuit regression, PPR, recursive partitioning regression, RPR, multivariate adaptive regression splines, MARS, alternating conditional expectations, ACE, additivity and variance stabilization, AVAS, locally weighted regression, LOESS, neural networks

Abstract

The ever-growing number of high-dimensional, superlarge databases requires effective analysis techniques to mine interesting information from the data. Development of new-wave methodologies for high-dimensional nonparametric regression has exploded over the last decade in an effort to meet these analysis demands. This paper reports on an extensive simulation experiment that compares the performance of ten different, commonly-used regression techniques: linear regression, stepwise linear regression, additive models (AM), projection pursuit regression (PPR), recursive partitioning regression (RPR), multivariate adaptive regression splines (MARS), alternating conditional expectations (ACE), additivity and variance stabilization (AVAS), locally weighted regression (LOESS), and neural networks. Each regression technique was used to analyze multiple datasets each having a unique embedded structure; the accuracy of each technique was determined by its ability to correctly identify the embedded structure averaged over all the datasets. Datasets used in the experiment were constructed to have a range of characteristics by varying the dimension of the data, the true dimension of the embedded structure, the sample size, the amount of noise, and the complexity of the embedded structure. Analyses of the results show that all of these properties affect the accuracy of each regression technique under investigation. A mapping from data characteristics to the most effective regression technique(s) is suggested.

Contents

1. Introduction

Regression analysis in high dimensions quickly becomes extremely unreliable; this phenomenon is called the "curse of dimensionality" (COD). There are three nearly equivalent formulations of the COD, each offering a useful perspective on the problem:

1. The number of possible regression structures increases faster than exponentially with dimension.
2. In high-dimensions, nearly all datasets are sparse.
3. In high dimensions, nearly all datasets show multicollinearity (and its nonparametric generalization, concurvity).

Detailed discussion of this topic and its consequences for regression may be found in Hastie and Tibshirani (1990) and in Scott and Wand (1991).

Historically, multivariate statistical analysis sidestepped the COD by imposing strong model assumptions that restricted the potential complexity of the fitted models, thereby allowing sample information to have non-local influence. But now there is growing demand for techniques that make weaker model assumptions and use larger datasets. This has led to the rapid development of a number of new methods, such as additive models (AM), projection pursuit regression (PPR), recursive partitioning regression (RPR), multivariate adaptive regression splines (MARS), alternating conditional expectations (ACE), additivity and variance stabilization (AVAS), locally weighted regression (LOESS), and neural networks. The comparative performance of these methods, however, is poorly understood.

2. Background

Currently, understanding of comparative regression performance is limited to a scattering of theoretical and simulation results. The key results for the most popular regression techniques (defined in Section 3.1) are as follows:

- Donoho and Johnstone (1989) make asymptotic comparisons in terms of the $\mathcal{L}^2$ norm criterion

$$\|\hat{f} - f\| = \int_{\mathbb{R}^p} [\hat{f(\boldsymbol{x})} - f(\boldsymbol{x})]^2 \phi(\boldsymbol{x})\, d\boldsymbol{x}$$

  where p is the dimension of the space and ϕ is the density of the standard normal distribution (i.e., they use a weighted mean integrated squared error (MISE) criterion). Thus the criterion judges an estimator according to the squared distance between its graph and the true graph, with standard normal weighting to downplay disagreement far out in the tails. They find that projection-based regression methods (e.g., PPR, MARS) perform significantly better for radial functions, whereas kernel-based regression (e.g., LOESS) is superior for harmonic functions. Radial functions are constant on hyperspheres centered at $\mathbf{0}$ (e.g., a ripple in a pond), whereas harmonic functions vary periodically on such hyperspheres (e.g., an Elizabethan ruffle).

- Friedman (1991a) reports simulation studies of MARS alone, and related work is described by Barron and Xiao (1991), Owen (1991), Breiman (1991a) and Gu and Wahba (1991). Friedman examines several criteria; the main ones are scaled versions of mean integrated squared error (MISE), predictive-squared error (PSE), and a criterion based on the ratio of a generalized cross-validation (GCV) error estimate to PSE. The most useful conclusions are the following:

  1. When the data are pure noise in 5 and 10 dimensions, for sample sizes of 50, 100, and 200, MARS and AM are roughly comparable and unlikely to find spurious structure.
  2. When the data are generated from the additive function of five variables

  $$Y = 0.1\exp(4X_1) + \frac{4}{1+\exp(-20X_2+10)} + 3X_3 + 2X_4 + X_5$$

     with five additional noise variables and sample sizes of 50, 100, and 200, MARS had a slight but clear tendency to overfit, especially at the smallest sample sizes.

  No simulations were done to compare MARS against other techniques. Breiman (1991a) notes that Friedman's examples (except for the pure noise case) have high signal-to-noise ratios.

- Tibshirani (1988) gives theoretical reasons why AVAS has superior properties to ACE (but notes that consistency and behavior under model misspecification are open questions). He describes a simulation experiment that compares ACE to AVAS in terms of weighted MISE on samples of size 100; the model is $Y = \exp(X_1 + cX_2)$ with X_1, X_2 independent $N(0,1)$ and c taking a range of values to vary the correlation between Y and X_1. He finds that AVAS and ACE are similar, but AVAS performs better than ACE when correlation is low.

- Breiman (1991b) developed nonparametric regression code that describes simulation results for the Π-method, which fits a sum of products. His experiment used the following five functions:

  $$\begin{aligned} Y &= \exp[X_1 \sin(\pi X_2)] \\ Y &= 3\sin(X_1X_2) \\ Y &= \frac{40\exp[8((X_1-.5)^2+(X_2-.5)^2)]}{\exp[8((X_1-.2)^2+(X_2-.7)^2)]\exp[8((X_1-.7)^2+(X_2-.2)^2)]} \\ Y &= \exp[X_1X_2\sin(\pi X_3)] \\ Y &= X_1X_2X_3 \end{aligned}$$

  Evaluation is based on mean squared error averaged over the n data locations. The explanatory variables are independent draws from uniform distributions whose support contains the interesting functional behavior (the support is the region on which the probability density is strictly greater than zero); to each observation, Breiman adds normal noise, choosing the variance so that the signal-to-noise ratio ranges from .9 for the first function to 3.1 for the third. Although the Π-method was not explicitly compared in a simulation study against the methods considered in this paper, Breiman made theoretical and heuristic comparisons, as did discussants, especially Friedman (1991b) and Gu (1991). Their broad conclusions include (1) model parsimony is increasingly valuable in high dimensions, (2) hierarchical models based on sums of piecewise-linear terms are relatively good, and (3) data can be found for which almost any method excels.

- Barron (1991, 1993) shows that in a somewhat narrow sense, the mean integrated squared error of neural network estimates for the class of functions whose Fourier transform $\tilde{g}$ satisfies $\int |\omega||\tilde{g}(\omega)|\, d\omega < c$, for some fixed c, has order $\mathcal{O}(1/m) + \mathcal{O}(mp/n)\ln n$, where m is the number of nodes, p is the dimension, and n is the sample size. This is linear in the dimension, evading the COD; similar results were subsequently obtained by Zhao and Atkeson (1992) for PPR, and it is likely that the result holds for MARS, too. These results may be less applicable than they seem; Barron's class of functions excludes such standard cases as hyperflats, becoming smoother as dimension increases.

- Ripley (1993, 1996) describes simulation studies of neural network procedures, usually in contrast with traditionally statistical methods. Generally, he finds that neural networks perform poorly and are computationally burdensome, more so for regression than classification problems.

- De Veaux, Psichogios, and Ungar (1993) compared MARS and a neural network on two functions, finding that MARS was faster and more accurate in terms of MISE.

- Hastie and Tibshirani (1990) survey many of the new methods. They treat theory and real datasets rather than simulation, but their account of the strategies behind the development of the new methodologies was central to the design of the experiment described in this paper.

These short, often asymptotic, explorations do not provide sufficient understanding for a practitioner to make an informed choice among regression techniques. By contrast, classification is better understood; see Ripley (1994a, 1994b, 1996) and Sutherland et al. (1993) for comparative evaluations of neural networks against more traditional statistical methodologies.

In an effort to fill this gap in the understanding of the comparative performance of regression techniques, a designed simulation experiment was used to contrast ten of the most prominent regression methods. The basis for the comparison is the mean integrated squared error (MISE) of each of the different techniques, assessed across a range of conditions. MISE was chosen because it is the criterion used in most previous studies, because it has an interpretable bias-variance decomposition, and because it reflects essentially all discrepancies between the fitted and true surfaces.

The experiment was run on a DecStation 3000 and an HP Apollo 715/75 over a period of nearly 19 months, using standard code, as described in Section 3. The results from the simulation experiment are summarized in Section 4; a complete set of results is provided in Appendix A. Conclusions drawn from the results are discussed in Section 5.

3. Experimental design

The experiment was a $10 \times 5 \times 3^4$ factorial design whose six factors were regression method, function, dimension, sample size, noise, and model sparseness. The levels (or values) each factor was allowed to take in the experiment are as follows:

Regression method. The ten levels of this factor are linear regression, stepwise linear regression, MARS, AM, projection pursuit regression, ACE, AVAS, recursive partitioning regression (this is very similar to CART), LOESS, and a neural network technique. See Section 3.1 for a description of each of these regression techniques.

Function. This factor determines the functional relationship that is embedded in the data. The five kinds of functions that were examined were hyperflats, multivariate normals with zero correlation, multivariate normals with all correlations .8, two-component mixtures of multivariate normals with zero correlation, and a function proportional to the product of the explanatory variables. The equations for these functions are, respectively, as follows:

$$f(\boldsymbol{X}_i) = \frac{1}{p}\sum_{j=1}^{p} X_{i,j} \qquad \text{(Linear)}$$

$$f(\boldsymbol{X}_i) = (\tfrac{1}{2\pi})^{\frac{p}{2}}(\tfrac{1}{|.25\boldsymbol{I}|})^{\frac{1}{2}}(\exp^{-\frac{1}{2}(\boldsymbol{X}_i)^T(.25\boldsymbol{I})^{-1}(\boldsymbol{X}_i)}) \qquad \text{(Gaussian)}$$

$$f(\boldsymbol{X}_i) = (\tfrac{1}{2\pi})^{\frac{p}{2}}(\tfrac{1}{|.25\boldsymbol{A}|})^{\frac{1}{2}}(\exp^{-\frac{1}{2}(\boldsymbol{X}_i)^T(.25\boldsymbol{A})^{-1}(\boldsymbol{X}_i)}) \qquad \text{(Correlated Gaussian)}$$

$$\begin{aligned} f(\boldsymbol{X}_i) = {} & \tfrac{1}{2}(\tfrac{1}{2\pi})^{\frac{p}{2}}(\tfrac{1}{|.16\boldsymbol{I}|})^{\frac{1}{2}}(\exp^{-\frac{1}{2}(\boldsymbol{X}_i)^T(.16\boldsymbol{I})^{-1}(\boldsymbol{X}_i)})+ \\ & \tfrac{1}{2}(\tfrac{1}{2\pi})^{\frac{p}{2}}(\tfrac{1}{|.16\boldsymbol{I}|})^{\frac{1}{2}}(\exp^{-\frac{1}{2}(\boldsymbol{X}_i-\mathbf{1})^T(.16\boldsymbol{I})^{-1}(\boldsymbol{X}_i-\mathbf{1})}) \end{aligned} \qquad \text{(Mixture)}$$

$$f(\boldsymbol{X}_i) = (\prod_{j=1}^{p} X_{i,j})^{\frac{1}{p}} \qquad \text{(Product)}$$

where p is the dimension, $\mathbf{1}$ is a p-dimensional vector of ones, and $\boldsymbol{A}$ is a covariance matrix with the off-diagonal entries set to .8 and the diagonal entries set to 1. These functions will be referred to, respectively, as Linear, Gaussian, Correlated Gaussian, Mixture, and Product. Figure 1 shows graphical representations of bivariate versions of these functions (i.e., two explanatory variables and one response variable).

Dimension. The three levels of this factor take the dimension of the explanatory variable space to be $p = 2, 6, 12$.

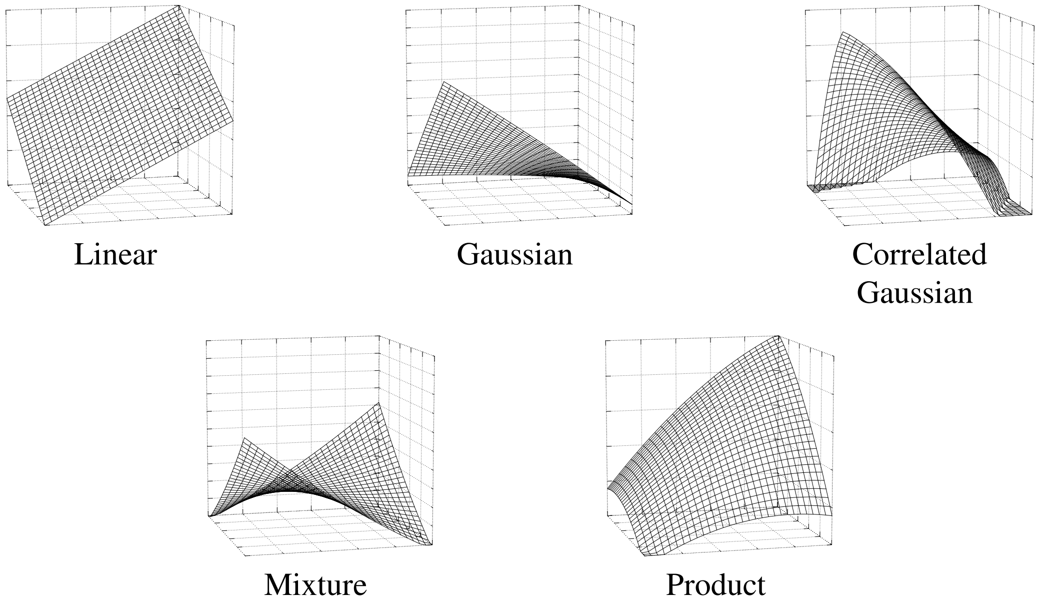

Figure 1: Graphical representations of bivariate versions of the functions used in the simulation experiment (i.e., two explanatory variables and one response variable).

	$p = 2$	$p = 6$	$p = 12$
$k = 4$	16	256	16,384
$k = 10$	40	640	40,960
$k = 25$	100	1,600	102,400

Table 1: Values of n for different values of dimension ($p = 2, 6, 12$) for small ($k = 4$), medium ($k = 10$), and large ($k = 25$) sample sizes.

Sample size. The three levels of this factor take the sample size to be $n = 2^p k$, where p is the dimension and $k = 4, 10, 25$. This scales across dimension, so that the different values of k correspond to small, medium, and large samples, respectively. Table 1 shows the specific values of n for different values of p and k.

Noise. This factor determines the variance in the additive Gaussian error associated with each observation. The standard deviations of the error variance are $\sigma = 0.02, 0.1, 0.5$.

Model sparseness. This factor determines the proportion of explanatory variables that are functionally related to the response variable. The different levels consist of all variables, half of the variables, and none of the variables. When none of the variables are explanatory, then the Constant function $f(\boldsymbol{X}_i) = 1.0$ is used regardless of the level of the Function factor.

Note that not all combinations of this design are realizable. Specifically, when model sparseness is set so that none of the variables pertain to the response variable, then the level of function is irrelevant. Also, LOESS, neural networks, and sometimes AVAS required too much memory or time when the sample size and/or dimension factors were large, resulting in additional missing combinations. These issues will manifest in the results reported in Section 4 and in Appendix A.

For a particular combination of factor levels, the simulation experiment proceeds as follows:

1. Generate a uniform random sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ inside the unit hypercube in $\mathbb{R}^p$.
2. Generate a sample of random errors $\epsilon_1, \ldots, \epsilon_n$, all independent and identically distributed (iid) $N(0, \sigma^2)$.
3. Calculate $Y_i = f(\boldsymbol{X}_i) + \epsilon_i$, where $f : \mathbb{R}^p \to \mathbb{R}$ is the target function determined by appropriate combinations of levels of function, dimension, and model sparseness.
4. Apply the regression method to obtain $\hat{f}$, an estimate of f.
5. Estimate the integrated squared error of $\hat{f}$ over the unit cube (via Monte Carlo, on 10,000 random uniform points). Call this M.
6. Repeat the first five steps 20 times. The average of the 20 resulting M values is an estimate of the MISE; the standard error of this estimate is also calculated for use in subsequent comparisons.

Both the random data sample and the Monte Carlo integration sample were reused for all regression methods. These variance reduction techniques improve the accuracy of contrasts between the methods.

From each combination of factor levels, an estimate of the MISE and its standard error was obtained. Regression methods whose MISE values are significantly lower than competing methods are superior. Note, however, that the COD implies that values of M for large p are less accurate than for smaller p (with all other factors remaining constant); this needs to be taken into account when interpreting the results. The goal is to understand which regression methods are best for which levels (or combinations of levels) of function, dimension, model sparseness, sample size, and noise.

3.1. Regression methods

Multiple linear regression (MLR) and stepwise linear regression (SLR) are standard methods that have been used for decades. These methods are included as performance benchmarks for the simulation study; it is assumed that most readers are already familiar with both techniques. For MLR and SLR, the experiment used the commercial code available from SAS, with the SAS defaults for entering and removing variables in SLR (i.e., using the `SELECTION=STEPWISE` option of `PROC REG`). A study of SLR performance is given by Frank and Friedman (1993).

For the remaining regression techniques, the simplest options and defaults were used consistently: the fitted models were not primed to include polynomial or product terms, MARS and LOESS made first-order fits, and settings did not vary from one level of function to another. (Readers who want details on the smoothers used by the regression methods and the default parameter settings should examine the corresponding documentation.)

The code for the regression techniques used in the experiment (with the exception of MLR and SLR) was assembled into a package called DRAT (Data Regression via Assembled Techniques); the DRAT package is available at `ftp://ftp.cs.cmu.edu/user/bobski/drat/`.

3.1.1. Additive model

The additive model (AM) has been developed by several authors; Buja, Hastie and Tibshirani (1989) describe the background and early development. The simplest AM has the form

$$\mathrm{E}[Y_i] = \theta_0 + \sum_{j=1}^{p} f_j(X_{ij}) \tag{1}$$

where the functions f_j are unknown.

Since the f_j are estimated from the data, one avoids the traditional assumption of linearity in the explanatory variables; however, AM retains the assumption that explanatory variable effects are additive. Thus the response is modeled as the sum of arbitrary smooth univariate functions of the explanatory variables, but not as the sum of multivariate functions of the explanatory variables. One needs a reasonably large sample size to estimate each f_j, but under the model posited in Equation (1), the sample size requirement grows only linearly in p.

The backfitting algorithm, described in Hastie and Tibshirani (1990), is the key procedure used to fit an AM: it is guaranteed to find the best fit between a given model and the data. Operationally, the backfitting algorithm proceeds as follows:

1. At the initialization step, define functions $f_j^{(0)} \equiv 1$ and set $\theta_0 = \bar{Y}$.
2. At the ith iteration, estimate $f_j^{(i+1)}$ by
$$f_j^{(i+1)} = Sm(\boldsymbol{Y} - \theta_0 - \sum_{k \neq j} f_k^i \mid X_{1j}, \ldots, X_{nj})$$
for $j = 1, \ldots, p$.
3. Check whether $|f_j^{(i+1)} - f_j^{(i)}| < \delta$ for all $j = 1, \ldots, p$, where δ is the convergence tolerance. If not, return to step 2; otherwise, use the $f_j^{(i)}$ as the additive functions f_j in the model.

This algorithm requires a smoothing operation (such as kernel smoothing or nearest-neighbor averaging), indicated by $Sm(\cdot \mid \cdot)$. For large classes of smoothing functions, the backfitting algorithm converges to a unique solution.

3.1.2. Projection pursuit regression

The AM considers sums of functions taking arguments in the natural coordinates of the space of explanatory variables. When the underlying function is additive with respect to variables formed by linear combinations of the original explanatory variables, the AM is inappropriate. Projection pursuit regression (PPR) was designed by Friedman and Stuetzle (1981) to handle such cases.

PPR employs the backfitting algorithm and a conventional numerical search routine, such as Gauss-Newton, to fit a model of the form

$$\mathrm{E}[Y_i] = \sum_{k=1}^{r} f_k(\boldsymbol{\alpha}_k^{\mathrm{T}} \boldsymbol{X}_i)$$

where the $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ determine a set of r linear combinations of the explanatory variables. These linear combinations are analogous to those used in principal components analysis (cf. Flury, 1988). A salient difference is that these vectors need not be orthogonal; they are chosen to maximize the predictive accuracy of the model as assessed through generalized cross-validation.

Specifically, the PPR alternately calls two routines. The first conditions upon a set of pseudovariables that are linear combinations of the original variables; these are used in the backfitting algorithm to find an AM that sums functions whose arguments are the pseudovariables (which need not be orthogonal). The second routine conditions upon the estimated AM functions, and searches for linear combinations of the original variables that maximize the fit. Alternating iterative application of these methods converges, very generally, to a unique solution.

PPR can be hard to interpret when $r > 1$. If r is allowed to grow without bound, PPR is consistent. Unlike AM, PPR is invariant to affine transformations of the data; this is appealing when the measurements impose no natural basis.

3.1.3. Recursive partitioning regression

Recursive partitioning regression (RPR) methods have become popular since the advent of the CART (Classification And Regression Trees) methodology, developed by Breiman, Friedman, Olshen and Stone (1984). This project is concerned with regression problems, in which the basic RPR algorithm fits a model of the form

$$\mathrm{E}[Y_i] = \sum_{j=1}^{M} \theta_j I_{R_j}(\boldsymbol{X}_i)$$

where the $R_1, \ldots, R_M$ are rectangular regions that partition $\mathbb{R}^p$, and $I_{R_j}(\boldsymbol{X}_i)$ is an indicator function taking the value 1 if and only if $\boldsymbol{X}_i \in R_j$, and otherwise is zero.

RPR is designed to be very good at finding local low-dimensional structure in functions that show high-dimensional global dependence. It is consistent and has a powerful graphic representation as a decision tree which increases interpretability. However, many elementary functions are awkward for RPR, and it is difficult to discover when the fitted piecewise-constant model approximates a standard smooth function. The RPR code available on Statlib (accessible at `http://lib.stat.cmu.edu/`) was used, rather than CART, to enable inclusion of nonproprietary code in DRAT.

3.1.4. Multivariate adaptive regression splines

Friedman (1991a) describes a method that combines the PPR with RPR, using multivariate adaptive regression splines. This procedure fits a weighted sum of multivariate spline basis functions, also known as tensor-spline basis functions, and the model takes the form

$$\mathrm{E}[Y_i] = \sum_{k=0}^{q} a_k B_k(X_1, \ldots, X_n)$$

where the coefficients a_k are determined in the course of generalized cross-validation fitting. The constant term follows by setting $B_0(X_1, \ldots, X_n) \equiv 1$, and the other multivariate splines are products of univariate spline basis functions:

$$B_k(x_1, \ldots, x_n) = \prod_{s=1}^{r_k} b(x_{i(s,k)} | t_{s,k}) \qquad 1 \leq k \leq r$$

Here the subscript $i(s,k)$ indicates a particular explanatory variable, and the basis spline in that variable has a knot at $t_{s,k}$. The values of q, the $r_1, \ldots, r_q$, the knot sets and the appropriate explanatory variables for inclusion are all determined adaptively from the data.

Multivariate adaptive regression splines (MARS) admits an ANOVA-like decomposition that can be represented in a table and similarly interpreted. MARS is designed to perform well whenever the true function has low local dimension. The procedure automatically accommodates interactions between variables and variable selection.

3.1.5. Alternating conditional expectations

Another extension of AM permits functional transformation of the response variable, as well as the p explanatory variables. The alternating conditional expectations (ACE) algorithm, developed by Breiman and Friedman (1985), fits the model

$$\mathrm{E}[g(Y_i)] = \theta_0 + \sum_{j=1}^{p} f_j(X_{ij}) \tag{2}$$

where all conditions are as given for Equation (1), except g is an unspecified function, scaled to satisfy the technically necessary constraint that var$[g(Y)] = 1$ (otherwise, the zero transformation would be trivially perfect).

Given variables Y_i and $\boldsymbol{X}_i$, one wants g and $f_1, \ldots, f_p$ such that $\mathrm{E}[g(Y_i)|\boldsymbol{X}_i] - \sum_{j=1}^{p} f_j(X_{ij})$ resembles independent error (without loss of generality, the constant term θ_0 can be ignored).

Formally, one solves

$$(\hat{g}, \hat{f}_1, \ldots, \hat{f}_p) = \text{argmin}_{(g, f_1, \ldots, f_p)} \left\{ \sum_{i=1}^{n} \left[g(Y_i) - \sum_{j=1}^{p} f_j(X_{ij}) \right]^2 \right\}$$

where $\hat{g}$ satisfies the unit variance constraint. Operationally, one proceeds as follows:

1. Estimate g by $g^{(0)}$, obtained by applying a smoother to the Y_i values and standardizing the variance. Set $f_j^{(0)} \equiv 1$ for all $j = 1, \ldots, p$.
2. Conditional on $g^{(k-1)}(Y_i)$, apply the backfitting algorithm to find estimates $f_1^{(k)}, \ldots, f_p^{(k)}$.
3. Conditional on the sum of $f_1^{(k)}, \ldots, f_p^{(k)}$, obtain $g^{(k)}$ by applying the backfitting algorithm (this interchanges the role of the explanatory and response variables). Standardize the new function to have unit variance.
4. Test whether $e^{(k)} - e^{(k-1)} = 0$, where

$$e^{(k)} = n^{-1} \sum_{i=1}^{n} \left[g^{(k)}(Y_i) - \sum_{j=1}^{p} f_j^{(k)}(X_{ij}) \right]^2$$

   If it is zero, set $\hat{g} = g^{(k)}$, $\hat{f}_j = f_j^{(k)}$; otherwise, go to step 2.

Steps 2 and 3 calculate smoothed expectations, each conditional upon functions of either the response or the explanatory variables; this alternation gives the method its name.

The ACE analysis finds sets of functions for which the linear correlation of the transformed response variable and the sum of the transformed explanatory variables is maximized. Thus ACE is closer kin to correlation analysis, and the multiple correlation coefficient, than to regression. Since ACE does not aim directly at regression, it has some undesirable features; for example, it treats the response and explanatory variables symmetrically, small changes can lead to radically different solutions (cf. Buja and Kass, 1985), and it does not reproduce model transformations. To increase fairness of comparison, the experiment used an implementation of ACE slightly modified to include a stepwise selection rule mimicking that of SLR.

3.1.6. Additivity and variance stabilization

To overcome some of the potential drawbacks of the ACE methodology, Tibshirani (1988) invented a variation called additivity and variance stabilization (AVAS), which imposes a variance-stabilizing transformation in the ACE backfitting loop for the explanatory variables. AVAS avoids at least two of the deficiencies of ACE in regression applications: it reproduces model transformations and it removes the symmetry between response and explanatory variables.

3.1.7. Locally weighted regression

Cleveland (1979) proposed a locally weighted regression (LOESS) technique. Rather than simply taking a local average, LOESS fits a model of the form $\mathrm{E}[Y] = \boldsymbol{\theta}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x}$ where

$$\hat{\boldsymbol{\theta}}(\boldsymbol{x}) = \mathrm{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^p} \sum_{i=1}^{n} w_i(\boldsymbol{x})(Y_i - \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{X}_i)^2$$

and w_i is a kernel function that weights the influence of the ith observation according to the (oriented) distance of $\boldsymbol{X}_i$ from $\boldsymbol{x}$.

Cleveland and Devlin (1988) generalize LOESS to include polynomial regression, rather than just multiple linear regression, in fitting Y_i to the data, but the improvement seems small. LOESS has good consistency properties, but can be inefficient at discovering some relatively simple structures in data.

3.1.8. Neural networks

Many neural network (NN) techniques exist, but from a statistical regression standpoint (cf. Barron and Barron, 1988), nearly all variants fit models that are weighted sums of sigmoidal functions whose arguments involve linear combinations of the data. A typical feed-forward network uses a model of the form

$$\mathrm{E}[Y] = \beta_0 + \sum_{i=1}^{m} \beta_i f(\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x} + \gamma_{i0})$$

where $f(\cdot)$ is a logistic function and the β_0, γ_{i0}, and $\boldsymbol{\alpha}_i$ are estimated from the data. Formally, this approach is similar to that in PPR. The choice of m determines the number of hidden nodes in the network, and affects the smoothness of the fit; in most cases the user determines this parameter, but for the experiment, m is also estimated from the data.

The particular implementation of the neural net strategy that was employed is Cascor, developed by Fahlman and Lebiere (1990) and used in a similar large-scale simulation comparison of classification methods, described in Sutherland et al. (1993). It was chosen because it learns more rapidly than standard feedforward nets with backpropagation training, because it was used previously in a major comparison, and because it adaptively chooses the number of hidden nodes, thereby making the analysis more automatic. However, Cascor is not necessarily a good indicator of all neural network strategies. Recent work (Doering, Galicki, and Witte, 1997) suggests that in some cases Cascor does not find optimal weights, and thus some alternative implementation of neural net methods may achieve better performance.

Neural nets are widely used, although their performance properties, compared to alternative regression methods, have not been thoroughly studied. Ripley (1993) describes one assessment which finds that neural net methods are not generally competitive. Another difficulty with neural nets is that the resulting model is hard to interpret.

3.2. Other design considerations

The other factors used in the simulation experiment (i.e., function, dimension, sample size, noise, and model sparseness) reflect conventional criteria for performance comparison. Only a few comments seem necessary:

Function. The different levels of function were chosen to reflect the range of structure that practitioners typically encounter in applications. As shown in Figure 1, these include essentially flat surfaces (Linear), surfaces with one or more bumps (Gaussian and Correlated Gaussian), surfaces with structure that is additive in either the natural explanatory variables or linear combinations of them (Mixture), and surfaces that incorporate multiplicative interactions (Product). Also, these choices exercise each of the methods; by their constructions, ACE and AVAS should excel on the Product function, NN and PPR should do well on the Correlated Gaussian function, LOESS and AM should handle the Gaussian function, and RPR and MARS should do well with the Mixture function.

Dimension. The values taken by the dimension factor may seem small. However, previous experience with the impact of the curse of dimensionality (COD) suggests that this is the correct arena for comparing the different methods. In higher dimensions, all methods perform so poorly that comparison is difficult.

Sample size and noise. The values of the sample size and noise factors are typical of previous simulation studies. Qualitatively, these two effects are similar, since a large sample with large noise is informationally comparable to a smaller sample with smaller noise.

Model sparseness. Variable selection is a key concern, both in practice and in theory. Including the model sparseness factor enables users to assess the regression methods with respect to this. But the automatic selection rules may not be comparable across implementations, and thus our results compare default performances, rather than the best that an expert might coax by tuning.

The guiding principle behind the choice of these factor levels is to explore the range of situations that arise in applications, and thereby assist practitioners who have some prior sense of the kinds of regression structures they face.

4. Results

The results of the study consist of the estimated MISE and its variance for $10 \times 5 \times 3^4$ different situations (less a few, since (1) when model sparseness sets all variables to be spurious, the function level becomes irrelevant, and (2) some programs took several days to run, or exhausted the available memory on the computer, with large dimensions and/or sample sizes). These data are complex, and are reported in several ways.

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
MOD	SMALL	ALL	MLR	(27.95)	7.02	(3.00)	.26	(.09)	.01
MOD	SMALL	ALL	SLR	(32.14)	10.65	(3.00)	.26	(.09)	.01
MOD	SMALL	ALL	ACE	53.12	7.36	15.06	1.14	.41	.02
MOD	SMALL	ALL	AM	(27.91)	7.01	(3.00)	.26	(.09)	.01
MOD	SMALL	ALL	MARS	158.23	35.32	18.36	1.88	.30	.02
MOD	SMALL	ALL	RPR	1248.37	307.78	176.32	10.89	61.53	.78
MOD	SMALL	ALL	PPR	55.08	13.14	19.24	8.68	.11	.01
MOD	SMALL	ALL	LOESS	59.50	9.12	9.14	.52	*	*
MOD	SMALL	ALL	AVAS	79.40	18.14	14.73	.95	.38	.02
MOD	SMALL	ALL	NN	124.03	13.05	100.42	2.43	*	*
MOD	SMALL	HALF	MLR	27.95	7.02	3.00	.26	(.09)	.01
MOD	SMALL	HALF	SLR	(23.23)	6.75	(2.39)	.30	(.08)	.01
MOD	SMALL	HALF	ACE	42.42	5.92	14.58	1.12	.40	.02
MOD	SMALL	HALF	AM	27.68	7.00	3.00	.26	(.09)	.01
MOD	SMALL	HALF	MARS	(17.99)	3.24	11.00	1.68	.40	.02
MOD	SMALL	HALF	RPR	1821.76	64.32	239.41	4.50	100.04	2.47
MOD	SMALL	HALF	PPR	35.31	5.61	13.95	8.09	.11	.01
MOD	SMALL	HALF	LOESS	59.50	9.12	9.14	.52	*	*
MOD	SMALL	HALF	AVAS	74.93	14.13	15.02	1.08	.36	.01
MOD	SMALL	HALF	NN	103.08	8.41	94.92	2.65	*	*

Table 2: A subset of the results for the case where function is Linear, sample size is small ($k = 4$), and noise is moderate ($\sigma = 0.1$). Dimension level is indicated by p. All numbers have been multiplied by 10,000. Asterisks denote cases in which no data were available. The parenthesized MISE values were not significantly different from the best regression method under a two-sample t-test with $\alpha = .05$.

Table 2 gives a subset of the results for the case where function is Linear, sample size is small ($k = 4$), and noise is moderate ($\sigma = 0.1$); the MISE values have been multiplied by 10,000 for ease of reading. Much of this information is later expressed in Figure 3, where these MISE values are the data points in each of the six graphs for the case where sample size is small. (The complete set of results is reported in Appendix A; the subset shown in Table 2 appears on page 30.)

A more powerful comparison could be made by taking account of the variance reduction induced by the sample reuse; in that case, most of the MISE values are significantly different, and the method with minimum MISE is strongly favored. But that degree of scrutiny ensures that "*Le mieux est l'ennemi du bien*;" it seems better service to highlight all methods that work well, rather than to emphasize one that is marginally best.

When one has little information about the application, a method that never does badly may be preferred to one that is sometimes the best, but sometimes among the worst. The overall performance of the competing regression methods can be ascertained by taking the ratio of the MISE for a given method to the MISE for the best method, for each combination of factor levels, and then averaging

Method	$p = 2$	$p = 6$	$p = 12$
MLR	2.40	3.21	15.03
SLR	1.77	1.70	8.35
ACE	7.49	30.43	165.70
AM	2.49	4.68	29.58
MARS	1.08	15.86	72.58
RPR	39.38	21.92	1.00
PPR	6.71	14.53	41.20
LOESS	4.85	9.74	*
AVAS	7.77	26.83	97.07
NN	28.46	644.08	*

Table 3: Averages of the MISE ratios for each regression method, broken out by dimension level (indicated by p), for the case where model sparseness sets all variables to be spurious (i.e., the Constant function was used regardless of function level). Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

Method	All Explanatory			Half Explanatory		
	$p = 2$	$p = 6$	$p = 12$	$p = 2$	$p = 6$	$p = 12$
MLR	1.00	1.00	1.00	1.19	1.27	1.08
SLR	1.12	1.10	1.13	1.04	1.00	1.00
ACE	5.07	5.95	5.35	4.20	7.17	5.71
AM	1.00	1.01	1.00	1.22	1.29	2.19
MARS	2.57	5.57	8162.43	1.73	6.92	300.42
RPR	782.77	1053.40	22217.08	2536.65	3049.21	36891.63
PPR	4.11	3.59	1.45	3.13	4.11	1.55
LOESS	1.94	2.98	*	2.29	3.80	*
AVAS	11.25	5.95	4.51	6.67	7.02	4.70
NN	22.11	96.85	*	44.95	157.34	*

Table 4: Averages of the MISE ratios for each regression method, broken out by dimension (indicated by p) and model sparseness levels, for the case where function is Linear. Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

Method	All Explanatory			Half Explanatory		
	$p = 2$	$p = 6$	$p = 12$	$p = 2$	$p = 6$	$p = 12$
MLR	4.10	2.02	7.51	2.66	9.43	4.95
SLR	4.15	1.91	4.40	2.58	9.40	9.00
ACE	2.84	6.73	75.19	2.20	3.20	2.41
AM	4.10	2.02	7.51	2.66	9.42	4.95
MARS	1.91	4.82	34.71	1.64	1.81	4.80
RPR	49.37	12.31	1.01	395.90	42.71	9.34
PPR	2.89	3.56	19.16	1.94	4.01	1.96
LOESS	1.27	2.43	*	2.73	2.38	*
AVAS	3.71	6.25	17.39	2.48	3.09	1.25
NN	8.06	92.55	*	12.77	26.50	*

Table 5: Averages of the MISE ratios for each regression method, broken out by dimension (indicated by p) and model sparseness levels, for the case where function is Gaussian. Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

Method	All Explanatory			Half Explanatory		
	$p = 2$	$p = 6$	$p = 12$	$p = 2$	$p = 6$	$p = 12$
MLR	4.59	2.29	1.51	2.66	7.04	1.75
SLR	4.69	2.30	1.62	2.58	7.02	1.86
ACE	4.41	1.63	1.00	2.20	6.24	1.01
AM	4.59	2.29	1.51	2.66	7.03	1.75
MARS	2.39	2.12	1.44	1.64	2.17	1.36
RPR	15.53	2.68	1.52	395.90	10.63	1.62
PPR	4.55	2.16	1.46	1.94	6.44	1.48
LOESS	1.18	1.78	*	2.73	3.23	*
AVAS	4.05	1.65	1.03	2.48	5.18	1.06
NN	3.56	1.62	*	12.80	3.19	*

Table 6: Averages of the MISE ratios for each regression method, broken out by dimension (indicated by p) and model sparseness levels, for the case where function is Correlated Gaussian. Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

Method	All Explanatory			Half Explanatory		
	$p = 2$	$p = 6$	$p = 12$	$p = 2$	$p = 6$	$p = 12$
MLR	8.94	2.53	5.07	3.99	6.99	3.17
SLR	8.61	2.33	3.13	3.61	6.91	3.14
ACE	10.47	6.39	45.69	4.13	8.21	3.61
AM	8.94	2.52	5.07	3.99	6.98	3.17
MARS	1.70	4.28	22.43	1.14	1.60	2.42
RPR	27.66	12.98	1.00	27.62	11.18	3.12
PPR	3.07	2.88	12.74	4.25	1.96	1.10
LOESS	1.35	1.97	*	3.36	1.37	*
AVAS	10.88	5.99	13.05	4.05	8.51	3.07
NN	7.67	57.26	*	16.53	15.84	*

Table 7: Averages of the MISE ratios for each regression method, broken out by dimension (indicated by p) and model sparseness levels, for the case where function is Mixture. Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

Method	All Explanatory			Half Explanatory		
	$p = 2$	$p = 6$	$p = 12$	$p = 2$	$p = 6$	$p = 12$
MLR	2.69	8.96	16.66	1.19	12.32	23.81
SLR	2.78	9.02	16.67	1.04	12.26	31.36
ACE	2.13	1.74	1.01	4.20	1.70	1.00
AM	2.69	8.96	16.67	1.22	12.31	23.81
MARS	1.81	4.70	57.66	1.73	2.20	15.32
RPR	36.30	31.70	61.11	2536.65	70.23	85.83
PPR	3.27	9.39	16.24	3.13	11.87	22.58
LOESS	1.20	7.14	*	2.29	6.61	*
AVAS	3.81	3.35	5.25	6.67	5.58	7.06
NN	6.40	20.96	*	43.39	17.79	*

Table 8: Averages of the MISE ratios for each regression method, broken out by dimension (indicated by p) and model sparseness levels, for the case where function is Product. Averages were taken over the noise and sample size levels. Asterisks denote cases in which no data were available.

these ratios over the levels of noise and sample size. Table 3 shows this analysis for the case where model sparseness sets all variables to be spurious (i.e., the Constant function was used regardless of function level); the results are broken out by dimension levels. Tables 4, 5, 6, 7, and 8 show this analysis for the cases where function is set, respectively, to Linear, Gaussian, Correlated Gaussian, Mixture, and Product; the results are broken out by dimension and model sparseness levels. Asterisks denote cases in which no data were available.

This tabular information is difficult to apprehend; graphs enable a stronger sense of comparison. To that end, two figures are shown, all for the moderate level of noise ($\sigma = 0.1$), that describe the performance of the methods across dimension and sample size levels. The figures do not include error bars, since (1) this would complicate the images, and (2) the error bars could not take account of the variance reduction attained by sample reuse. The correlation between most methods is very high, and thus visually distinct curves may safely be regarded as statistically distinct.

First consider the case when all explanatory variables are spurious. Here the best predictive rule is $\mathbb{E}[y]$, but many methods overfit. Figure 2 shows the relationship among the methods that attained the smallest MISE. Again, to simplify the graph labels, the MISE has been multiplied by 10,000. Table 9 provides a key for associating regression techniques with the lines on the graphs in Figure 2: the regression techniques are ordered in the table for each graph according to their position when sample size is small. For example, the graph lines in the middle graph of Figure 2, from better to worse MISE (i.e., from smaller to larger MISE values) for the small sample size, are associated with the regression methods are SLR, MLR and AM (i.e., the graph lines for MLR and AM are indistinguishable), LOESS, PPR, MARS, AVAS, ACE, and RPR. Methods not shown have such large MISEs that their inclusion would compress the scale among the good performers, making visual distinctions difficult.

Note that the MISE values typically decrease with dimension, suggesting that the increase in sample size to match the level of dimension is too generous. But cross-dimensional comparisons are not straightforward, since MISE is not a dimensionless quantity (cf. Scott, 1992). Within the dimension levels, RPR does very well when dimension is high ($p = 12$), and badly otherwise; MARS does well when dimension is low ($p = 2$), and degrades for larger values of dimension. SLR, MLR, and AM are consistently competitive. The theoretical minimum that the optimal procedure could be expected to attain is $10^4\sigma^2/2^p k$, where $\sigma = .1$ and p and k are determined by the dimension and sample size levels. For low dimension ($p = 2$), some methods come close to this bound; for larger p, the curse of dimensionality is apparent.

Figure 3 shows six graphs for the case where function is Linear: the lefthand column pertains to the case where model sparseness sets all variables to be explanatory, and the righthand column pertains to the case where model sparseness sets half the variables to be explanatory. As before, all MISE values are multiplied by 10,000. Table 10 provides a key for associating regression techniques with the lines on the graphs in Figure 3: the regression techniques are ordered in the table for each graph according to their position when sample size is small. For example, the graphs lines in the top left of Figure 3, from better to worse MISE (i.e., from smaller to larger MISE values) for the small sample size, are associated with the regression methods MLR and AM (i.e., the graph lines for MLR and AM are indistinguishable), SLR, PPR, AVAS, ACE. As before, methods with very large MISEs are not included, to enhance visual resolution of comparative performance.

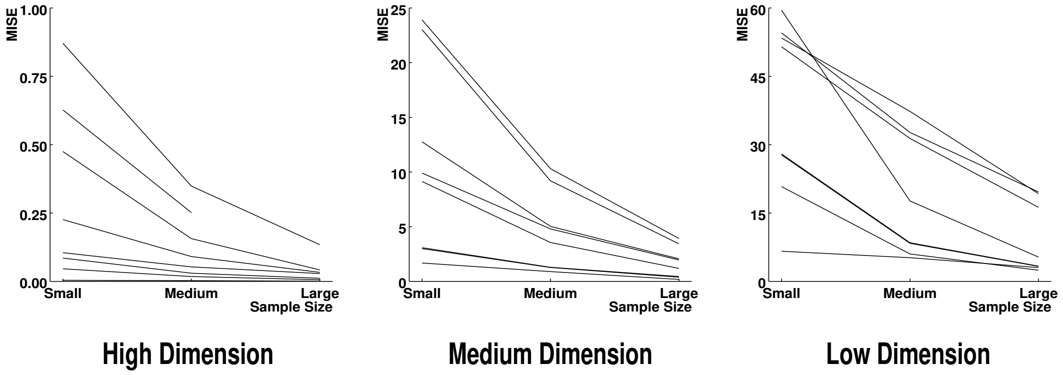

Figure 2: Graphs of MISE values for the case where model sparseness sets all variables to be spurious (i.e., the Constant function was used regardless of function level), broken out by dimension and sample size levels. Each line connects the MISE values for a particular regression method. All MISE values have been multiplied by 10,000. The key to associate regression techniques with graph lines is in Table 9.

ACE	*Larger MISE*	RPR	*Larger MISE*	
AVAS	↑	ACE	↑	LOESS
MARS		AVAS		ACE
PPR		MARS		AVAS
AM		PPR		PPR
MLR		LOESS		MLR,AM
SLR	↓	MLR,AM	↓	SLR
RPR	*Smaller MISE*	SLR	*Smaller MISE*	MARS
High Dimension		**Medium Dimension**		**Low Dimension**

Table 9: The key to associate regression techniques with lines on the graphs in Figure 2. This table is laid out in blocks, similar to the graphs in the figure. Each block in the table lists the regression techniques associated with the lines of the corresponding graph for the case of small sample size.

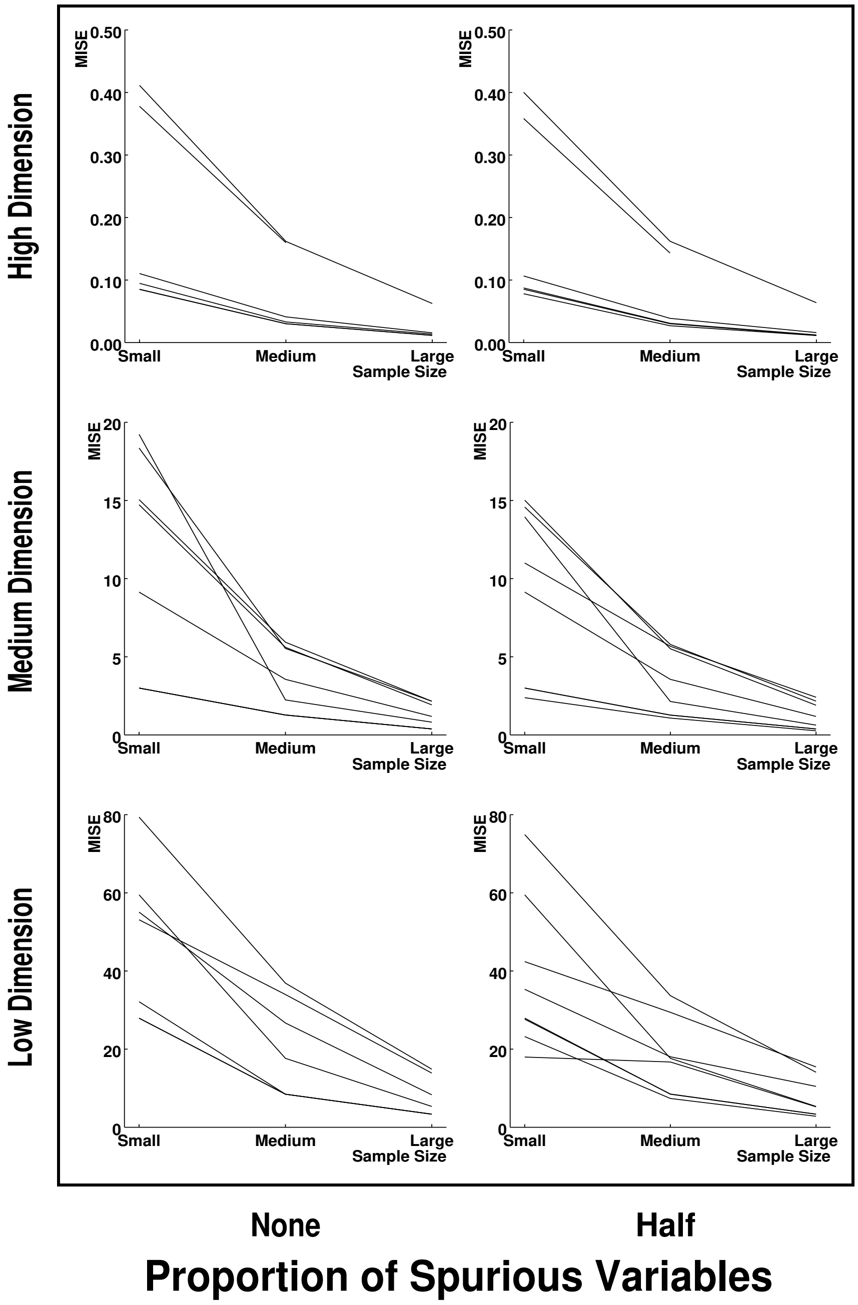

Figure 3: Graphs of MISE values for the cases where model sparseness sets all variables to be explanatory (lefthand column) and half the variables to be explanatory (righthand column), broken out by dimension and sample size levels. Each line connects the MISE values for a particular regression method. All MISE values have been multiplied by 10,000. The key to associate regression techniques with graph lines is in Table 10.

High Dimension	ACE	*Larger MISE*	ACE
	AVAS		AVAS
	PPR		PPR
	SLR		MLR,AM
	MLR,AM	*Smaller MISE*	SLR
		Larger MISE	AVAS
Medium Dimension	PPR		ACE
	MARS		PPR
	ACE		MARS
	AVAS		LOESS
	LOESS		MLR,AM
	MLR,SLR,AM	*Smaller MISE*	SLR
		Larger MISE	AVAS
Low Dimension	AVAS		LOESS
	LOESS		ACE
	PPR		PPR
	ACE		MLR,AM
	SLR		SLR
	MLR,AM	*Smaller MISE*	MARS
	None		**Half**
		Proportion of Spurious Variables	

Table 10: The key to associate regression techniques with lines on the graphs in Figure 3. This table is laid out in blocks, similar to the graphs in the figure. Each block in the table lists the regression techniques associated with the lines of the corresponding graph for the case of small sample size.

Unsurprisingly, MLR excels in the left column, and SLR in the right. The overall shapes of the lines within graphs are consistent with increasing sample size, except for the the odd performance of MARS in the lower right graph. It appears MARS has been tuned, during its design, to handle this paradigm situation when dimension is low ($p = 2$). More generally, the tables indicate that MARS shows marked variability in performance for linear functions; this is reasonable, since its design employs local linear fits. When all knots are removed, one supposes that MARS acts like SLR; but when some knots, by chance, persist, then MARS cannot employ the global information available to SLR, MLR, AM, PPR, ACE, or AVAS.

Within dimension, the graphs for the two levels of model sparseness are quite similar. This suggests that the variable-selection overhead roughly cancels the advantage from fitting a simpler model. Presumably this correspondence would weaken if the numbers of independent variables or the levels of dimension were changed.

Similar figures (not shown here) for the other function levels typically have larger MISE values, and show pronounced differences between the two levels of model sparseness. The insights drawn from these other figures are reflected in the discussion in Section 5.

5. Conclusions

The tables and figures shown or alluded to in Section 4 describe, for different types of functions, the performance of the regression methods under examination. A reverse index is now presented: the performance of each method in different situations is summarized.

- MLR, SLR, and AM perform similarly over all situations considered, and represent broadly safe choices. They are never disastrous, though rarely the best (except for MLR when the function is Linear and all variables are explanatory). For the Constant function, SLR shows less overfit than MLR, which is better than AM; however, it is easy to find functions for which AM would outperform both MLR and SLR. SLR is usually slightly better with spurious variables, but its strategy becomes notably less effective as the number of spurious variables increases, especially for non-linear functions. All three methods have greatest relative difficulty with the Product function, which has substantial curvature.
- On theoretical grounds ACE and AVAS should be similar, but this is not always borne out. ACE is decisively better for the Product function, and AVAS for the Constant function. ACE and AVAS are the best methods for the Product function (as expected—the log transformation produces a linear relationship), but among the worst for the Constant function and for the Mixture function; in other cases, their performance is not remarkable. Both methods are fairly robust to spurious variables.
- Contrary to expectation, MARS does not show well in higher dimensions, especially when all variables are explanatory, and especially for the Linear function. However, for lower levels of dimension, MARS shows adequate performance across the different function levels.

	$p = 2$	$p = 6$	$p = 12$
All Explanatory			
Linear	MLR,SLR	MLR,SLR	MLR,SLR
Gaussian	LOESS,MARS	SLR	RPR
Correlated Gaussian	LOESS	LOESS,NN	ACE,AVAS
Mixture	LOESS	LOESS	RPR
Product	ACE	ACE	ACE
Half Explanatory			
Linear	MLR,SLR	MLR,SLR	MLR,SLR
Gaussian	MARS,PPR	MARS,PPR	MARS,PPR
Correlated Gaussian	MARS	MARS	MARS
Mixture	MARS	MARS	MARS
Product	ACE	ACE	ACE
None Explanatory	MARS	SLR	RPR

Table 11: The most effective regression technique(s) for each combination of dimension (indicated by p), number of explanatory variables, and underlying functional relationship.

MARS is well-calibrated for the Constant function when $p = 2$, but finds spurious structure for larger values, which may account for some of its failures.

- RPR was consistently bad in low levels of dimension, but sometimes stunningly successful in high levels of dimension, especially when all variables were explanatory. Surprisingly, its variable-selection capability was not very successful (MARS's implementation clearly outperforms it). Perhaps the CART program, with its flexible pruning, would surpass RPR, but previous experience with CART suggests such an improvement is dubious. Unsurprisingly, RPR's design made it uncompetitive on the Linear function.

- PPR and NN are theoretically similar methods, but PPR was clearly superior in all cases except for the Correlated Gaussian function. This may reflect peculiarities of the Cascor implementation of neural nets. PPR was often among the best when the function was Gaussian, Correlated Gaussian, or Mixture, but among the worst with the Product function and when all variables were spurious. PPR's variable selection was generally good. In contrast, NN was generally poor, except for the Correlated Gaussian function when $p = 2, 6$ and all variables are explanatory and when $p = 6$ and half the variables are explanatory. The Correlated Gaussian function lends itself to approximation by a small number of sigmoidal functions whose orientations are determined by the data.

- LOESS does well in low levels of dimension with the Gaussian, Correlated Gaussian, and Mixture function. It is not as successful with the other function levels, especially the Constant function. Often, it is not bad in higher levels of dimension, though its relative performance tends to deteriorate.

Additional comparative observations on performance are:

- For the Constant function, MARS is good when $p = 2$, SLR is good when $p = 6$, and RPR is good when $p = 12$. For the Linear function, MLR and SLR are consistently strong. For the Gaussian function, with all variables explanatory, LOESS and MARS are good when $p = 2$, SLR is good when $p = 6$, and RPR is good when $p = 12$; when half of the variables are explanatory, MARS and PPR perform well. For the Correlated Gaussian function, with all variables explanatory, LOESS works well for $p = 2$, LOESS and NN for $p = 6$, and ACE or AVAS for $p = 12$; with half the variables explanatory, MARS is reliably good. For the Mixture function, with all variables explanatory, LOESS works well for $p \leq 6$, and RPR for $p = 12$; with half of the variables explanatory, MARS is consistently good. There is considerable variability for the product function, but ACE is broadly superior. Table 11 summarizes these observations.

- Two kinds of variable-selection strategies were used by the methods: global variable selection, as practiced by SLR, ACE, AVAS, and PPR, and local variable reduction, as practiced by MARS and RPR. Generally, the latter does best in high levels of dimension, but performance depends on the level of function.

- LOESS, NN, and sometimes AVAS proved infeasible in high levels of dimension. The number of local minimizations in LOESS grew exponentially with p. Cascor's demands were high because of the cross-validated selection of the hidden nodes; alternative NN methods fix these a priori, making fewer computational demands, but this is equivalent to imposing strong, though complex, modeling assumptions. Typically, fitting a single high-dimensional dataset with either LOESS or NN took more than two hours. AVAS was faster, but the combination of high dimension and large sample size also required substantial time.

These findings are broadly consistent with those of previous authors, but perhaps more comprehensive.

6. Summary

To restate the most important conclusions, MLR, SLR, and AM are blue-chip methods, that rarely do badly. When the response function is rough, they tend to fit the average, which is often a sensible default. Obviously, a method that in the same circumstances tended to fit the median might have more attractive robustness properties.

NN is unreliable; it can do well, but most often has very large MISE compared to other methods. (Part of this may be that Cascor is not an effective implementation of neural net strategy.) The only cases in which NN was decently competitive were the Correlated Gaussian function with $p = 2, 6$ and all variables explanatory, and $p = 6$ with half of the variables explanatory. However, in practice, NN methods are reported to be very effective. One conjecture is that this is because for many applications, the response function has a sigmoidal shape over the domain of interest, as happened in the example with the Correlated Gaussian function.

MARS is less able in higher dimensions than recent enthusiasm suggests, but it handles a broad range of cases well, and rarely has relatively large MISE. Insofar as MARS combines PPR and RPR strategies, it appears to have protected itself against the worst failures of both, but not attained the best performances of either; such compromise is probably unavoidable. A further hybrid that employs the ACE transformation strategy would potentially be effective.

As a final recommendation for practice, analysts are urged to set aside a portion of the original data, and use each of the reasonable methods to predict the holdouts. The method which does the best job in this test ought to be the method of choice for the unknown situation in hand. This obviates the need for strong prior knowledge about the form of the function, and reduces the inclination to engage in philosophical disputes on the merits of competing strategies for multivariate nonparametric regression.

A. Complete results of the simulation experiment

The following tables contain the complete results from the simulation experiment, broken out by levels of function, noise, sample size (indicated by n), model sparseness, regression method, and dimension (indicated by p). For each combination of factor levels, the MISE and its standard error are reported. All MISE values have been multiplied by 10,000. Asterisks denote cases in which no data were available (the methods took too long to run or had excessive memory demands).

When model sparseness sets all variables to be spurious, the Constant function is used regardless of function level. To eliminate redundancy, the results for the case where all variables are spurious appear as their own set of tables for the Constant function, and these results are omitted from the tables for each function level.

Function: Constant

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	SMALL	NONE	MLR	1.12	.28	.12	.01	.00	.00
LOW	SMALL	NONE	SLR	.83	.28	.07	.01	.00	.00
LOW	SMALL	NONE	ACE	2.18	.35	.96	.06	.04	.00
LOW	SMALL	NONE	AM	1.12	.28	.14	.02	.01	.00
LOW	SMALL	NONE	MARS	.26	.07	.51	.09	.02	.00
LOW	SMALL	NONE	RPR	7.17	1.68	3.27	.36	.00	.00
LOW	SMALL	NONE	PPR	2.03	.27	.47	.05	.01	.00
LOW	SMALL	NONE	LOESS	2.38	.36	.37	.02	*	*
LOW	SMALL	NONE	AVAS	2.20	.28	.90	.06	.02	.00
LOW	SMALL	NONE	NN	4.57	.49	4.21	.12	*	*
LOW	MED	NONE	MLR	.34	.06	.05	.01	.00	.00
LOW	MED	NONE	SLR	.24	.06	.04	.01	.00	.00
LOW	MED	NONE	ACE	1.31	.19	.41	.02	.01	.00
LOW	MED	NONE	AM	.37	.07	.10	.02	.00	.00
LOW	MED	NONE	MARS	.21	.08	.20	.04	.01	.00
LOW	MED	NONE	RPR	7.30	1.62	.58	.20	.00	.00
LOW	MED	NONE	PPR	1.27	.14	.19	.02	.00	.00
LOW	MED	NONE	LOESS	.71	.13	.14	.00	*	*
LOW	MED	NONE	AVAS	1.46	.19	.37	.03	.01	.00
LOW	MED	NONE	NN	4.32	.30	4.20	.08	*	*
LOW	LARGE	NONE	MLR	.13	.02	.02	.00	.00	.00
LOW	LARGE	NONE	SLR	.10	.02	.01	.00	.00	.00
LOW	LARGE	NONE	ACE	.79	.08	.16	.01	.01	.00
LOW	LARGE	NONE	AM	.21	.04	.04	.01	.00	.00
LOW	LARGE	NONE	MARS	.12	.08	.08	.01	.00	.00
LOW	LARGE	NONE	RPR	5.36	1.49	.00	.00	.00	.00
LOW	LARGE	NONE	PPR	.60	.05	.08	.01	.00	.00
LOW	LARGE	NONE	LOESS	.21	.02	.05	.00	*	*
LOW	LARGE	NONE	AVAS	.79	.10	.14	.01	.00	.00
LOW	LARGE	NONE	NN	4.52	.16	4.19	.04	*	*
MOD	SMALL	NONE	MLR	27.95	7.02	3.00	.26	.09	.01
MOD	SMALL	NONE	SLR	20.81	6.93	1.68	.30	.05	.01
MOD	SMALL	NONE	ACE	54.54	8.83	23.91	1.46	.87	.05
MOD	SMALL	NONE	AM	27.75	6.99	3.08	.29	.11	.02
MOD	SMALL	NONE	MARS	6.62	1.74	12.77	2.25	.48	.03
MOD	SMALL	NONE	RPR	178.52	42.19	81.75	8.87	.00	.00
MOD	SMALL	NONE	PPR	51.49	6.54	9.91	1.17	.23	.02
MOD	SMALL	NONE	LOESS	59.50	9.12	9.14	.52	*	*
MOD	SMALL	NONE	AVAS	53.41	6.95	23.02	1.47	.63	.02
MOD	SMALL	NONE	NN	117.30	11.59	102.09	2.33	*	*

Function: Constant (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	NONE	MLR	8.48	1.38	1.28	.13	.03	.00
MOD	MED	NONE	SLR	6.03	1.52	.90	.14	.02	.00
MOD	MED	NONE	ACE	32.70	4.69	10.31	.63	.35	.01
MOD	MED	NONE	AM	8.36	1.43	1.30	.14	.05	.01
MOD	MED	NONE	MARS	5.21	2.05	5.03	1.12	.16	.02
MOD	MED	NONE	RPR	181.89	40.63	14.42	5.12	.00	.00
MOD	MED	NONE	PPR	31.38	3.46	4.80	.54	.09	.01
MOD	MED	NONE	LOESS	17.67	3.21	3.56	.12	*	*
MOD	MED	NONE	AVAS	37.31	4.85	9.19	.65	.25	.01
MOD	MED	NONE	NN	112.31	7.98	105.34	2.31	*	*
MOD	LARGE	NONE	MLR	3.36	.53	.38	.03	.01	.00
MOD	LARGE	NONE	SLR	2.47	.59	.19	.04	.01	.00
MOD	LARGE	NONE	ACE	19.65	2.00	3.93	.24	.13	.01
MOD	LARGE	NONE	AM	3.32	.55	.45	.04	.03	.01
MOD	LARGE	NONE	MARS	3.04	1.99	2.06	.25	.04	.00
MOD	LARGE	NONE	RPR	140.70	37.26	.06	.02	.00	.00
MOD	LARGE	NONE	PPR	16.26	1.46	1.96	.18	.03	.00
MOD	LARGE	NONE	LOESS	5.34	.53	1.19	.06	*	*
MOD	LARGE	NONE	AVAS	19.26	2.52	3.44	.21	.04	.00
MOD	LARGE	NONE	NN	113.65	4.70	104.47	1.25	*	*
HIGH	SMALL	NONE	MLR	698.86	175.55	75.12	6.58	2.13	.25
HIGH	SMALL	NONE	SLR	520.34	173.18	42.09	7.45	1.16	.26
HIGH	SMALL	NONE	ACE	1363.56	220.85	597.75	36.47	21.80	1.17
HIGH	SMALL	NONE	AM	699.52	175.67	75.38	6.58	2.15	.25
HIGH	SMALL	NONE	MARS	165.58	43.50	319.15	56.27	11.88	.79
HIGH	SMALL	NONE	RPR	4582.92	1043.68	2060.34	233.91	.11	.03
HIGH	SMALL	NONE	PPR	1287.09	163.50	262.36	19.06	5.46	.57
HIGH	SMALL	NONE	LOESS	1487.55	227.92	228.57	13.02	*	*
HIGH	SMALL	NONE	AVAS	1365.34	187.20	567.22	38.60	15.53	.54
HIGH	SMALL	NONE	NN	2979.35	318.09	2656.58	63.22	*	*
HIGH	MED	NONE	MLR	212.10	34.48	31.89	3.37	.76	.06
HIGH	MED	NONE	SLR	150.78	38.06	22.44	3.58	.45	.07
HIGH	MED	NONE	ACE	817.43	117.19	257.85	15.66	8.75	.29
HIGH	MED	NONE	AM	212.00	34.66	32.06	3.42	.79	.07
HIGH	MED	NONE	MARS	130.19	51.14	125.72	27.96	3.91	.46
HIGH	MED	NONE	RPR	4518.82	1022.41	360.48	128.12	.07	.03
HIGH	MED	NONE	PPR	787.15	86.64	104.46	10.07	2.11	.16
HIGH	MED	NONE	LOESS	441.64	80.37	89.01	3.11	*	*
HIGH	MED	NONE	AVAS	905.29	115.21	229.53	16.12	6.41	.34
HIGH	MED	NONE	NN	2774.71	203.46	2635.23	55.04	*	*

Function: Constant (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	LARGE	NONE	MLR	84.07	13.20	9.62	.75	.29	.02
HIGH	LARGE	NONE	SLR	61.66	14.73	4.67	1.02	.16	.03
HIGH	LARGE	NONE	ACE	491.16	50.03	98.28	5.85	3.35	.15
HIGH	LARGE	NONE	AM	83.76	13.33	9.70	.73	.33	.03
HIGH	LARGE	NONE	MARS	75.95	49.77	51.58	6.14	1.05	.08
HIGH	LARGE	NONE	RPR	3496.65	933.24	1.50	.52	.03	.01
HIGH	LARGE	NONE	PPR	388.70	33.36	43.99	2.66	.83	.06
HIGH	LARGE	NONE	LOESS	133.52	13.29	29.72	1.49	*	*
HIGH	LARGE	NONE	AVAS	514.20	62.38	83.33	4.78	.68	.05
HIGH	LARGE	NONE	NN	2933.91	104.83	2658.15	27.90	*	*

Function: Linear

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	SMALL	ALL	MLR	1.12	.28	.12	.01	.00	.00
LOW	SMALL	ALL	SLR	1.12	.28	.12	.01	.00	.00
LOW	SMALL	ALL	ACE	11.08	1.95	.68	.05	.02	.00
LOW	SMALL	ALL	AM	1.11	.28	.12	.01	.00	.00
LOW	SMALL	ALL	MARS	2.19	.43	.48	.05	.01	.00
LOW	SMALL	ALL	RPR	602.38	85.66	113.45	6.13	57.80	.52
LOW	SMALL	ALL	PPR	9.66	1.82	.30	.03	.00	.00
LOW	SMALL	ALL	LOESS	2.38	.36	.37	.02	*	*
LOW	SMALL	ALL	AVAS	29.48	4.91	1.11	.39	.02	.00
LOW	SMALL	ALL	NN	21.51	3.15	1.76	.11	*	*
LOW	SMALL	HALF	MLR	1.12	.28	.12	.01	.00	.00
LOW	SMALL	HALF	SLR	.93	.27	.10	.01	.00	.00
LOW	SMALL	HALF	ACE	4.57	1.05	.66	.04	.02	.00
LOW	SMALL	HALF	AM	1.13	.29	.12	.01	.00	.00
LOW	SMALL	HALF	MARS	2.11	.65	.46	.08	.02	.00
LOW	SMALL	HALF	RPR	1651.73	26.01	214.46	3.36	93.15	1.54
LOW	SMALL	HALF	PPR	3.86	.98	.29	.02	.00	.00
LOW	SMALL	HALF	LOESS	2.38	.36	.37	.02	*	*
LOW	SMALL	HALF	AVAS	16.83	5.73	.86	.05	.02	.00
LOW	SMALL	HALF	NN	29.71	2.04	5.19	.45	*	*
LOW	MED	ALL	MLR	.34	.06	.05	.01	.00	.00
LOW	MED	ALL	SLR	.34	.06	.05	.01	.00	.00
LOW	MED	ALL	ACE	3.18	.67	.28	.01	.01	.00
LOW	MED	ALL	AM	.34	.05	.05	.01	.00	.00
LOW	MED	ALL	MARS	.64	.13	.20	.02	13.82	1.57
LOW	MED	ALL	RPR	627.72	71.44	98.72	2.75	58.43	.81
LOW	MED	ALL	PPR	2.37	.56	.12	.01	.00	.00
LOW	MED	ALL	LOESS	.71	.13	.14	.00	*	*
LOW	MED	ALL	AVAS	10.88	2.53	.28	.02	.01	.00
LOW	MED	ALL	NN	9.93	1.22	1.41	.06	*	*
LOW	MED	HALF	MLR	.34	.06	.05	.01	.00	.00
LOW	MED	HALF	SLR	.30	.06	.04	.01	.00	.00
LOW	MED	HALF	ACE	1.75	.33	.26	.02	.01	.00
LOW	MED	HALF	AM	.41	.07	.05	.01	.00	.00
LOW	MED	HALF	MARS	.64	.13	.20	.02	.00	.00
LOW	MED	HALF	RPR	1636.55	12.22	209.65	2.44	90.57	1.90
LOW	MED	HALF	PPR	1.18	.28	.11	.01	.00	.00
LOW	MED	HALF	LOESS	.71	.13	.14	.00	*	*
LOW	MED	HALF	AVAS	2.79	.65	.28	.02	.01	.00
LOW	MED	HALF	NN	24.22	1.80	3.74	.31	*	*

Function: Linear (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	LARGE	ALL	MLR	.13	.02	.02	.00	.00	.00
LOW	LARGE	ALL	SLR	.13	.02	.02	.00	.00	.00
LOW	LARGE	ALL	ACE	.74	.11	.10	.00	.00	.00
LOW	LARGE	ALL	AM	.14	.02	.02	.00	.00	.00
LOW	LARGE	ALL	MARS	.44	.23	.07	.01	27.14	1.51
LOW	LARGE	ALL	RPR	560.75	20.64	93.19	1.77	58.02	.62
LOW	LARGE	ALL	PPR	.52	.09	.04	.00	.00	.00
LOW	LARGE	ALL	LOESS	.21	.02	.05	.00	*	*
LOW	LARGE	ALL	AVAS	2.76	.67	.09	.01	.00	.00
LOW	LARGE	ALL	NN	7.34	.60	1.09	.07	*	*
LOW	LARGE	HALF	MLR	.13	.02	.02	.00	.00	.00
LOW	LARGE	HALF	SLR	.11	.02	.01	.00	.00	.00
LOW	LARGE	HALF	ACE	.52	.06	.09	.00	.00	.00
LOW	LARGE	HALF	AM	.14	.02	.02	.00	.00	.00
LOW	LARGE	HALF	MARS	.16	.04	.10	.01	1.16	1.16
LOW	LARGE	HALF	RPR	1645.60	8.02	207.54	1.21	91.64	1.45
LOW	LARGE	HALF	PPR	.31	.04	.03	.00	.00	.00
LOW	LARGE	HALF	LOESS	.21	.02	.05	.00	*	*
LOW	LARGE	HALF	AVAS	.92	.37	.09	.01	.00	.00
LOW	LARGE	HALF	NN	20.97	1.41	3.00	.28	*	*
MOD	SMALL	ALL	MLR	27.95	7.02	3.00	.26	.09	.01
MOD	SMALL	ALL	SLR	32.14	10.65	3.00	.26	.09	.01
MOD	SMALL	ALL	ACE	53.12	7.36	15.06	1.14	.41	.02
MOD	SMALL	ALL	AM	27.91	7.01	3.00	.26	.09	.01
MOD	SMALL	ALL	MARS	158.23	35.32	18.36	1.88	.30	.02
MOD	SMALL	ALL	RPR	1248.37	307.78	176.32	10.89	61.53	.78
MOD	SMALL	ALL	PPR	55.08	13.14	19.24	8.68	.11	.01
MOD	SMALL	ALL	LOESS	59.50	9.12	9.14	.52	*	*
MOD	SMALL	ALL	AVAS	79.40	18.14	14.73	.95	.38	.02
MOD	SMALL	ALL	NN	124.03	13.05	100.42	2.43	*	*
MOD	SMALL	HALF	MLR	27.95	7.02	3.00	.26	.09	.01
MOD	SMALL	HALF	SLR	23.23	6.75	2.39	.30	.08	.01
MOD	SMALL	HALF	ACE	42.42	5.92	14.58	1.12	.40	.02
MOD	SMALL	HALF	AM	27.68	7.00	3.00	.26	.09	.01
MOD	SMALL	HALF	MARS	17.99	3.24	11.00	1.68	.40	.02
MOD	SMALL	HALF	RPR	1821.76	64.32	239.41	4.50	100.04	2.47
MOD	SMALL	HALF	PPR	35.31	5.61	13.95	8.09	.11	.01
MOD	SMALL	HALF	LOESS	59.50	9.12	9.14	.52	*	*
MOD	SMALL	HALF	AVAS	74.93	14.13	15.02	1.08	.36	.01
MOD	SMALL	HALF	NN	103.08	8.41	94.92	2.65	*	*

Function: Linear (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	ALL	MLR	8.48	1.38	1.28	.13	.03	.00
MOD	MED	ALL	SLR	8.48	1.38	1.28	.13	.03	.00
MOD	MED	ALL	ACE	33.98	3.92	5.94	.23	.16	.01
MOD	MED	ALL	AM	8.47	1.37	1.28	.13	.03	.00
MOD	MED	ALL	MARS	19.60	4.13	5.54	.52	14.45	1.47
MOD	MED	ALL	RPR	869.96	103.93	124.07	4.45	62.80	.75
MOD	MED	ALL	PPR	26.71	4.34	2.24	.26	.04	.00
MOD	MED	ALL	LOESS	17.67	3.21	3.56	.12	*	*
MOD	MED	ALL	AVAS	36.89	4.44	5.61	.29	.16	.01
MOD	MED	ALL	NN	99.61	8.50	99.59	2.03	*	*
MOD	MED	HALF	MLR	8.48	1.38	1.28	.13	.03	.00
MOD	MED	HALF	SLR	7.38	1.44	1.08	.13	.03	.00
MOD	MED	HALF	ACE	29.42	4.45	5.79	.22	.16	.01
MOD	MED	HALF	AM	8.52	1.42	1.27	.13	.03	.00
MOD	MED	HALF	MARS	16.73	3.84	5.67	.53	.10	.01
MOD	MED	HALF	RPR	1708.05	41.80	219.15	3.15	97.91	2.56
MOD	MED	HALF	PPR	18.07	3.28	2.14	.22	.04	.00
MOD	MED	HALF	LOESS	17.67	3.21	3.56	.12	*	*
MOD	MED	HALF	AVAS	33.71	4.57	5.51	.32	.14	.01
MOD	MED	HALF	NN	97.55	11.51	94.53	1.76	*	*
MOD	LARGE	ALL	MLR	3.36	.53	.38	.03	.01	.00
MOD	LARGE	ALL	SLR	3.36	.53	.38	.03	.01	.00
MOD	LARGE	ALL	ACE	13.88	1.15	2.16	.09	.06	.00
MOD	LARGE	ALL	AM	3.37	.53	.39	.03	.01	.00
MOD	LARGE	ALL	MARS	4.35	.67	2.16	.20	26.11	1.37
MOD	LARGE	ALL	RPR	828.79	60.61	106.04	3.90	61.95	.57
MOD	LARGE	ALL	PPR	8.31	.90	.82	.07	.02	.00
MOD	LARGE	ALL	LOESS	5.34	.53	1.19	.06	*	*
MOD	LARGE	ALL	AVAS	14.87	1.57	1.92	.11	.03	.00
MOD	LARGE	ALL	NN	93.48	4.39	99.50	1.05	*	*
MOD	LARGE	HALF	MLR	3.36	.53	.38	.03	.01	.00
MOD	LARGE	HALF	SLR	2.84	.56	.27	.04	.01	.00
MOD	LARGE	HALF	ACE	15.45	1.64	2.17	.08	.06	.00
MOD	LARGE	HALF	AM	3.38	.52	.38	.03	.01	.00
MOD	LARGE	HALF	MARS	5.26	1.54	2.41	.34	1.19	1.17
MOD	LARGE	HALF	RPR	1671.40	30.36	211.66	2.36	101.42	2.14
MOD	LARGE	HALF	PPR	10.49	1.78	.63	.05	.02	.00
MOD	LARGE	HALF	LOESS	5.34	.53	1.19	.06	*	*
MOD	LARGE	HALF	AVAS	14.09	1.65	1.91	.10	.03	.00
MOD	LARGE	HALF	NN	86.01	5.20	92.99	1.07	*	*

Function: Linear (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	SMALL	ALL	MLR	698.86	175.55	75.12	6.58	2.13	.25
HIGH	SMALL	ALL	SLR	853.34	167.67	106.59	5.29	2.37	.21
HIGH	SMALL	ALL	ACE	1149.79	161.82	517.59	33.79	11.98	.68
HIGH	SMALL	ALL	AM	700.48	176.17	75.06	6.58	2.13	.25
HIGH	SMALL	ALL	MARS	701.76	76.61	458.09	81.63	14.58	1.06
HIGH	SMALL	ALL	RPR	7167.58	1568.85	2020.65	161.21	69.49	.18
HIGH	SMALL	ALL	PPR	1437.90	244.03	337.96	24.76	3.64	.39
HIGH	SMALL	ALL	LOESS	1487.55	227.92	228.57	13.02	*	*
HIGH	SMALL	ALL	AVAS	1269.11	184.66	488.20	34.53	10.27	.52
HIGH	SMALL	ALL	NN	2874.32	273.84	2533.64	55.51	*	*
HIGH	SMALL	HALF	MLR	698.86	175.55	75.12	6.58	2.13	.25
HIGH	SMALL	HALF	SLR	659.82	169.01	65.97	10.33	1.95	.22
HIGH	SMALL	HALF	ACE	1306.61	209.63	534.01	38.54	10.67	.52
HIGH	SMALL	HALF	AM	700.57	176.14	75.03	6.56	2.14	.25
HIGH	SMALL	HALF	MARS	980.79	99.23	545.75	154.81	11.99	.47
HIGH	SMALL	HALF	RPR	6120.10	1335.74	2477.94	304.14	137.36	1.87
HIGH	SMALL	HALF	PPR	1285.05	255.48	442.21	32.85	3.04	.46
HIGH	SMALL	HALF	LOESS	1487.55	227.92	228.57	13.02	*	*
HIGH	SMALL	HALF	AVAS	1243.75	181.84	475.59	37.67	9.62	.36
HIGH	SMALL	HALF	NN	2806.13	347.00	2636.77	69.95	*	*
HIGH	MED	ALL	MLR	212.10	34.48	31.89	3.37	.76	.06
HIGH	MED	ALL	SLR	316.35	40.70	43.89	5.53	.83	.08
HIGH	MED	ALL	ACE	870.34	121.43	198.99	10.42	4.26	.18
HIGH	MED	ALL	AM	211.72	34.35	31.88	3.37	.75	.06
HIGH	MED	ALL	MARS	533.57	40.58	187.33	17.19	143.64	127.11
HIGH	MED	ALL	RPR	5910.82	974.39	720.56	167.88	69.47	.19
HIGH	MED	ALL	PPR	880.51	125.50	184.61	15.33	1.29	.13
HIGH	MED	ALL	LOESS	441.64	80.37	89.01	3.11	*	*
HIGH	MED	ALL	AVAS	870.12	122.10	178.82	11.76	4.05	.22
HIGH	MED	ALL	NN	2919.89	185.31	2585.78	33.27	*	*
HIGH	MED	HALF	MLR	212.10	34.48	31.89	3.37	.76	.06
HIGH	MED	HALF	SLR	231.02	59.73	26.97	3.13	.67	.08
HIGH	MED	HALF	ACE	816.30	136.00	180.50	6.83	4.25	.21
HIGH	MED	HALF	AM	211.16	34.40	31.89	3.37	.75	.06
HIGH	MED	HALF	MARS	406.73	74.99	146.96	20.75	2.63	.17
HIGH	MED	HALF	RPR	6062.59	941.35	767.26	130.32	138.27	.36
HIGH	MED	HALF	PPR	674.30	131.88	218.03	27.72	1.23	.10
HIGH	MED	HALF	LOESS	441.64	80.37	89.01	3.11	*	*
HIGH	MED	HALF	AVAS	847.22	128.02	160.28	11.18	3.71	.21
HIGH	MED	HALF	NN	2814.44	165.28	2634.68	47.74	*	*

Function: Linear (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	LARGE	ALL	MLR	84.07	13.20	9.62	.75	.29	.02
HIGH	LARGE	ALL	SLR	104.18	18.84	10.10	1.08	.34	.03
HIGH	LARGE	ALL	ACE	416.96	46.59	72.52	3.60	1.67	.06
HIGH	LARGE	ALL	AM	84.05	13.21	9.60	.74	.29	.02
HIGH	LARGE	ALL	MARS	270.37	35.99	92.97	9.55	25.72	1.20
HIGH	LARGE	ALL	RPR	4472.38	777.70	169.59	11.66	69.43	.18
HIGH	LARGE	ALL	PPR	310.89	39.82	41.44	7.12	.42	.03
HIGH	LARGE	ALL	LOESS	133.52	13.29	29.72	1.49	*	*
HIGH	LARGE	ALL	AVAS	385.59	40.45	59.32	4.16	.85	.05
HIGH	LARGE	ALL	NN	2849.57	94.08	2622.47	21.95	*	*
HIGH	LARGE	HALF	MLR	84.07	13.20	9.62	.75	.29	.02
HIGH	LARGE	HALF	SLR	70.99	13.89	6.77	.90	.28	.03
HIGH	LARGE	HALF	ACE	335.56	43.33	62.65	2.91	1.62	.08
HIGH	LARGE	HALF	AM	84.00	13.21	9.61	.74	.29	.02
HIGH	LARGE	HALF	MARS	81.51	17.81	74.02	7.58	.52	.05
HIGH	LARGE	HALF	RPR	4230.49	550.92	268.37	15.27	138.23	.35
HIGH	LARGE	HALF	PPR	286.81	24.59	25.29	2.35	.42	.03
HIGH	LARGE	HALF	LOESS	133.52	13.29	29.72	1.49	*	*
HIGH	LARGE	HALF	AVAS	341.80	34.96	52.74	3.40	.79	.05
HIGH	LARGE	HALF	NN	2771.59	119.81	2641.75	29.16	*	*

Function: Gaussian

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
LOW	SMALL	ALL	MLR	45.54	2.85	1.73	.02	.01	.00
LOW	SMALL	ALL	SLR	45.54	2.85	1.73	.02	.00	.00
LOW	SMALL	ALL	ACE	29.18	4.93	1.75	.06	.03	.00
LOW	SMALL	ALL	AM	45.52	2.84	1.72	.02	.01	.00
LOW	SMALL	ALL	MARS	27.65	5.31	2.30	.14	.02	.00
LOW	SMALL	ALL	RPR	368.20	25.80	5.54	.39	.00	.00
LOW	SMALL	ALL	PPR	25.57	2.90	2.29	.32	.01	.00
LOW	SMALL	ALL	LOESS	9.64	1.17	.76	.05	*	*
LOW	SMALL	ALL	AVAS	39.62	9.01	1.94	.09	.01	.00
LOW	SMALL	ALL	NN	34.29	5.35	5.24	.16	*	*
LOW	SMALL	HALF	MLR	6.83	.35	22.15	.10	1.56	.02
LOW	SMALL	HALF	SLR	6.57	.35	22.06	.11	3.89	.93
LOW	SMALL	HALF	ACE	3.42	.46	5.07	.33	.25	.01
LOW	SMALL	HALF	AM	6.85	.36	22.14	.10	1.56	.02
LOW	SMALL	HALF	MARS	5.03	.32	1.47	.07	1.00	.05
LOW	SMALL	HALF	RPR	1060.87	15.00	88.00	1.44	2.82	.07
LOW	SMALL	HALF	PPR	3.11	.52	7.26	.08	.34	.00
LOW	SMALL	HALF	LOESS	6.70	.79	4.82	.15	*	*
LOW	SMALL	HALF	AVAS	9.82	3.48	5.62	.47	.16	.01
LOW	SMALL	HALF	NN	18.53	1.75	8.72	.50	*	*
LOW	MED	ALL	MLR	34.51	.64	1.66	.02	.00	.00
LOW	MED	ALL	SLR	34.51	.64	1.66	.02	.00	.00
LOW	MED	ALL	ACE	9.60	1.48	.98	.05	.01	.00
LOW	MED	ALL	AM	34.52	.64	1.65	.02	.00	.00
LOW	MED	ALL	MARS	3.73	.34	1.58	.06	.01	.00
LOW	MED	ALL	RPR	378.44	11.40	3.24	.12	.00	.00
LOW	MED	ALL	PPR	14.78	.65	.50	.01	.00	.00
LOW	MED	ALL	LOESS	4.06	.41	.52	.02	*	*
LOW	MED	ALL	AVAS	21.99	2.71	1.14	.06	.00	.00
LOW	MED	ALL	NN	11.52	1.31	5.00	.10	*	*
LOW	MED	HALF	MLR	5.78	.18	21.70	.10	1.56	.02
LOW	MED	HALF	SLR	5.68	.18	21.65	.10	2.53	.66
LOW	MED	HALF	ACE	1.58	.23	2.32	.10	.22	.00
LOW	MED	HALF	AM	5.78	.18	21.68	.09	1.56	.02
LOW	MED	HALF	MARS	1.97	.25	.91	.04	1.45	.03
LOW	MED	HALF	RPR	1046.53	10.33	87.21	.69	2.84	.06
LOW	MED	HALF	PPR	1.24	.17	6.53	.05	.33	.00
LOW	MED	HALF	LOESS	4.46	.24	3.80	.06	*	*
LOW	MED	HALF	AVAS	1.71	.23	2.68	.15	.13	.00
LOW	MED	HALF	NN	13.19	1.47	3.87	.35	*	*

Function: Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	LARGE	ALL	MLR	31.61	.21	1.60	.02	.00	.00
LOW	LARGE	ALL	SLR	31.61	.21	1.60	.02	.00	.00
LOW	LARGE	ALL	ACE	5.45	.99	.53	.03	.01	.00
LOW	LARGE	ALL	AM	31.60	.21	1.61	.02	.00	.00
LOW	LARGE	ALL	MARS	2.69	.12	1.35	.03	.00	.00
LOW	LARGE	ALL	RPR	401.10	13.03	2.93	.09	.00	.00
LOW	LARGE	ALL	PPR	12.68	.23	.39	.01	.00	.00
LOW	LARGE	ALL	LOESS	2.91	.13	.39	.01	*	*
LOW	LARGE	ALL	AVAS	15.42	2.78	.60	.03	*	*
LOW	LARGE	ALL	NN	8.49	.51	4.48	.06	*	*
LOW	LARGE	HALF	MLR	5.38	.10	21.50	.09	1.56	.02
LOW	LARGE	HALF	SLR	5.33	.10	21.49	.09	2.55	.68
LOW	LARGE	HALF	ACE	.59	.06	1.29	.05	.21	.00
LOW	LARGE	HALF	AM	5.38	.10	21.51	.10	1.56	.02
LOW	LARGE	HALF	MARS	.72	.08	.69	.03	1.87	.02
LOW	LARGE	HALF	RPR	1050.93	4.86	87.90	.75	2.78	.06
LOW	LARGE	HALF	PPR	.52	.05	6.34	.03	1.31	.45
LOW	LARGE	HALF	LOESS	3.91	.16	3.51	.05	*	*
LOW	LARGE	HALF	AVAS	.70	.09	1.63	.07	*	*
LOW	LARGE	HALF	NN	10.49	.95	1.72	.10	*	*
MOD	SMALL	ALL	MLR	76.27	10.85	4.55	.27	.09	.01
MOD	SMALL	ALL	SLR	93.10	15.16	4.45	.23	.05	.01
MOD	SMALL	ALL	ACE	88.49	10.47	25.75	1.62	.81	.03
MOD	SMALL	ALL	AM	76.30	10.86	4.42	.27	.09	.01
MOD	SMALL	ALL	MARS	234.11	20.54	13.75	2.08	.45	.03
MOD	SMALL	ALL	RPR	522.56	59.84	101.45	16.08	.01	.00
MOD	SMALL	ALL	PPR	73.44	11.71	12.04	.76	.22	.02
MOD	SMALL	ALL	LOESS	63.58	9.91	9.50	.60	*	*
MOD	SMALL	ALL	AVAS	91.04	10.89	23.52	1.76	.21	.01
MOD	SMALL	ALL	NN	140.16	15.37	104.60	2.35	*	*
MOD	SMALL	HALF	MLR	31.92	5.37	25.02	.35	1.64	.02
MOD	SMALL	HALF	SLR	27.75	4.95	24.23	.36	3.45	.77
MOD	SMALL	HALF	ACE	41.55	5.82	28.20	1.89	1.18	.03
MOD	SMALL	HALF	AM	31.90	5.37	24.83	.32	1.64	.02
MOD	SMALL	HALF	MARS	69.51	27.58	23.01	3.09	1.10	.07
MOD	SMALL	HALF	RPR	1154.00	60.09	149.73	6.72	3.50	.04
MOD	SMALL	HALF	PPR	45.70	7.57	30.25	5.86	.47	.01
MOD	SMALL	HALF	LOESS	59.41	8.10	13.70	.82	*	*
MOD	SMALL	HALF	AVAS	50.06	6.95	26.80	1.54	.68	.03
MOD	SMALL	HALF	NN	109.50	13.47	109.00	3.23	*	*

Function: Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	ALL	MLR	43.17	2.13	2.88	.14	.03	.00
MOD	MED	ALL	SLR	43.17	2.13	3.22	.12	.02	.00
MOD	MED	ALL	ACE	57.24	6.72	11.36	.70	.31	.02
MOD	MED	ALL	AM	43.16	2.12	2.78	.12	.03	.00
MOD	MED	ALL	MARS	46.88	5.87	13.34	4.95	.16	.02
MOD	MED	ALL	RPR	599.96	100.59	15.93	3.01	.01	.00
MOD	MED	ALL	PPR	44.22	6.96	6.69	.62	.08	.01
MOD	MED	ALL	LOESS	20.76	3.29	3.93	.16	*	*
MOD	MED	ALL	AVAS	59.59	7.17	9.77	.61	.07	.00
MOD	MED	ALL	NN	124.75	9.66	106.70	2.20	*	*
MOD	MED	HALF	MLR	14.33	1.65	22.93	.19	1.59	.02
MOD	MED	HALF	SLR	13.28	1.68	22.72	.20	2.86	.69
MOD	MED	HALF	ACE	30.05	4.08	15.00	.45	.62	.02
MOD	MED	HALF	AM	14.32	1.65	22.75	.15	1.59	.02
MOD	MED	HALF	MARS	16.38	2.91	7.42	.71	1.10	.07
MOD	MED	HALF	RPR	1189.13	46.66	101.11	2.02	3.50	.04
MOD	MED	HALF	PPR	19.41	3.19	11.04	1.57	.38	.01
MOD	MED	HALF	LOESS	20.93	3.06	7.12	.21	*	*
MOD	MED	HALF	AVAS	28.94	4.26	13.51	.55	.40	.02
MOD	MED	HALF	NN	94.25	8.45	106.44	2.09	*	*
MOD	LARGE	ALL	MLR	34.85	.61	1.97	.03	.01	.00
MOD	LARGE	ALL	SLR	34.85	.61	2.22	.06	.01	.00
MOD	LARGE	ALL	ACE	28.25	2.36	4.64	.27	.14	.01
MOD	LARGE	ALL	AM	34.84	.61	2.01	.05	.01	.00
MOD	LARGE	ALL	MARS	9.01	.84	4.18	.23	.04	.00
MOD	LARGE	ALL	RPR	450.43	21.62	3.85	.08	.00	.00
MOD	LARGE	ALL	PPR	21.93	1.31	2.55	.27	.03	.00
MOD	LARGE	ALL	LOESS	7.53	.61	1.51	.06	*	*
MOD	LARGE	ALL	AVAS	29.83	3.09	4.15	.20	*	*
MOD	LARGE	ALL	NN	101.40	3.85	107.99	1.35	*	*
MOD	LARGE	HALF	MLR	8.72	.56	21.84	.10	1.57	.02
MOD	LARGE	HALF	SLR	8.37	.59	21.74	.09	2.43	.60
MOD	LARGE	HALF	ACE	13.30	1.28	8.11	.47	.39	.01
MOD	LARGE	HALF	AM	8.68	.55	21.89	.11	1.57	.02
MOD	LARGE	HALF	MARS	7.49	.89	3.31	.27	1.31	.05
MOD	LARGE	HALF	RPR	1073.66	18.64	91.82	1.15	3.50	.04
MOD	LARGE	HALF	PPR	9.41	.81	7.23	.10	.35	.00
MOD	LARGE	HALF	LOESS	9.09	.62	4.65	.13	*	*
MOD	LARGE	HALF	AVAS	12.32	1.34	6.91	.35	*	*
MOD	LARGE	HALF	NN	89.33	4.48	103.93	1.14	*	*

Function: Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	SMALL	ALL	MLR	766.65	188.43	76.32	6.59	2.13	.25
HIGH	SMALL	ALL	SLR	647.82	169.75	46.23	7.66	1.15	.26
HIGH	SMALL	ALL	ACE	1551.57	232.27	614.53	33.03	20.23	.76
HIGH	SMALL	ALL	AM	766.34	188.18	73.15	6.57	2.13	.25
HIGH	SMALL	ALL	MARS	622.15	167.11	285.07	81.80	10.51	.90
HIGH	SMALL	ALL	RPR	4158.61	849.36	2220.23	221.85	.11	.03
HIGH	SMALL	ALL	PPR	1303.87	200.92	254.14	26.07	5.11	.43
HIGH	SMALL	ALL	LOESS	1475.68	227.83	228.76	13.36	*	*
HIGH	SMALL	ALL	AVAS	1408.40	191.68	571.39	38.97	5.11	.21
HIGH	SMALL	ALL	NN	2860.39	311.03	2589.91	69.50	*	*
HIGH	SMALL	HALF	MLR	694.06	165.60	97.05	6.79	3.68	.25
HIGH	SMALL	HALF	SLR	562.98	170.48	103.87	8.89	3.74	.22
HIGH	SMALL	HALF	ACE	1175.59	172.64	592.49	38.10	20.58	.78
HIGH	SMALL	HALF	AM	693.05	165.24	93.73	6.64	3.67	.25
HIGH	SMALL	HALF	MARS	798.05	81.37	327.89	41.84	15.04	1.14
HIGH	SMALL	HALF	RPR	5265.97	815.08	2447.36	402.99	3.60	.05
HIGH	SMALL	HALF	PPR	1209.30	177.62	335.56	32.97	6.58	.43
HIGH	SMALL	HALF	LOESS	1465.39	220.76	233.65	14.05	*	*
HIGH	SMALL	HALF	AVAS	1211.21	181.22	554.94	40.69	6.33	.21
HIGH	SMALL	HALF	NN	2804.79	235.82	2621.24	61.22	*	*
HIGH	MED	ALL	MLR	249.39	35.10	33.47	3.39	.76	.06
HIGH	MED	ALL	SLR	301.51	33.65	26.39	3.55	.45	.07
HIGH	MED	ALL	ACE	867.83	122.72	266.16	15.42	8.11	.43
HIGH	MED	ALL	AM	249.56	35.23	31.25	3.13	.76	.06
HIGH	MED	ALL	MARS	443.10	61.51	142.83	30.28	3.93	.47
HIGH	MED	ALL	RPR	5102.35	1289.80	242.47	50.55	.08	.03
HIGH	MED	ALL	PPR	761.53	106.34	113.54	9.12	2.04	.14
HIGH	MED	ALL	LOESS	443.40	79.12	89.36	3.30	*	*
HIGH	MED	ALL	AVAS	908.21	139.91	229.17	15.91	1.72	.09
HIGH	MED	ALL	NN	2977.07	235.48	2635.09	50.56	*	*
HIGH	MED	HALF	MLR	219.98	35.41	53.56	3.44	2.31	.07
HIGH	MED	HALF	SLR	254.58	53.36	53.90	4.02	2.37	.09
HIGH	MED	HALF	ACE	778.16	116.41	241.67	15.31	8.05	.40
HIGH	MED	HALF	AM	219.35	35.37	50.89	3.02	2.31	.07
HIGH	MED	HALF	MARS	439.34	69.47	191.43	18.12	5.76	.34
HIGH	MED	HALF	RPR	5845.81	1013.19	914.35	227.30	3.57	.05
HIGH	MED	HALF	PPR	683.83	103.67	185.04	9.44	2.32	.12
HIGH	MED	HALF	LOESS	442.44	79.10	92.09	3.15	*	*
HIGH	MED	HALF	AVAS	815.80	115.49	203.72	12.81	2.98	.10
HIGH	MED	HALF	NN	2848.33	229.72	2635.57	47.00	*	*

Function: Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
HIGH	LARGE	ALL	MLR	115.59	12.58	11.16	.74	.29	.02
HIGH	LARGE	ALL	SLR	140.94	14.34	9.05	.97	.16	.03
HIGH	LARGE	ALL	ACE	466.63	51.21	93.81	6.36	3.81	.07
HIGH	LARGE	ALL	AM	115.59	12.60	11.98	1.06	.29	.02
HIGH	LARGE	ALL	MARS	273.54	24.77	57.97	11.17	1.05	.08
HIGH	LARGE	ALL	RPR	3625.81	565.99	5.05	.53	.03	.01
HIGH	LARGE	ALL	PPR	381.91	50.82	52.60	3.80	.90	.07
HIGH	LARGE	ALL	LOESS	133.17	13.02	29.93	1.47	*	*
HIGH	LARGE	ALL	AVAS	407.17	46.38	87.12	5.50	*	*
HIGH	LARGE	ALL	NN	2831.63	127.99	2602.72	22.35	*	*
HIGH	LARGE	HALF	MLR	89.99	13.18	30.93	.78	1.85	.04
HIGH	LARGE	HALF	SLR	77.50	13.81	28.35	.88	2.35	.56
HIGH	LARGE	HALF	ACE	397.87	38.43	90.92	4.97	3.85	.10
HIGH	LARGE	HALF	AM	89.62	13.12	31.75	1.10	1.85	.04
HIGH	LARGE	HALF	MARS	153.39	39.29	65.84	6.74	2.01	.11
HIGH	LARGE	HALF	RPR	4073.16	573.28	126.83	6.94	3.52	.04
HIGH	LARGE	HALF	PPR	374.01	33.02	46.97	6.31	1.28	.31
HIGH	LARGE	HALF	LOESS	137.53	13.49	33.20	1.42	*	*
HIGH	LARGE	HALF	AVAS	382.00	46.42	79.45	4.42	*	*
HIGH	LARGE	HALF	NN	2887.99	98.16	2602.01	23.92	*	*

Function: Correlated Gaussian

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	SMALL	ALL	MLR	598.41	31.62	1622.14	19.62	4043.74	264.05
LOW	SMALL	ALL	SLR	658.00	40.19	1653.15	19.41	4294.16	258.44
LOW	SMALL	ALL	ACE	584.82	47.84	1178.51	20.45	2749.71	228.44
LOW	SMALL	ALL	AM	598.48	31.74	1622.00	19.64	4043.73	264.05
LOW	SMALL	ALL	MARS	562.44	74.47	1723.48	206.00	3763.94	258.38
LOW	SMALL	ALL	RPR	1257.14	77.28	1938.70	53.01	4079.39	266.34
LOW	SMALL	ALL	PPR	570.85	41.47	1626.69	31.40	3915.18	257.59
LOW	SMALL	ALL	LOESS	154.81	16.25	1288.97	34.28	*	*
LOW	SMALL	ALL	AVAS	620.15	51.36	1150.57	20.77	2772.06	228.44
LOW	SMALL	ALL	NN	346.36	65.32	1247.36	56.60	*	*
LOW	SMALL	HALF	MLR	6.83	.35	766.05	2.53	1577.24	21.30
LOW	SMALL	HALF	SLR	6.57	.35	761.93	2.71	1679.06	50.17
LOW	SMALL	HALF	ACE	3.42	.46	690.31	16.30	1027.14	23.50
LOW	SMALL	HALF	AM	6.85	.36	765.93	2.50	1577.25	21.30
LOW	SMALL	HALF	MARS	5.03	.32	219.33	10.44	1122.25	20.03
LOW	SMALL	HALF	RPR	1060.87	15.00	1108.85	11.41	1463.01	14.99
LOW	SMALL	HALF	PPR	3.11	.52	757.46	16.10	1325.45	18.67
LOW	SMALL	HALF	LOESS	6.70	.79	374.55	8.33	*	*
LOW	SMALL	HALF	AVAS	9.82	3.48	592.37	15.20	997.33	23.60
LOW	SMALL	HALF	NN	21.18	2.37	107.92	4.27	*	*
LOW	MED	ALL	MLR	478.19	8.69	1606.58	20.87	4042.17	263.75
LOW	MED	ALL	SLR	478.19	8.69	1612.55	21.91	4287.93	263.05
LOW	MED	ALL	ACE	435.68	28.83	1113.04	20.88	2729.60	225.76
LOW	MED	ALL	AM	478.16	8.69	1606.33	20.88	4042.17	263.75
LOW	MED	ALL	MARS	142.45	13.77	1604.37	175.31	3879.00	259.65
LOW	MED	ALL	RPR	1320.84	38.73	1788.81	36.49	4079.03	266.18
LOW	MED	ALL	PPR	435.06	16.59	1497.57	26.33	3910.41	256.76
LOW	MED	ALL	LOESS	88.15	5.77	1219.76	29.07	*	*
LOW	MED	ALL	AVAS	367.85	27.55	1079.09	26.08	2751.82	226.22
LOW	MED	ALL	NN	69.18	10.26	597.25	21.41	*	*
LOW	MED	HALF	MLR	5.78	.18	751.27	2.05	1576.74	21.16
LOW	MED	HALF	SLR	5.68	.18	749.50	2.13	1705.44	55.19
LOW	MED	HALF	ACE	1.58	.23	703.26	12.79	1027.25	23.59
LOW	MED	HALF	AM	5.78	.18	751.15	1.99	1576.75	21.17
LOW	MED	HALF	MARS	1.97	.25	203.99	5.73	1219.29	21.51
LOW	MED	HALF	RPR	1046.53	10.33	1076.28	9.25	1470.91	17.89
LOW	MED	HALF	PPR	1.24	.17	693.45	15.16	1301.80	7.75
LOW	MED	HALF	LOESS	4.46	.24	332.70	5.92	*	*
LOW	MED	HALF	AVAS	1.71	.23	543.76	12.48	997.15	23.84
LOW	MED	HALF	NN	11.02	1.60	51.27	1.35	*	*

Function: Correlated Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	LARGE	ALL	MLR	450.12	2.93	1592.70	18.31	4041.97	263.88
LOW	LARGE	ALL	SLR	450.12	2.93	1592.70	18.31	4301.33	258.64
LOW	LARGE	ALL	ACE	366.84	29.71	1064.20	20.12	2725.90	225.52
LOW	LARGE	ALL	AM	450.16	2.93	1592.75	18.31	4041.97	263.88
LOW	LARGE	ALL	MARS	135.24	13.93	1312.92	38.80	3920.50	261.50
LOW	LARGE	ALL	RPR	1340.71	29.41	1761.54	26.19	4079.17	266.21
LOW	LARGE	ALL	PPR	411.03	18.23	1431.86	29.08	3907.11	257.10
LOW	LARGE	ALL	LOESS	79.81	2.83	1210.05	16.15	*	*
LOW	LARGE	ALL	AVAS	293.48	14.69	1031.11	21.07	*	*
LOW	LARGE	ALL	NN	38.77	1.97	297.34	7.45	*	*
LOW	LARGE	HALF	MLR	5.38	.10	745.41	2.04	1576.48	21.14
LOW	LARGE	HALF	SLR	5.33	.10	744.65	2.00	1692.56	52.27
LOW	LARGE	HALF	ACE	.59	.06	707.41	7.60	1026.09	23.82
LOW	LARGE	HALF	AM	5.38	.10	745.42	2.03	1576.49	21.14
LOW	LARGE	HALF	MARS	.72	.08	209.46	8.09	1317.54	22.46
LOW	LARGE	HALF	RPR	1050.93	4.86	1050.16	8.19	1461.56	14.03
LOW	LARGE	HALF	PPR	.52	.05	624.12	15.18	1375.47	31.18
LOW	LARGE	HALF	LOESS	3.91	.16	322.92	2.71	*	*
LOW	LARGE	HALF	AVAS	.70	.09	529.66	7.14	*	*
LOW	LARGE	HALF	NN	11.20	.88	39.73	.58	*	*
MOD	SMALL	ALL	MLR	638.37	40.57	1622.09	18.92	4043.89	264.07
MOD	SMALL	ALL	SLR	684.88	44.17	1656.90	20.09	4292.36	258.15
MOD	SMALL	ALL	ACE	618.46	46.52	1257.78	23.49	2711.47	227.18
MOD	SMALL	ALL	AM	639.25	40.86	1621.22	19.20	4043.90	264.07
MOD	SMALL	ALL	MARS	690.32	59.93	1637.32	62.72	3764.32	258.11
MOD	SMALL	ALL	RPR	1317.69	127.37	1961.95	57.79	4079.42	266.36
MOD	SMALL	ALL	PPR	573.28	44.07	1656.19	30.81	3915.74	257.56
MOD	SMALL	ALL	LOESS	195.62	19.40	1304.09	40.17	*	*
MOD	SMALL	ALL	AVAS	641.36	53.20	1260.74	21.45	2780.78	230.59
MOD	SMALL	ALL	NN	395.23	37.44	1289.73	53.39	*	*
MOD	SMALL	HALF	MLR	31.92	5.37	768.98	2.83	1577.31	21.30
MOD	SMALL	HALF	SLR	27.75	4.95	765.72	3.03	1705.83	54.65
MOD	SMALL	HALF	ACE	41.55	5.82	636.98	17.17	955.92	22.92
MOD	SMALL	HALF	AM	31.90	5.37	768.26	2.63	1577.31	21.30
MOD	SMALL	HALF	MARS	69.51	27.58	229.96	8.37	1122.89	20.09
MOD	SMALL	HALF	RPR	1154.00	60.09	1155.97	16.79	1473.54	16.21
MOD	SMALL	HALF	PPR	45.70	7.57	726.79	19.27	1324.93	18.80
MOD	SMALL	HALF	LOESS	59.41	8.10	386.10	9.73	*	*
MOD	SMALL	HALF	AVAS	50.06	6.95	609.56	14.01	992.03	22.85
MOD	SMALL	HALF	NN	109.11	15.60	226.01	6.92	*	*

Function: Correlated Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	ALL	MLR	485.34	10.37	1607.58	20.64	4042.19	263.76
MOD	MED	ALL	SLR	485.34	10.37	1615.91	21.57	4286.95	263.37
MOD	MED	ALL	ACE	445.56	23.11	1130.62	25.75	2690.31	223.15
MOD	MED	ALL	AM	485.26	10.36	1606.40	20.65	4042.18	263.76
MOD	MED	ALL	MARS	183.40	12.20	1503.30	91.00	3881.47	259.66
MOD	MED	ALL	RPR	1545.33	86.44	1789.52	39.30	4079.01	266.17
MOD	MED	ALL	PPR	447.66	21.64	1494.76	24.95	3910.77	256.76
MOD	MED	ALL	LOESS	98.25	8.50	1226.51	28.33	*	*
MOD	MED	ALL	AVAS	438.87	34.77	1165.97	21.43	2746.67	225.33
MOD	MED	ALL	NN	155.38	16.01	758.49	23.05	*	*
MOD	MED	HALF	MLR	14.33	1.65	752.21	2.50	1576.76	21.17
MOD	MED	HALF	SLR	13.28	1.68	750.28	2.52	1670.26	48.67
MOD	MED	HALF	ACE	30.05	4.08	573.58	11.55	954.23	22.91
MOD	MED	HALF	AM	14.32	1.65	751.59	2.21	1576.76	21.16
MOD	MED	HALF	MARS	16.38	2.91	215.66	6.27	1219.49	21.47
MOD	MED	HALF	RPR	1189.13	46.66	1089.25	11.09	1472.17	18.27
MOD	MED	HALF	PPR	19.41	3.19	691.13	19.61	1352.03	22.24
MOD	MED	HALF	LOESS	20.93	3.06	336.67	5.54	*	*
MOD	MED	HALF	AVAS	28.94	4.26	520.15	10.63	989.88	23.61
MOD	MED	HALF	NN	95.18	9.40	138.85	4.24	*	*
MOD	LARGE	ALL	MLR	452.68	3.50	1592.87	18.32	4042.00	263.88
MOD	LARGE	ALL	SLR	452.68	3.50	1592.87	18.32	4300.50	258.69
MOD	LARGE	ALL	ACE	346.26	19.41	1053.04	18.21	2686.97	223.45
MOD	LARGE	ALL	AM	452.72	3.52	1593.18	18.30	4041.99	263.88
MOD	LARGE	ALL	MARS	146.53	10.16	1437.32	166.77	3919.90	261.60
MOD	LARGE	ALL	RPR	1461.81	46.98	1765.35	20.55	4079.18	266.20
MOD	LARGE	ALL	PPR	403.53	10.87	1439.26	30.93	3907.42	257.10
MOD	LARGE	ALL	LOESS	82.19	3.57	1211.01	16.38	*	*
MOD	LARGE	ALL	AVAS	317.48	21.87	1082.54	19.33	*	*
MOD	LARGE	ALL	NN	95.77	7.32	417.64	10.15	*	*
MOD	LARGE	HALF	MLR	8.72	.56	745.69	2.05	1576.49	21.15
MOD	LARGE	HALF	SLR	8.37	.59	745.04	2.06	1604.09	29.61
MOD	LARGE	HALF	ACE	13.30	1.28	553.39	8.00	952.57	23.14
MOD	LARGE	HALF	AM	8.68	.55	745.73	2.04	1576.47	21.15
MOD	LARGE	HALF	MARS	7.49	.89	208.41	6.13	1318.48	22.36
MOD	LARGE	HALF	RPR	1073.66	18.64	1060.30	9.46	1464.13	13.60
MOD	LARGE	HALF	PPR	9.41	.81	671.39	17.22	1394.31	36.16
MOD	LARGE	HALF	LOESS	9.09	.62	324.35	2.64	*	*
MOD	LARGE	HALF	AVAS	12.32	1.34	505.61	6.73	*	*
MOD	LARGE	HALF	NN	90.34	4.55	98.21	1.24	*	*

Function: Correlated Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	SMALL	ALL	MLR	1374.89	243.50	1679.52	17.75	4046.28	264.18
HIGH	SMALL	ALL	SLR	1254.59	218.73	1709.14	18.82	4599.13	304.67
HIGH	SMALL	ALL	ACE	2052.25	245.68	1903.08	51.69	2642.32	222.89
HIGH	SMALL	ALL	AM	1375.59	243.10	1672.77	20.09	4046.29	264.18
HIGH	SMALL	ALL	MARS	1200.72	214.84	1848.00	32.85	3772.80	258.11
HIGH	SMALL	ALL	RPR	5376.83	1185.31	3591.36	227.02	4079.62	266.44
HIGH	SMALL	ALL	PPR	1894.11	247.50	1878.26	26.86	3919.92	257.42
HIGH	SMALL	ALL	LOESS	1542.14	220.89	1555.21	85.94	*	*
HIGH	SMALL	ALL	AVAS	1906.05	231.13	1911.42	51.53	2843.05	237.35
HIGH	SMALL	ALL	NN	2964.48	290.53	4017.83	196.86	*	*
HIGH	SMALL	HALF	MLR	694.06	165.60	841.30	10.56	1579.29	21.26
HIGH	SMALL	HALF	SLR	562.98	170.48	855.92	12.89	1663.22	52.21
HIGH	SMALL	HALF	ACE	1175.59	172.64	1113.79	48.60	800.72	19.38
HIGH	SMALL	HALF	AM	693.05	165.24	835.09	9.61	1579.32	21.27
HIGH	SMALL	HALF	MARS	798.05	81.37	946.28	115.36	1132.30	20.26
HIGH	SMALL	HALF	RPR	5265.97	815.08	3773.11	351.06	1471.88	18.48
HIGH	SMALL	HALF	PPR	1209.30	177.62	1135.67	34.90	1316.34	14.46
HIGH	SMALL	HALF	LOESS	1465.39	220.76	619.38	36.72	*	*
HIGH	SMALL	HALF	AVAS	1211.21	181.22	1120.12	50.12	915.10	19.57
HIGH	SMALL	HALF	NN	2980.87	329.27	2865.79	60.43	*	*
HIGH	MED	ALL	MLR	683.98	46.59	1637.11	21.39	4042.84	263.79
HIGH	MED	ALL	SLR	737.86	34.60	1656.57	21.81	4282.89	265.11
HIGH	MED	ALL	ACE	1310.41	136.27	1286.18	29.32	2586.70	212.92
HIGH	MED	ALL	AM	683.49	46.55	1629.46	20.54	4042.84	263.79
HIGH	MED	ALL	MARS	969.53	101.47	1539.72	17.93	3885.77	259.04
HIGH	MED	ALL	RPR	6894.90	1436.55	1882.34	67.58	4078.99	266.16
HIGH	MED	ALL	PPR	1218.27	128.68	1707.35	23.33	3911.49	256.38
HIGH	MED	ALL	LOESS	487.90	81.60	1328.66	34.14	*	*
HIGH	MED	ALL	AVAS	1277.44	148.94	1363.79	21.99	2723.64	221.82
HIGH	MED	ALL	NN	3040.78	254.25	3483.62	78.60	*	*
HIGH	MED	HALF	MLR	219.98	35.41	781.37	6.28	1577.40	21.19
HIGH	MED	HALF	SLR	254.58	53.36	776.36	7.92	1677.74	50.29
HIGH	MED	HALF	ACE	778.16	116.41	706.57	16.52	774.30	19.85
HIGH	MED	HALF	AM	219.35	35.37	776.72	4.14	1577.40	21.20
HIGH	MED	HALF	MARS	439.34	69.47	387.07	16.53	1221.02	21.25
HIGH	MED	HALF	RPR	5845.81	1013.19	1439.59	89.39	1445.29	20.23
HIGH	MED	HALF	PPR	683.83	103.67	853.65	28.32	1311.27	13.76
HIGH	MED	HALF	LOESS	442.44	79.10	424.85	11.39	*	*
HIGH	MED	HALF	AVAS	815.80	115.49	687.80	19.66	889.81	9.80
HIGH	MED	HALF	NN	2887.43	213.70	2784.16	55.70	*	*

Function: Correlated Gaussian (Cont.)

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
HIGH	LARGE	ALL	MLR	530.07	10.67	1601.13	18.37	4042.36	263.92
HIGH	LARGE	ALL	SLR	569.40	17.79	1609.15	18.53	4296.64	258.93
HIGH	LARGE	ALL	ACE	757.42	51.40	1023.01	25.47	2567.29	217.05
HIGH	LARGE	ALL	AM	529.65	10.66	1603.19	18.54	4042.36	263.92
HIGH	LARGE	ALL	MARS	433.62	45.75	1382.16	24.07	3919.93	262.06
HIGH	LARGE	ALL	RPR	3760.96	462.71	1675.16	29.92	4079.25	266.20
HIGH	LARGE	ALL	PPR	762.97	46.53	1505.06	33.32	3907.78	256.91
HIGH	LARGE	ALL	LOESS	196.60	16.78	1238.65	19.75	*	*
HIGH	LARGE	ALL	AVAS	735.05	60.12	1168.67	30.34	*	*
HIGH	LARGE	ALL	NN	2820.19	129.41	3242.85	48.54	*	*
HIGH	LARGE	HALF	MLR	89.99	13.18	754.45	2.52	1576.76	21.17
HIGH	LARGE	HALF	SLR	77.50	13.81	751.97	2.45	1718.31	65.68
HIGH	LARGE	HALF	ACE	397.87	38.43	495.18	7.91	760.21	20.40
HIGH	LARGE	HALF	AM	89.62	13.12	755.20	2.69	1576.74	21.17
HIGH	LARGE	HALF	MARS	153.39	39.29	248.03	8.93	1319.75	22.27
HIGH	LARGE	HALF	RPR	4073.16	573.28	1101.52	15.99	1452.50	14.76
HIGH	LARGE	HALF	PPR	374.01	33.02	654.27	16.69	1309.09	13.53
HIGH	LARGE	HALF	LOESS	137.53	13.49	354.33	4.82	*	*
HIGH	LARGE	HALF	AVAS	382.00	46.42	503.87	7.47	*	*
HIGH	LARGE	HALF	NN	2821.98	85.91	2692.19	26.89	*	*

Function: Mixture

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
LOW	SMALL	ALL	MLR	177.73	11.28	7.69	.10	.02	.00
LOW	SMALL	ALL	SLR	168.65	10.96	7.55	.11	.02	.00
LOW	SMALL	ALL	ACE	169.83	9.74	9.65	.24	.05	.00
LOW	SMALL	ALL	AM	177.80	11.32	7.69	.11	.02	.00
LOW	SMALL	ALL	MARS	90.99	12.85	13.88	4.36	.04	.00
LOW	SMALL	ALL	RPR	339.45	39.62	32.01	12.49	.02	.00
LOW	SMALL	ALL	PPR	50.46	6.90	6.37	.59	.02	.00
LOW	SMALL	ALL	LOESS	28.81	4.97	2.51	.14	*	*
LOW	SMALL	ALL	AVAS	157.17	10.09	9.31	.21	.02	.00
LOW	SMALL	ALL	NN	108.23	13.73	10.28	.53	*	*
LOW	SMALL	HALF	MLR	8.81	.68	89.08	.61	7.33	.10
LOW	SMALL	HALF	SLR	8.31	.74	88.17	.62	7.32	.10
LOW	SMALL	HALF	ACE	3.40	.55	92.48	1.06	7.27	.10
LOW	SMALL	HALF	AM	8.80	.68	89.07	.61	7.33	.10
LOW	SMALL	HALF	MARS	7.06	.34	8.48	2.43	4.16	.15
LOW	SMALL	HALF	RPR	20.60	3.05	157.96	24.11	7.32	.10
LOW	SMALL	HALF	PPR	4.53	1.07	11.36	.13	1.09	.02
LOW	SMALL	HALF	LOESS	6.10	.61	14.08	.49	*	*
LOW	SMALL	HALF	AVAS	3.42	.54	103.06	.96	7.17	.11
LOW	SMALL	HALF	NN	7.41	.92	20.41	.64	*	*
LOW	MED	ALL	MLR	133.93	1.84	7.60	.13	.02	.00
LOW	MED	ALL	SLR	128.46	1.65	7.56	.13	.02	.00
LOW	MED	ALL	ACE	141.80	6.27	7.95	.15	.03	.00
LOW	MED	ALL	AM	133.97	1.85	7.58	.13	.02	.00
LOW	MED	ALL	MARS	6.84	.57	6.97	.95	.02	.00
LOW	MED	ALL	RPR	303.08	27.62	27.80	12.83	.02	.00
LOW	MED	ALL	PPR	23.60	.82	2.43	.50	.02	.00
LOW	MED	ALL	LOESS	8.02	.48	2.12	.10	*	*
LOW	MED	ALL	AVAS	156.47	5.45	8.26	.13	.02	.00
LOW	MED	ALL	NN	27.89	2.47	7.42	.19	*	*
LOW	MED	HALF	MLR	7.10	.11	87.65	.56	7.32	.10
LOW	MED	HALF	SLR	6.76	.12	87.22	.54	7.32	.10
LOW	MED	HALF	ACE	1.36	.17	88.64	1.01	7.29	.10
LOW	MED	HALF	AM	7.10	.11	87.57	.55	7.32	.10
LOW	MED	HALF	MARS	1.04	.24	7.78	2.45	4.65	.12
LOW	MED	HALF	RPR	19.22	1.58	106.15	1.67	7.32	.10
LOW	MED	HALF	PPR	1.95	.48	10.50	.06	1.38	.30
LOW	MED	HALF	LOESS	2.95	.23	11.67	.21	*	*
LOW	MED	HALF	AVAS	1.39	.18	92.45	.86	7.21	.12
LOW	MED	HALF	NN	4.88	.36	10.18	.51	*	*

Function: Mixture (Cont.)

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
LOW	LARGE	ALL	MLR	124.38	.90	7.47	.11	.02	.00
LOW	LARGE	ALL	SLR	123.32	1.01	7.45	.11	.02	.00
LOW	LARGE	ALL	ACE	128.19	1.66	7.63	.10	.02	.00
LOW	LARGE	ALL	AM	124.38	.90	7.48	.11	.02	.00
LOW	LARGE	ALL	MARS	4.29	.12	7.06	1.45	.02	.00
LOW	LARGE	ALL	RPR	303.68	11.54	18.19	10.47	.02	.00
LOW	LARGE	ALL	PPR	20.01	.32	1.25	.02	.02	.00
LOW	LARGE	ALL	LOESS	5.79	.34	1.91	.06	*	*
LOW	LARGE	ALL	AVAS	132.87	5.95	7.69	.09	*	*
LOW	LARGE	ALL	NN	10.25	.68	6.04	.10	*	*
LOW	LARGE	HALF	MLR	6.62	.06	87.01	.53	7.32	.10
LOW	LARGE	HALF	SLR	6.55	.07	86.86	.53	7.32	.10
LOW	LARGE	HALF	ACE	.61	.06	86.33	.48	7.29	.10
LOW	LARGE	HALF	AM	6.63	.06	87.01	.53	7.32	.10
LOW	LARGE	HALF	MARS	.46	.04	7.46	.81	5.64	.11
LOW	LARGE	HALF	RPR	15.68	1.07	97.18	2.44	7.32	.10
LOW	LARGE	HALF	PPR	1.45	.48	13.11	2.93	1.07	.01
LOW	LARGE	HALF	LOESS	2.12	.14	10.83	.18	*	*
LOW	LARGE	HALF	AVAS	.67	.06	89.40	.58	*	*
LOW	LARGE	HALF	NN	4.54	.21	4.79	.21	*	*
MOD	SMALL	ALL	MLR	213.82	21.48	10.31	.29	.10	.01
MOD	SMALL	ALL	SLR	188.81	20.00	9.10	.31	.06	.01
MOD	SMALL	ALL	ACE	251.83	20.99	31.37	1.66	.82	.03
MOD	SMALL	ALL	AM	213.84	21.48	10.18	.29	.10	.01
MOD	SMALL	ALL	MARS	156.58	18.39	19.52	2.25	.47	.03
MOD	SMALL	ALL	RPR	681.04	151.76	107.05	7.76	.02	.00
MOD	SMALL	ALL	PPR	132.58	14.39	18.11	.82	.25	.02
MOD	SMALL	ALL	LOESS	83.66	12.37	11.08	.69	*	*
MOD	SMALL	ALL	AVAS	256.47	18.67	31.42	1.61	.22	.01
MOD	SMALL	ALL	NN	229.76	28.70	111.57	3.11	*	*
MOD	SMALL	HALF	MLR	36.10	7.55	91.72	.65	7.40	.10
MOD	SMALL	HALF	SLR	28.06	7.13	89.87	.64	7.37	.10
MOD	SMALL	HALF	ACE	58.03	7.62	123.11	2.84	8.12	.09
MOD	SMALL	HALF	AM	36.07	7.55	91.54	.56	7.40	.10
MOD	SMALL	HALF	MARS	20.04	6.86	36.33	6.86	3.92	.31
MOD	SMALL	HALF	RPR	190.43	52.48	240.31	26.34	7.32	.10
MOD	SMALL	HALF	PPR	70.84	10.07	57.97	7.67	4.39	.39
MOD	SMALL	HALF	LOESS	66.33	9.72	22.71	1.22	*	*
MOD	SMALL	HALF	AVAS	59.22	8.27	126.63	3.06	7.41	.11
MOD	SMALL	HALF	NN	133.53	13.19	129.33	3.77	*	*

Function: Mixture (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	ALL	MLR	141.58	3.33	8.79	.20	.05	.00
MOD	MED	ALL	SLR	134.43	2.78	8.36	.21	.03	.00
MOD	MED	ALL	ACE	188.59	13.14	18.17	.69	.33	.02
MOD	MED	ALL	AM	141.58	3.34	8.61	.16	.05	.00
MOD	MED	ALL	MARS	66.56	12.66	13.11	1.23	.18	.02
MOD	MED	ALL	RPR	471.57	54.75	23.95	5.68	.02	.00
MOD	MED	ALL	PPR	70.98	9.21	10.64	.75	.10	.00
MOD	MED	ALL	LOESS	24.98	3.50	5.48	.18	*	*
MOD	MED	ALL	AVAS	191.14	11.51	17.62	.64	.08	.00
MOD	MED	ALL	NN	146.08	12.71	109.82	2.08	*	*
MOD	MED	HALF	MLR	15.01	1.21	88.82	.60	7.35	.10
MOD	MED	HALF	SLR	11.78	1.39	88.15	.58	7.34	.10
MOD	MED	HALF	ACE	32.16	3.94	97.04	2.03	7.49	.10
MOD	MED	HALF	AM	15.01	1.20	88.42	.53	7.35	.10
MOD	MED	HALF	MARS	11.65	2.06	14.07	3.62	4.85	.21
MOD	MED	HALF	RPR	170.25	52.52	110.48	3.29	7.32	.10
MOD	MED	HALF	PPR	40.78	6.00	25.59	5.70	4.09	.48
MOD	MED	HALF	LOESS	20.74	3.27	14.63	.45	*	*
MOD	MED	HALF	AVAS	35.81	5.20	102.38	1.33	7.34	.10
MOD	MED	HALF	NN	119.83	10.51	117.16	1.98	*	*
MOD	LARGE	ALL	MLR	127.45	1.29	7.82	.11	.03	.00
MOD	LARGE	ALL	SLR	125.63	1.44	7.63	.12	.02	.00
MOD	LARGE	ALL	ACE	155.51	5.67	11.46	.36	.15	.01
MOD	LARGE	ALL	AM	127.43	1.28	7.89	.13	.03	.00
MOD	LARGE	ALL	MARS	19.18	5.78	8.80	.29	.05	.00
MOD	LARGE	ALL	RPR	429.72	38.26	7.53	.11	.02	.00
MOD	LARGE	ALL	PPR	35.36	2.95	7.01	.75	.05	.00
MOD	LARGE	ALL	LOESS	9.98	.96	3.02	.11	*	*
MOD	LARGE	ALL	AVAS	160.58	7.14	11.17	.27	*	*
MOD	LARGE	ALL	NN	113.34	5.22	108.51	.81	*	*
MOD	LARGE	HALF	MLR	9.87	.67	87.34	.51	7.33	.10
MOD	LARGE	HALF	SLR	8.96	.74	87.10	.50	7.33	.10
MOD	LARGE	HALF	ACE	19.16	1.94	87.86	.54	7.31	.10
MOD	LARGE	HALF	AM	9.87	.67	87.36	.52	7.33	.10
MOD	LARGE	HALF	MARS	8.65	.96	13.94	3.83	5.70	.13
MOD	LARGE	HALF	RPR	148.95	19.54	90.70	2.22	7.32	.10
MOD	LARGE	HALF	PPR	19.75	1.52	11.24	.14	3.17	.40
MOD	LARGE	HALF	LOESS	7.33	.68	11.92	.29	*	*
MOD	LARGE	HALF	AVAS	19.53	1.88	92.51	.57	*	*
MOD	LARGE	HALF	NN	115.13	4.76	110.22	.87	*	*

Function: Mixture (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	SMALL	ALL	MLR	930.98	214.06	81.11	6.57	2.15	.25
HIGH	SMALL	ALL	SLR	776.82	226.72	50.84	7.00	1.18	.26
HIGH	SMALL	ALL	ACE	1585.47	236.76	611.84	36.43	20.39	.79
HIGH	SMALL	ALL	AM	931.64	214.30	78.08	6.53	2.15	.25
HIGH	SMALL	ALL	MARS	515.47	206.26	290.54	61.96	11.52	.79
HIGH	SMALL	ALL	RPR	5392.70	1305.79	2398.97	232.29	.12	.03
HIGH	SMALL	ALL	PPR	1526.85	221.79	267.18	22.81	5.88	.48
HIGH	SMALL	ALL	LOESS	1500.36	231.67	229.46	13.60	*	*
HIGH	SMALL	ALL	AVAS	1627.60	202.14	570.32	36.10	5.13	.21
HIGH	SMALL	ALL	NN	2887.25	284.15	2664.36	59.03	*	*
HIGH	SMALL	HALF	MLR	709.26	177.14	162.57	7.14	9.44	.28
HIGH	SMALL	HALF	SLR	526.85	174.12	131.30	7.30	8.47	.29
HIGH	SMALL	HALF	ACE	1527.01	262.11	684.15	45.30	27.05	.76
HIGH	SMALL	HALF	AM	709.69	177.35	159.16	6.64	9.44	.28
HIGH	SMALL	HALF	MARS	172.50	43.30	382.14	54.02	19.46	.51
HIGH	SMALL	HALF	RPR	4488.50	1076.89	2061.83	148.46	7.42	.10
HIGH	SMALL	HALF	PPR	1378.45	202.11	332.15	25.24	12.51	.54
HIGH	SMALL	HALF	LOESS	1509.97	230.64	241.42	14.56	*	*
HIGH	SMALL	HALF	AVAS	1297.72	170.05	655.83	43.02	12.53	.33
HIGH	SMALL	HALF	NN	2793.76	282.68	2643.98	65.20	*	*
HIGH	MED	ALL	MLR	342.68	34.75	39.22	3.41	.77	.06
HIGH	MED	ALL	SLR	255.28	37.51	30.61	3.62	.47	.07
HIGH	MED	ALL	ACE	1098.36	163.18	267.46	16.30	8.11	.44
HIGH	MED	ALL	AM	342.77	34.74	36.60	3.13	.77	.06
HIGH	MED	ALL	MARS	277.60	56.34	150.07	31.21	3.94	.46
HIGH	MED	ALL	RPR	4573.85	1095.03	293.49	117.92	.09	.03
HIGH	MED	ALL	PPR	917.05	95.23	127.88	11.09	2.12	.12
HIGH	MED	ALL	LOESS	448.96	78.71	90.68	3.37	*	*
HIGH	MED	ALL	AVAS	1120.93	163.88	239.97	17.96	1.71	.09
HIGH	MED	ALL	NN	2922.62	209.02	2644.17	51.73	*	*
HIGH	MED	HALF	MLR	217.45	32.97	119.18	3.54	8.07	.12
HIGH	MED	HALF	SLR	147.54	37.62	109.46	3.83	7.77	.13
HIGH	MED	HALF	ACE	890.10	122.50	343.70	12.48	15.01	.39
HIGH	MED	HALF	AM	217.41	32.99	115.43	2.88	8.07	.12
HIGH	MED	HALF	MARS	137.80	50.96	193.82	27.47	10.67	.57
HIGH	MED	HALF	RPR	4524.86	1080.86	359.60	66.56	7.39	.10
HIGH	MED	HALF	PPR	819.49	121.93	184.90	9.31	8.10	.51
HIGH	MED	HALF	LOESS	448.90	80.38	97.82	3.34	*	*
HIGH	MED	HALF	AVAS	896.06	115.02	335.54	14.02	9.04	.14
HIGH	MED	HALF	NN	2888.37	199.72	2724.56	54.13	*	*

Function: Mixture (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	LARGE	ALL	MLR	207.38	12.51	16.99	.75	.31	.02
HIGH	LARGE	ALL	SLR	185.60	14.44	12.05	1.00	.17	.03
HIGH	LARGE	ALL	ACE	553.94	56.01	110.91	6.52	3.49	.15
HIGH	LARGE	ALL	AM	207.34	12.52	17.90	1.10	.31	.02
HIGH	LARGE	ALL	MARS	213.76	49.95	53.12	11.16	1.02	.09
HIGH	LARGE	ALL	RPR	3142.94	577.07	8.89	.53	.04	.01
HIGH	LARGE	ALL	PPR	429.75	44.17	49.26	3.85	.83	.04
HIGH	LARGE	ALL	LOESS	133.44	12.78	31.42	1.53	*	*
HIGH	LARGE	ALL	AVAS	561.32	49.81	95.66	4.91	*	*
HIGH	LARGE	ALL	NN	2905.53	128.48	2631.35	23.63	*	*
HIGH	LARGE	HALF	MLR	90.64	13.88	96.36	.85	7.61	.11
HIGH	LARGE	HALF	SLR	70.84	15.33	92.50	.96	7.48	.11
HIGH	LARGE	HALF	ACE	462.59	54.24	187.64	8.48	10.79	.19
HIGH	LARGE	HALF	AM	90.58	13.85	97.07	1.24	7.61	.11
HIGH	LARGE	HALF	MARS	46.67	15.17	84.37	10.76	7.53	.28
HIGH	LARGE	HALF	RPR	4039.17	1072.73	87.98	.71	7.35	.10
HIGH	LARGE	HALF	PPR	381.34	22.63	98.79	9.01	5.84	.53
HIGH	LARGE	HALF	LOESS	135.94	13.57	40.19	1.60	*	*
HIGH	LARGE	HALF	AVAS	463.60	53.29	174.34	5.47	*	*
HIGH	LARGE	HALF	NN	2901.43	120.27	2652.75	26.74	*	*

Function: Product

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
LOW	SMALL	ALL	MLR	71.88	3.97	40.41	.22	23.64	.10
LOW	SMALL	ALL	SLR	71.88	3.97	40.41	.22	23.64	.10
LOW	SMALL	ALL	ACE	53.06	4.52	4.06	.27	.84	.01
LOW	SMALL	ALL	AM	71.87	3.97	40.40	.22	23.64	.10
LOW	SMALL	ALL	MARS	57.09	4.88	20.11	.59	23.85	.15
LOW	SMALL	ALL	RPR	838.57	116.55	151.48	4.82	83.06	.60
LOW	SMALL	ALL	PPR	77.38	6.31	39.39	.30	22.99	.10
LOW	SMALL	ALL	LOESS	19.97	2.01	31.19	.63	*	*
LOW	SMALL	ALL	AVAS	111.98	15.20	18.44	1.06	7.55	.06
LOW	SMALL	ALL	NN	69.56	5.84	39.71	1.31	*	*
LOW	SMALL	HALF	MLR	1.12	.28	54.73	.30	39.38	.21
LOW	SMALL	HALF	SLR	.93	.27	54.52	.31	39.38	.21
LOW	SMALL	HALF	ACE	4.57	1.05	5.35	.39	.88	.02
LOW	SMALL	HALF	AM	1.13	.29	54.72	.30	39.39	.21
LOW	SMALL	HALF	MARS	2.11	.65	7.80	.47	10.60	.13
LOW	SMALL	HALF	RPR	1651.73	26.01	298.87	5.45	132.57	1.88
LOW	SMALL	HALF	PPR	3.86	.98	50.38	.33	37.32	.20
LOW	SMALL	HALF	LOESS	2.38	.36	29.68	.70	*	*
LOW	SMALL	HALF	AVAS	16.83	5.73	38.09	2.43	12.37	.10
LOW	SMALL	HALF	NN	30.98	2.69	27.35	.87	*	*
LOW	MED	ALL	MLR	57.09	1.33	39.75	.20	23.63	.10
LOW	MED	ALL	SLR	57.09	1.33	39.75	.20	23.63	.10
LOW	MED	ALL	ACE	23.31	2.34	1.80	.07	.81	.01
LOW	MED	ALL	AM	57.10	1.33	39.76	.20	23.63	.10
LOW	MED	ALL	MARS	12.14	1.62	16.09	.19	50.27	.19
LOW	MED	ALL	RPR	709.84	73.70	125.23	1.81	83.40	.57
LOW	MED	ALL	PPR	53.49	1.55	38.01	.20	22.97	.10
LOW	MED	ALL	LOESS	13.25	.73	29.89	.40	*	*
LOW	MED	ALL	AVAS	72.12	6.74	12.00	1.00	7.52	.05
LOW	MED	ALL	NN	38.11	1.84	25.89	.67	*	*
LOW	MED	HALF	MLR	.34	.06	53.76	.30	39.36	.22
LOW	MED	HALF	SLR	.30	.06	53.68	.30	39.36	.22
LOW	MED	HALF	ACE	1.75	.33	1.79	.11	.84	.02
LOW	MED	HALF	AM	.41	.07	53.76	.30	39.37	.22
LOW	MED	HALF	MARS	.64	.13	5.61	.14	14.38	.09
LOW	MED	HALF	RPR	1636.55	12.22	297.78	4.54	130.91	2.32
LOW	MED	HALF	PPR	1.18	.28	48.87	.27	37.30	.20
LOW	MED	HALF	LOESS	.71	.13	27.04	.30	*	*
LOW	MED	HALF	AVAS	2.79	.65	23.50	1.46	12.28	.12
LOW	MED	HALF	NN	21.54	1.44	16.67	.49	*	*

Function: Product (Cont.)

Noise	n	Var. Expl.	Method	MISE $p = 2$	St. Err.	MISE $p = 6$	St. Err.	MISE $p = 12$	St. Err.
LOW	LARGE	ALL	MLR	53.12	.61	39.43	.18	23.62	.10
LOW	LARGE	ALL	SLR	53.12	.61	39.43	.18	23.63	.10
LOW	LARGE	ALL	ACE	9.95	1.31	1.15	.02	.80	.01
LOW	LARGE	ALL	AM	53.11	.61	39.44	.18	23.63	.10
LOW	LARGE	ALL	MARS	9.57	.89	14.70	.09	66.20	.19
LOW	LARGE	ALL	RPR	578.02	14.77	115.89	1.49	83.31	.52
LOW	LARGE	ALL	PPR	46.98	.81	37.49	.18	22.97	.10
LOW	LARGE	ALL	LOESS	11.42	.41	29.43	.19	*	*
LOW	LARGE	ALL	AVAS	69.50	7.38	8.67	.35	*	*
LOW	LARGE	ALL	NN	25.74	1.38	17.15	.26	*	*
LOW	LARGE	HALF	MLR	.13	.02	53.46	.27	39.35	.22
LOW	LARGE	HALF	SLR	.11	.02	53.41	.27	81.34	41.88
LOW	LARGE	HALF	ACE	.52	.06	1.01	.03	.83	.02
LOW	LARGE	HALF	AM	.14	.02	53.46	.27	39.36	.22
LOW	LARGE	HALF	MARS	.16	.04	5.24	.19	46.49	.23
LOW	LARGE	HALF	RPR	1645.60	8.02	296.95	2.95	130.09	2.60
LOW	LARGE	HALF	PPR	.31	.04	48.46	.24	37.29	.20
LOW	LARGE	HALF	LOESS	.21	.02	26.44	.13	*	*
LOW	LARGE	HALF	AVAS	.92	.37	17.66	.75	*	*
LOW	LARGE	HALF	NN	20.13	1.30	13.38	.30	*	*
MOD	SMALL	ALL	MLR	100.90	11.42	43.12	.47	23.71	.10
MOD	SMALL	ALL	SLR	103.57	13.27	43.12	.47	23.73	.11
MOD	SMALL	ALL	ACE	95.44	11.77	25.84	1.57	1.81	.03
MOD	SMALL	ALL	AM	100.92	11.44	43.12	.47	23.71	.10
MOD	SMALL	ALL	MARS	188.31	38.04	37.58	1.80	24.14	.10
MOD	SMALL	ALL	RPR	1343.09	254.31	210.89	10.25	87.34	.81
MOD	SMALL	ALL	PPR	123.21	14.00	66.18	10.74	23.11	.11
MOD	SMALL	ALL	LOESS	67.66	10.54	40.57	1.97	*	*
MOD	SMALL	ALL	AVAS	116.75	14.17	26.72	1.59	8.35	.06
MOD	SMALL	ALL	NN	164.98	21.19	138.00	3.97	*	*
MOD	SMALL	HALF	MLR	27.95	7.02	57.53	.61	39.46	.21
MOD	SMALL	HALF	SLR	23.23	6.75	56.90	.64	82.06	42.62
MOD	SMALL	HALF	ACE	42.42	5.92	28.81	1.87	2.47	.06
MOD	SMALL	HALF	AM	27.68	7.00	57.54	.60	39.46	.21
MOD	SMALL	HALF	MARS	17.99	3.24	27.86	1.51	10.88	.18
MOD	SMALL	HALF	RPR	1821.76	64.32	322.14	6.40	139.62	2.72
MOD	SMALL	HALF	PPR	35.31	5.61	56.39	1.03	37.43	.20
MOD	SMALL	HALF	LOESS	59.50	9.12	39.24	2.02	*	*
MOD	SMALL	HALF	AVAS	74.93	14.13	32.94	1.87	11.55	.12
MOD	SMALL	HALF	NN	105.95	10.85	123.74	3.94	*	*

Function: Product (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
MOD	MED	ALL	MLR	66.02	3.20	41.01	.28	23.66	.10
MOD	MED	ALL	SLR	66.02	3.20	41.01	.28	23.65	.10
MOD	MED	ALL	ACE	64.93	8.59	12.93	.41	1.35	.02
MOD	MED	ALL	AM	66.00	3.20	40.99	.28	23.66	.10
MOD	MED	ALL	MARS	53.83	5.32	23.49	.75	50.35	.15
MOD	MED	ALL	RPR	935.97	113.40	149.67	4.94	88.12	.67
MOD	MED	ALL	PPR	77.98	6.14	39.99	.39	23.02	.11
MOD	MED	ALL	LOESS	28.58	3.82	33.91	.94	*	*
MOD	MED	ALL	AVAS	80.66	11.65	12.58	.46	8.20	.04
MOD	MED	ALL	NN	121.31	11.58	127.91	2.33	*	*
MOD	MED	HALF	MLR	8.48	1.38	55.00	.37	39.39	.22
MOD	MED	HALF	SLR	7.38	1.44	54.65	.40	39.38	.22
MOD	MED	HALF	ACE	29.42	4.45	15.99	.45	1.90	.04
MOD	MED	HALF	AM	8.52	1.42	54.99	.37	39.39	.22
MOD	MED	HALF	MARS	16.73	3.84	11.32	.58	14.50	.09
MOD	MED	HALF	RPR	1708.05	41.80	309.07	5.08	142.02	2.84
MOD	MED	HALF	PPR	18.07	3.28	50.89	.30	37.33	.20
MOD	MED	HALF	LOESS	17.67	3.21	30.71	.74	*	*
MOD	MED	HALF	AVAS	33.71	4.57	19.18	.73	11.40	.11
MOD	MED	HALF	NN	100.45	9.21	114.59	2.22	*	*
MOD	LARGE	ALL	MLR	56.31	.83	39.80	.20	23.64	.10
MOD	LARGE	ALL	SLR	56.31	.83	39.80	.20	23.64	.10
MOD	LARGE	ALL	ACE	32.12	2.16	6.26	.19	1.14	.02
MOD	LARGE	ALL	AM	56.32	.84	39.82	.20	23.64	.10
MOD	LARGE	ALL	MARS	21.13	2.48	17.17	.30	66.08	.18
MOD	LARGE	ALL	RPR	836.93	53.83	126.06	1.45	89.00	.69
MOD	LARGE	ALL	PPR	55.26	1.33	38.29	.19	22.99	.10
MOD	LARGE	ALL	LOESS	15.89	1.03	30.53	.39	*	*
MOD	LARGE	ALL	AVAS	43.47	2.62	7.46	.29	*	*
MOD	LARGE	ALL	NN	103.61	4.91	120.83	1.61	*	*
MOD	LARGE	HALF	MLR	3.36	.53	53.85	.29	39.37	.22
MOD	LARGE	HALF	SLR	2.84	.56	53.69	.28	39.37	.22
MOD	LARGE	HALF	ACE	15.45	1.64	8.51	.33	1.72	.03
MOD	LARGE	HALF	AM	3.38	.52	53.85	.29	39.37	.22
MOD	LARGE	HALF	MARS	5.26	1.54	7.26	.25	46.54	.21
MOD	LARGE	HALF	RPR	1671.40	30.36	298.77	3.47	141.37	2.18
MOD	LARGE	HALF	PPR	10.49	1.78	49.12	.28	37.31	.20
MOD	LARGE	HALF	LOESS	5.34	.53	27.63	.26	*	*
MOD	LARGE	HALF	AVAS	14.09	1.65	13.87	.67	*	*
MOD	LARGE	HALF	NN	80.32	5.59	107.00	1.52	*	*

Function: Product (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	SMALL	ALL	MLR	782.71	187.67	114.38	6.86	25.71	.28
HIGH	SMALL	ALL	SLR	971.83	179.85	153.10	6.91	26.01	.21
HIGH	SMALL	ALL	ACE	1305.65	210.77	560.80	40.88	15.69	.61
HIGH	SMALL	ALL	AM	782.31	187.22	114.44	6.85	25.71	.28
HIGH	SMALL	ALL	MARS	779.90	76.07	507.79	42.72	33.72	.89
HIGH	SMALL	ALL	RPR	7371.20	1669.68	2707.78	300.11	108.27	.32
HIGH	SMALL	ALL	PPR	1476.85	194.41	398.97	25.05	26.52	.45
HIGH	SMALL	ALL	LOESS	1448.58	223.28	263.06	18.32	*	*
HIGH	SMALL	ALL	AVAS	1295.49	174.28	465.34	34.34	14.34	.31
HIGH	SMALL	ALL	NN	2819.40	260.57	2687.31	61.70	*	*
HIGH	SMALL	HALF	MLR	698.86	175.55	129.25	7.33	41.45	.35
HIGH	SMALL	HALF	SLR	659.82	169.01	114.20	7.87	41.29	.27
HIGH	SMALL	HALF	ACE	1306.61	209.63	526.24	36.84	17.03	.56
HIGH	SMALL	HALF	AM	700.57	176.14	129.23	7.33	41.47	.36
HIGH	SMALL	HALF	MARS	980.79	99.23	334.45	30.12	22.57	.63
HIGH	SMALL	HALF	RPR	6120.10	1335.74	2310.32	273.92	191.22	4.08
HIGH	SMALL	HALF	PPR	1285.05	255.48	472.08	34.03	40.25	.37
HIGH	SMALL	HALF	LOESS	1487.55	227.92	262.61	18.38	*	*
HIGH	SMALL	HALF	AVAS	1243.75	181.84	483.22	36.49	19.47	.38
HIGH	SMALL	HALF	NN	3039.75	329.04	2657.71	59.48	*	*
HIGH	MED	ALL	MLR	273.57	37.01	71.79	3.58	24.39	.14
HIGH	MED	ALL	SLR	394.80	46.34	83.78	5.51	24.42	.15
HIGH	MED	ALL	ACE	846.36	109.96	224.10	10.84	7.30	.22
HIGH	MED	ALL	AM	273.52	36.96	71.77	3.59	24.40	.14
HIGH	MED	ALL	MARS	643.89	95.39	210.45	13.74	1577.79	.00
HIGH	MED	ALL	RPR	5893.80	843.44	662.97	154.52	108.25	.32
HIGH	MED	ALL	PPR	989.77	136.33	217.71	20.30	24.38	.19
HIGH	MED	ALL	LOESS	444.41	78.75	122.34	5.84	*	*
HIGH	MED	ALL	AVAS	971.14	133.62	192.52	10.93	11.36	.14
HIGH	MED	ALL	NN	2926.93	190.67	2648.25	38.39	*	*
HIGH	MED	HALF	MLR	212.10	34.48	85.65	3.56	40.12	.25
HIGH	MED	HALF	SLR	231.02	59.73	79.98	3.73	40.01	.24
HIGH	MED	HALF	ACE	816.30	136.00	209.48	7.45	8.41	.30
HIGH	MED	HALF	AM	211.16	34.40	85.64	3.57	40.13	.25
HIGH	MED	HALF	MARS	406.73	74.99	197.53	23.20	17.12	.21
HIGH	MED	HALF	RPR	6062.59	941.35	753.35	123.34	199.10	3.17
HIGH	MED	HALF	PPR	674.30	131.88	265.69	30.94	38.40	.24
HIGH	MED	HALF	LOESS	441.64	80.37	117.41	5.19	*	*
HIGH	MED	HALF	AVAS	847.22	128.02	182.39	8.70	16.26	.19
HIGH	MED	HALF	NN	2900.10	206.36	2666.66	45.04	*	*

Function: Product (Cont.)

Noise	n	Var. Expl.	Method	MISE $p=2$	St. Err.	MISE $p=6$	St. Err.	MISE $p=12$	St. Err.
HIGH	LARGE	ALL	MLR	136.84	11.75	49.04	.81	23.92	.10
HIGH	LARGE	ALL	SLR	152.20	17.55	49.77	1.34	23.98	.11
HIGH	LARGE	ALL	ACE	419.14	42.06	78.62	3.73	3.63	.11
HIGH	LARGE	ALL	AM	136.81	11.75	49.01	.79	23.92	.10
HIGH	LARGE	ALL	MARS	284.24	47.30	106.30	5.55	66.72	.22
HIGH	LARGE	ALL	RPR	4044.88	689.21	216.21	13.36	108.20	.31
HIGH	LARGE	ALL	PPR	372.50	54.94	130.13	15.93	23.49	.10
HIGH	LARGE	ALL	LOESS	140.80	13.47	58.81	2.25	*	*
HIGH	LARGE	ALL	AVAS	408.95	39.24	70.31	3.47	*	*
HIGH	LARGE	ALL	NN	2843.91	106.97	2660.98	25.75	*	*
HIGH	LARGE	HALF	MLR	84.07	13.20	63.21	.97	39.65	.22
HIGH	LARGE	HALF	SLR	70.99	13.89	60.32	1.10	39.64	.22
HIGH	LARGE	HALF	ACE	335.56	43.33	81.34	3.43	4.69	.13
HIGH	LARGE	HALF	AM	84.00	13.21	63.18	.95	39.66	.21
HIGH	LARGE	HALF	MARS	81.51	17.81	93.46	5.69	47.40	.17
HIGH	LARGE	HALF	RPR	4230.49	550.92	355.19	9.94	285.83	79.69
HIGH	LARGE	HALF	PPR	286.81	24.59	77.85	9.29	37.76	.20
HIGH	LARGE	HALF	LOESS	133.52	13.29	56.39	1.87	*	*
HIGH	LARGE	HALF	AVAS	341.80	34.96	69.51	4.01	*	*
HIGH	LARGE	HALF	NN	2902.25	95.07	2624.17	29.90	*	*

References

[1] BARRON, A.R. (1991). "Complexity Regularization with Applications to Artificial Neural Networks," in *Nonparametric Functional Estimation*, ed. by G. Roussas, Kluwer, Dordrecht, 561-576.

[2] BARRON, A.R. (1993). "Universal Approximation Bounds for Superpositions of a Sigmoidal Function," *IEEE Transactions on Information Theory*, Vol. 39, 930-945.

[3] BARRON, A.R. and BARRON, R.L. (1988). "Statistical Learning Networks: A Unifying View," in *Computing Science and Statistics: Proceedings of the Twentieth Symposium on the Interface*, ed. by E.J. Wegman, D.T. Gantz and J.J. Miller, 192-203, American Statistical Association, Alexandria, VA.

[4] BARRON, A.R. and XIAO, XIANGYU (1991). Discussion of "Multivariate Adaptive Regression Splines," by J.H. Friedman, *Annals of Statistics*, Vol. 19, 67-81.

[5] BREIMAN, L. (1991a). Discussion of "Multivariate Adaptive Regression Splines," by J.H. Friedman, *Annals of Statistics*, Vol. 19, 82-90.

[6] BREIMAN, L. (1991b). "The Π Method for Estimating Multivariate Functions From Noisy Data," *Technometrics*, Vol. 33, 124-144.

[7] BREIMAN, L., FRIEDMAN, J., OLSHEN, R.A. and STONE, C. (1984). *Classification and Regression Trees*, Wadsworth, Belmont, CA.

[8] BREIMAN, L. and FRIEDMAN, J. (1985). "Estimating optimal transformations for multiple regression and correlation" (with discussion). *Journal of the American Statistical Association* 80, 580-619.

[9] BUJA, A., HASTIE, T. J. and TIBSHIRANI, R. (1989). "Linear Smoothers and Additive Models," (with discussion) *Annals of Statistics*, Vol. 17, 453-555.

[10] BUJA, A. and KASS, R. (1985). Discussion of "Estimating optimal transformations for multiple regression and correlation," by Breiman and Friedman, *Journal of the American Statistical Association* 80, 602-607.

[11] CLEVELAND, W. (1979). "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, Vol. 74, 829-836.

[12] CLEVELAND, W. and DEVLIN, S. (1988). "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association*, Vol. 83, 596-610.

[13] DE VEAUX, R. D., PSICHOGIOS, D. C., and UNGAR, L. H. (1993). "A Comparison of Two Nonparametric Estimation Schemes: MARS and Neural Networks," *Computers in Chemical Engineering*, **17**, 819-837.

[14] DONOHO, D.L. and JOHNSTONE, I. (1989). "Projection Based Approximation and a Duality with Kernel Methods," *Annals of Statistics*, Vol. 17, 58-106.

[15] DOERING, A., GALICKI, M., and WITTE, H. (1997). "Admissibility and Optimality of the Cascade-Correlation Algorithm." In *Artificial Neural Networks: Proceedings of the Seventh International Conference*, W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, eds. Springer-Verlag, Berlin, 505-510.

[16] FAHLMAN, S. E. and LEBIERE, C. (1990). The Cascade-Correlation Learning Architecture. In *Advance in Neural Information Processing Systems 2*, ed. by D. Touretzky. Morgan Kaufman, San Mateo, CA, pp. 525-533.

[17] FLURY, B. (1988). *Common Principal Components and Related Multivariate Models*, John Wiley and Sons, Inc., NY.

[18] FRANK, I. and FRIEDMAN, J.H. (1993). "A Statistical View of Some Chemometric Regression Tools," (with discussion), *Technometrics*, **35**, 109-148.

[19] FRIEDMAN, J.H. (1991a). "Multivariate additive regression splines". *Annals of Statistics*, Vol. 19, 1-66.

[20] FRIEDMAN, J.H. (1991b). Discussion of "The Π Method for Estimating Multivariate Functions From Noisy Data," by L. Breiman, in *Technometrics*, Vol. 33, 145-148.

[21] FRIEDMAN, J.H. and STUETZLE, W. (1981). "Projection pursuit regression". *Journal of the American Statistical Association*, Vol. 76, 817-23.

[22] GU, C. (1991). Discussion of "The Π Method for Estimating Multivariate Functions From Noisy Data," by L. Breiman, in *Technometrics*, Vol. 33, 149-154.

[23] HASTIE, T.J. and TIBSHIRANI, R.J. (1990). *Generalized Additive Models*, Chapman and Hall, New York.

[24] RIPLEY, B. D. (1993). Statistical aspects of neural networks. In *Networks and Chaos—Statistical and Probabilistic Aspects*, O. E. Barndorff-Nielsen, J. L. Jensen, and W. S. Kendall, eds. Chapman and Hall, London, pp. 40-123.

[25] RIPLEY, B. D. (1994a). Flexible non-linear approaches to classification. In *From Statistics to Neural Networks. Theory and Pattern Recognition Applications.* V. Cherkassky, J. H. Friedman, and H. Wechsler, eds. Springer-Verlag, New York, NY.

[26] RIPLEY, B. D. (1994b). Neural networks and related methods for classification (with discussion). *Journal of the Royal Statistical Society, Series B*, **56**, 409-456.

[27] RIPLEY, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press.

[28] SCOTT, D.W. (1992). *Multivariate Density Estimation*. Wiley, New York.

[29] SCOTT, D.W. and WAND, M.P. (1991). "Feasibility of Multivariate Density Estimates," *Biometrika*, Vol. 78, 197-206.

[30] SUTHERLAND, A, HENERY, R., MOLINA, R., TAYLOR, C. C., and KING, R. (1993). "Statistical Methods in Learning." In *IPMU '92—Advanced Methods in Artificial Intelligence*, B. Bouchon-Meunier, L. Valverde, and R. R. Yager, eds. Springer-Verlag, Berlin, 173-182.

[31] TIBSHIRANI, R. (1988). "Estimating optimal transformations for regression via additivity and variance stabilization". *Journal of the American Statistical Association*, Vol. 83, 394-405.

[32] ZHAO, Y and ATKESON, C. G. (1992). Some approximation properties of projection pursuit networks. In *Advances in Neural Information Processing Systems 4*, J. Moody, S. J. Hanson, and R. P. Lippmann, eds. Morgan Kaufmann, San Mateo, CA, pp. 936-943.