Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computational Statistics & Data Analysis 48 (2005) 809–819

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Hybrid local polynomial wavelet shrinkage: wavelet regression with automatic boundary adjustment

Hee-Seok Oh[a], Thomas C.M. Lee[b,*]

[a]*Department of Statistics, Seoul National University, Seoul 151-742, South Korea*
[b]*Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, USA*

Received 4 February 2004; received in revised form 2 April 2004; accepted 3 April 2004

Abstract

An usual assumption underlying the use of wavelet shrinkage is that the regression function is assumed to be either periodic or symmetric. However, such an assumption is not always realistic. This paper proposes an effective method for correcting the boundary bias introduced by the inappropriateness of such periodic or symmetric assumption. The idea is to combine wavelet shrinkage with local polynomial regression, where the latter regression technique is known to possess excellent boundary properties. Simulation results from both the univariate and bivariate settings provide strong evidence that the proposed method is extremely effective in terms of correcting boundary bias.

Keywords: Back-fitting; Boundary problem; Local polynomial regression; Polynomial wavelet regression; Wavelet shrinkage

1. Introduction

Consider the following regression model:

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

* Corresponding author. Tel.: +1-970-4912185; fax: +1-970-4917895.
E-mail address: tlee@stat.colostate.edu (T.C.M. Lee).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2004.04.002

where $x_i = (i-1)/n$ and the ε_i's are independent and identical $N(0, \sigma^2)$ random errors. The unknown function f, mostly smooth, is suspected to have a few discontinuities or abrupt changes. Under this situation one popular method to estimate f is wavelet shrinkage; see, for examples, the seminal papers Donoho and Johnstone (1994, 1995).

However, wavelet shrinkage suffers from boundary problems that are caused by the application of the wavelet transformation to a finite signal. Many approaches have been proposed to overcome these problems. Perhaps the most popular approach is to impose some additional constraints, such as periodicity or symmetry, on f. This approach can be easily implemented, but such additional constraints may not always be realistic, especially so for two-dimensional data such as images. More recently, Oh et al. (2001) (see also Lee and Oh, 2003; Naveau and Oh, 2003) propose a simple method called polynomial wavelet regression (PWR) for handling these boundary problems. The idea of PWR is to estimate f with the sum of a set of wavelet basis functions, $\hat{f}_W$, and a low-order (global) polynomial, $\hat{f}_P$. That is,

$$\hat{f}_{PW}(x) = \hat{f}_P(x) + \hat{f}_W(x), \tag{2}$$

where $\hat{f}_{PW}$ is the PWR estimate for f. The hope is that, once $\hat{f}_P$ is removed from the data y_i, the remaining signal hidden in $y_i - \hat{f}_P(x_i)$ can be well estimated using wavelet regression with say periodic boundary assumption. In practice, one needs to determine the order of the polynomial for $\hat{f}_P$. Simulation results from Lee and Oh (2003) suggest the use of BIC (Schwarz, 1978).

The use of PWR for resolving boundary problems works very well if $\hat{f}_P$ is able to remove the "non-periodicity" in y_i. However, due to the global nature of $\hat{f}_P$, for those cases when f has complex boundary conditions or has some abrupt changing objects present near the boundaries, PWR does not work well. The goal of this article is to proposal a new method which will also work well under these situations. Fig. 1 provides some illustrative examples. The left column displays various curve estimates for a regression function that can be well estimated using PWR. It can be seen that both the PWR and the proposed method (to be described in the next section) produce good estimates, while a classical wavelet regression estimate with periodic assumption suffers from edge effects. On the other hand, the right column presents a situation that both the PWR and classical wavelet regression fail at the boundaries, but the proposed method is still able to procedure good boundary estimates.

The basic idea behind the proposed method is to introduce a local polynomial regression component to the wavelet shrinkage. Since local polynomial regression produces excellent boundary handling (Fan, 1992; Hastie and Loader, 1993), it is expected that the addition of this component to wavelet shrinkage will result in equally well boundary properties. Results from numerical experiments strongly support this claim. Besides producing promising empirical results, other desirable properties of the proposed method include: it is easy to implement, computationally fast, and can be straightforwardly extended to higher dimensional settings.

The rest of this article is organized as follows. Section 2 presents the proposed method. Results from a simulation study are reported in Section 3. In Section 4 the two-dimensional setting is considered. Conclusion is offered in Section 5.

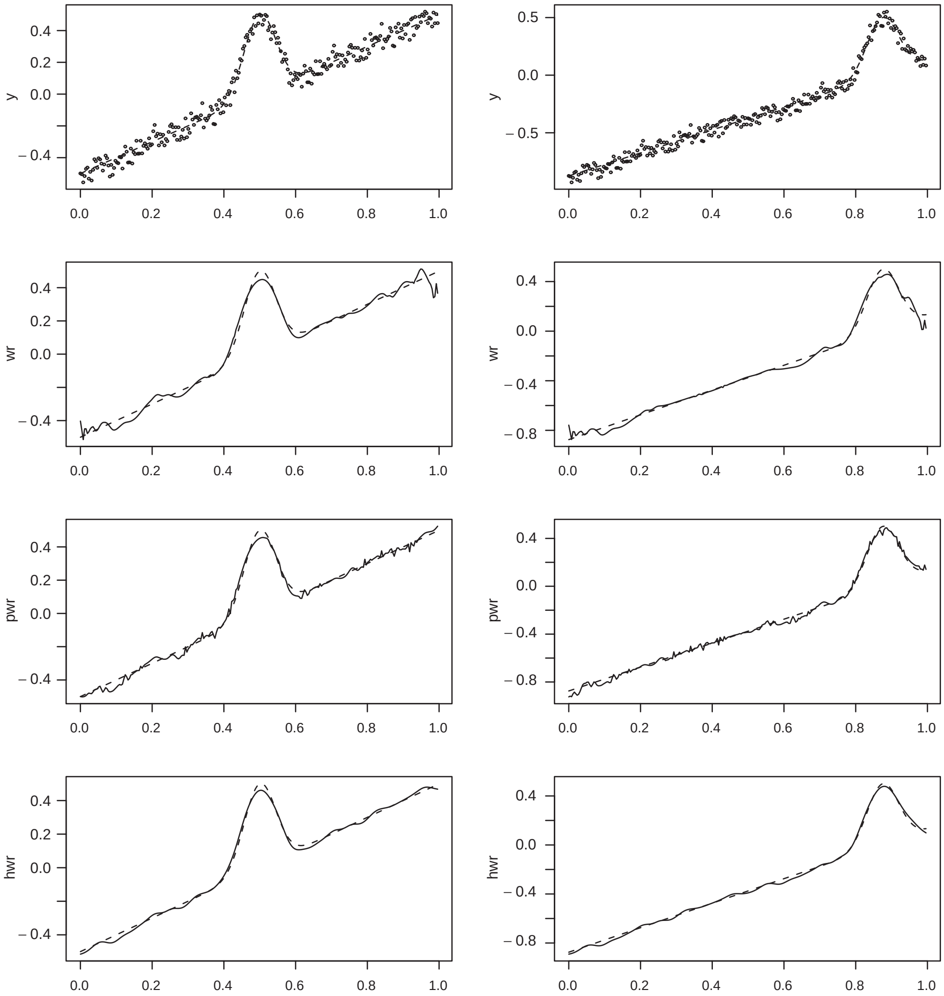

Fig. 1. Plots of two test functions (dashed lines) together with different curve estimates (solid lines). First row: noisy data sets; second row: classical wavelet regression estimates with periodic boundary conditions; third row: polynomial wavelet regression estimates; fourth row: the proposed hybrid local polynomial wavelet regression estimates.

2. Hybrid local polynomial wavelet shrinkage

This section presents the proposed method for improving boundary adjustment in wavelet regression. Driven by the fact that local polynomial regression is extremely effective in adapting to boundary conditions (e.g., see Fan, 1992; Hastie and Loader, 1993), we propose replacing the global polynomial fit $\hat{f}_P$ in PWR with a local polynomial fit $\hat{f}_{\mathrm{LP}}$. We call the new resulting curve estimate $\hat{f}_{\mathrm{H}}$ a hybrid estimate:

$$\hat{f}_{\mathrm{H}}(x) = \hat{f}_{\mathrm{LP}}(x) + \hat{f}_{\mathrm{W}}(x). \tag{3}$$

Now the expectation is that, the use of local polynomial regression should automatically provide boundary bias correction when applied with wavelet shrinkage.

To obtain the hybrid estimate $\hat{f}_{\mathrm{H}}$, we need to estimate the two components: $\hat{f}_{\mathrm{LP}}(x)$ and $\hat{f}_{\mathrm{W}}(x)$. Inspired by the back-fitting algorithm of Hastie and Tibshirani (1990), we propose the following iterative algorithm for computing $\hat{f}_{\mathrm{LP}}(x)$, $\hat{f}_{\mathrm{W}}(x)$ and hence $\hat{f}_{\mathrm{H}}$.

1. Obtain an initial estimate $\hat{f}^0$ for f, and set $\hat{f}^0_{\mathrm{LP}} = \hat{f}^0$.
2. For $j = 1, \ldots,$ iterate the following steps:
   (a) Apply wavelet shrinkage to $y_i - \hat{f}^{j-1}_{\mathrm{LP}}$ and obtain $\hat{f}^j_{\mathrm{W}}$.
   (b) Estimate $\hat{f}^j_{\mathrm{LP}}$ by fitting local polynomial regression to $y_i - \hat{f}^j_{\mathrm{W}}$.
3. Stop if $\hat{f}_{\mathrm{H}} = \hat{f}^j_{\mathrm{LP}} + \hat{f}^j_{\mathrm{W}}$ converges.

To use the above algorithm, one needs to choose the initial curve estimate $\hat{f}^0$ in Step 1 and the smoothing parameter for the local polynomial fit $\hat{f}^j_{\mathrm{LP}}$ in Step 2(b). For computing $\hat{f}^0$, we use Friedman's supersmoother (available as `supsmu` in *R* or *S*-Plus), while for the smoothing parameter for computing $\hat{f}^j_{\mathrm{LP}}$, we use cross-validation.

Our numerical experiences suggest that the above algorithm converges very quickly. Note that the algorithm can be straightforwardly extended to higher dimensional problems, in which the proportion of "boundary regions" is much higher then one-dimensional curve fitting problems.

3. Simulation study

This section reports results from a simulation study that was designed to assess the practical performance of the proposed method. The *R*-codes used to carry out this simulation are available from the website http://www.stat.ualberta.ca/~heeseok/hybrid.html

3.1. Experimental setup

In this simulation the following three methods are studied:

1. `wr`: the classical wavelet shrinkage with periodic boundary correction,
2. `pwr`: the polynomial wavelet regression developed by Oh et al. (2001) and Lee and Oh (2003),
3. `hwr`: the hybrid wavelet shrinkage proposed in Section 2.

Throughout the whole simulation and for all the above three methods, the empirical Bayes procedure `EBayesThresh` of Johnstone and Silverman (2003) was used as the thresholding rule.

Table 1
Formulae of the test functions. All have the same domain $x \in [0, 1]$

Test function	Formula	
1-1	$\frac{1}{4}[(4x-0.3)+2\exp\{-16(4x-0.3)^2\}]$	
1-2	$\frac{1}{4}[(4x-2.0)+2\exp\{-16(4x-2.0)^2\}]$	
1-3	$\frac{1}{4}[(4x-3.5)+2\exp\{-16(4x-3.5)^2\}]$	
2-1	$\frac{1}{3}\sin(4\pi x)-\frac{1}{12}\operatorname{sgn}(x-0.3)-\frac{1}{12}\operatorname{sgn}(0.72-x)$	
2-2	$\frac{1}{3}\sin(4\pi x)-\frac{1}{12}\operatorname{sgn}(x-0.3)-\frac{1}{12}\operatorname{sgn}(0.90-x)$	
2-3	$\frac{1}{3}\sin(4\pi x)-\frac{1}{12}\operatorname{sgn}(x-0.3)-\frac{1}{12}\operatorname{sgn}(0.95-x)$	
3-1	$4x(1-\sin x)$	$x \in [0, 0.47) \cup (0.52, 1]$
	0.625	$x \in [0.47, 0.52]$
3-2	$4x(1-\sin x)$	$x \in [0, 0.73) \cup (0.78, 1]$
	0.625	$x \in [0.73, 0.78]$
3-3	$4x(1-\sin x)$	$x \in [0, 0.88) \cup (0.93, 1]$
	0.625	$x \in [0.88, 0.93]$

We considered three test functions. These test functions have been used by previous authors and their descriptions can be found in, respectively, Fan and Gijbels (1995), Donoho and Johnstone (1994) and Oh et al. (2001). Each function has some abrupt changing features such as discontinuities or sharp bumps. For each of these three test functions, we produced two additional variants by shifting the locations of these abrupt changing features. Therefore 9 test functions were used in total. They are listed in Table 1 and displayed in Fig. 2. Note that it is reasonable to assume a periodic boundary condition for Test Function 2, while for Test Functions 1 and 3, some boundary adjustment is strongly preferred. As mentioned previously, it is not expected that the addition of a low-order global polynomial term will eliminate boundary bias for those cases that there are abrupt changing features near the boundaries, such as those in Test Functions 1-1, 1-3 and 3-3.

Two levels of signal-to-noise ratio (snr) were chosen: snr $= 8$ and 6, where snr is defined as snr$=\|\boldsymbol{f}\|/\sigma$. We considered two different sample sizes: $n=512$ and 256. For each combination of test function, snr and n, 100 sets of observations contaminated by Gaussian noise were simulated. For each simulated data set, the above three methods were applied to estimate the test function.

The mean-squared error (MSE) was used as the numerical measure for assessing the quality of an estimate $\hat{f}$: $\mathrm{MSE}(\hat{f})=n^{-1}\sum_{i=1}^{n}\{f(x_i)-\hat{f}(x_i)\}^2$. For those experimental settings with snr $= 8$ and $n = 256$, boxplots of the MSEs for all the curve estimates are given in Fig. 3. Results for the remaining experimental combinations were similar and hence are omitted.

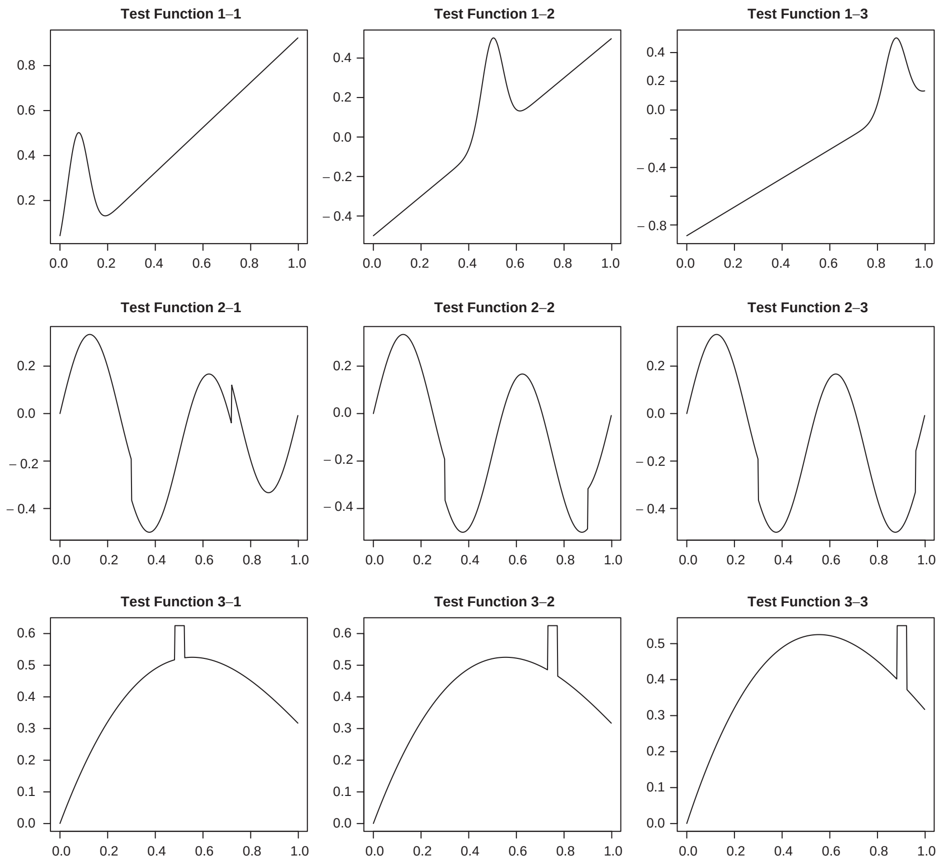

Fig. 2. The nine test functions used in the univariate simulation.

To gain an idea of how the three methods perform near the boundaries, we calculated the following MSE values for observations near the boundaries

$$\mathrm{MSE}_{\Delta}(\hat{f}) = \frac{1}{2\Delta} \sum_{i \in \mathcal{N}(\Delta)} \{f(x_i) - \hat{f}(x_i)\}^2 \quad (\Delta = 1, 2, \ldots, [n/2];\ x_i = i/n),$$

where $\mathcal{N}(\Delta) = \{1, \ldots, \Delta, n - \Delta + 1, \ldots, n\}$. Fig. 4 displays the averaged values (over 100 replicates) of $\mathrm{MSE}_{\Delta}(\hat{f})$ versus Δ for Test Functions 1-3 and 3-3, with $\mathrm{snr} = 8$ and $n = 512$.

3.2. Results

From the simulation results, the following empirical observations can be made:

1. When the periodic boundary assumption is satisfied, both `hwr` and `wr` gave very similar results and outperformed `pwr` (Test Function 2).

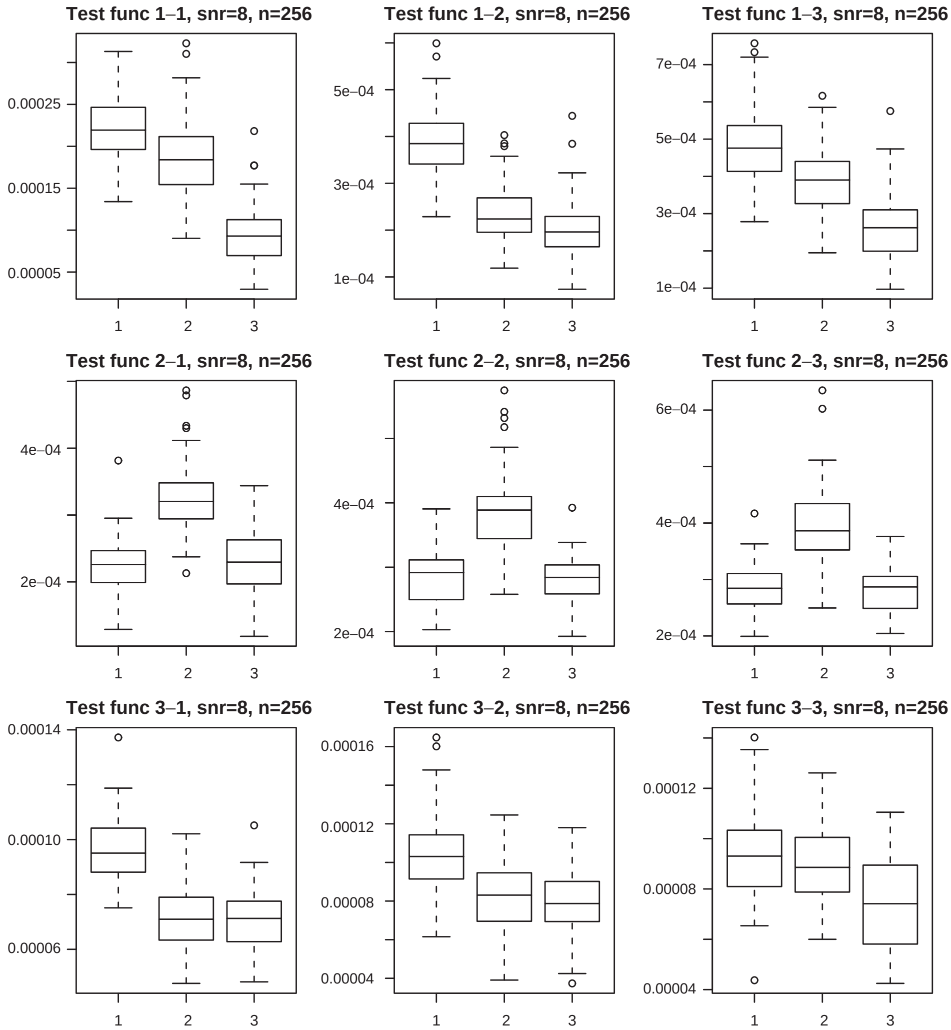

Fig. 3. Boxplots of MSE($\hat{f}$) from the univariate simulation study, with snr $= 8$ and $n = 256$.

2. When the periodic boundary assumption is *not* satisfied, both `hwr` and `pwr` outperformed `wr` (Test Functions 1 and 3).
3. When the periodic boundary assumption is *not* satisfied and when the abrupt changing features are located *near* the boundaries, `hwr` outperformed `pwr`. On the other hand, if the abrupt changing features are far away from the boundaries, both `hwr` and `pwr` gave similar results.

Therefore, the simulation results seem to suggest that hybrid local polynomial wavelet shrinkage is a very preferable method for handling boundary problems for wavelet regression.

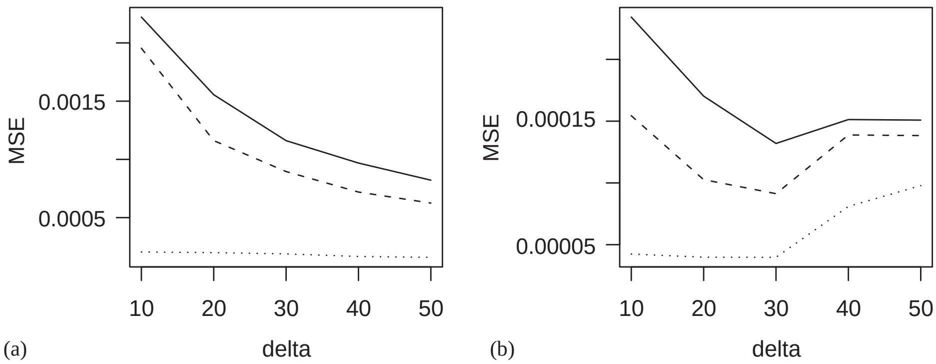

Fig. 4. The averaged values of $\mathrm{MSE}_{\Delta}(\hat{f})$ versus Δ. The solid line is for `wr`, the broken line is for `pwr` and the dotted line is for `hwr`. (a) Test Function 1-3; (b) Test Function 3-3.

4. Two-dimensional fitting

In higher dimensional problems, such as image denoising, the proportion of boundary observations is much higher than univariate curve estimation problems. Thus the need for boundary adjustment is even stronger. The above proposed method can be extended straightforwardly to such higher dimensional settings. This section present some simulation results obtained from two-dimensional surface fittings.

Displayed in Fig. 5 are the five two-dimensional test functions used in this study. These five functions were constructed by adding two rectangular blocks to the five test functions used in Hwang et al. (1994). To be specific, let $f(x, y)$ denote any one of the five functions of Hwang et al. (1994). Then the corresponding new test function $g(x, y)$ used in the current simulation is given by

$$g(x,y) = \begin{cases} 3 & \text{if } \frac{1}{8} < x < \frac{1}{4}, \frac{1}{16} < y < \frac{3}{16}, \\ 5 & \text{if } \frac{1}{16} < x < \frac{3}{16}, \frac{3}{4} < y < \frac{7}{8}, \\ f(x,y) & \text{otherwise.} \end{cases}$$

The observations were generated by adding independent Gaussian noise to each of the test functions sampled on a square grid of 256×256 regularly spaced grid points. For each test function the number of replicates was 100 and snr was set to 6.

For each of the simulated data sets, two surface estimates were obtained. The first one was obtained by applying a classical two-dimensional wavelet thresholding estimation method, with a thresholding value $0.3\hat{\sigma}\sqrt{2\log 256^2}$; i.e. the universal thresholding rule of Donoho and Johnstone (1994) multiplied by 0.3. This multiplier of 0.3 was suggested by Kovac and Silverman (2000), and $\hat{\sigma}$ is a robust estimate of the Gaussian noise standard deviation. The second surface estimate were obtained with the proposed hybrid method. The same thresholding value was used and the smoothing parameter for the local linear regression was chosen by cross-validation.

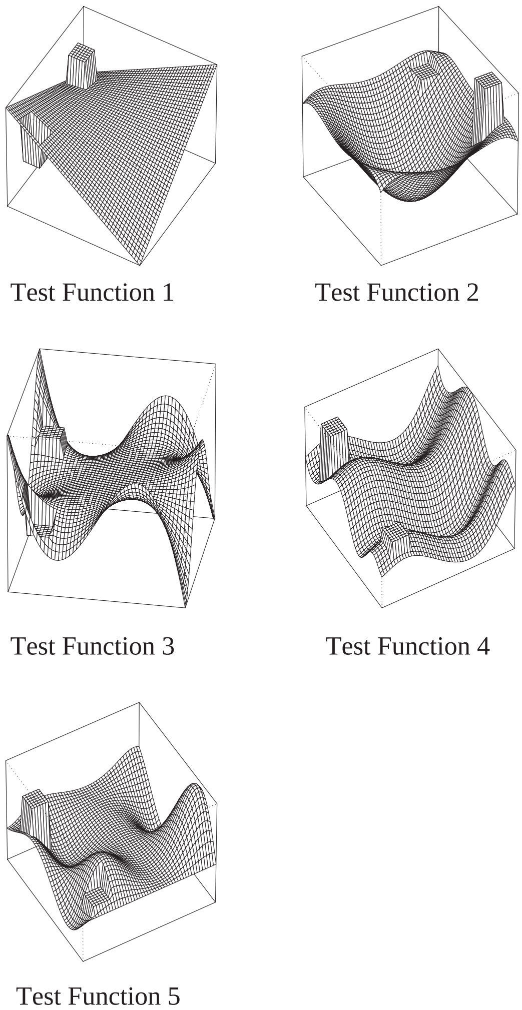

Fig. 5. Perspective plots of two-dimensional test functions.

Two MSE values were computed for each estimated surface. The first one was calculated over the whole domain of the test functions, while the second one was calculated over those observations that were at most 5 sample points away from the edges of the test functions. The following ratio was then computed for both types of MSE values:

$$\frac{\text{MSE for classical estimate}}{\text{MSE for hybrid estimate}}.$$

Boxplots of the logarithms of these two types of MSE ratios are given in Fig. 6. From these plots, one can see that the hybrid method always outperformed the classical method. The improvement was even more substantial near the boundaries.

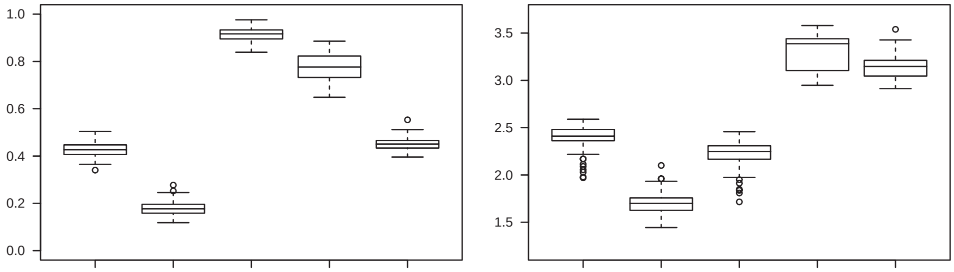

Fig. 6. Boxplots of the log of two MSE ratios. Left: for MSE values calculated over the entire domain of the test functions; right: for MSE values calculated over the boundary regions of the test functions. In each panel the five boxplots, from left to right, correspond respectively to the 2D Test Functions 1–5.

5. Conclusions

In this paper a hybrid wavelet shrinkage method is proposed for reducing the boundary bias that is commonly found in wavelet shrinkage. The proposed method is based on a coupling of classical wavelet shrinkage and local polynomial regression. The empirical performance of the method was tested on different numerical experiments, including both the univariate and bivariate settings. Results from these experiments illustrate the improvement of the hybrid wavelet shrinkage over the classical wavelet shrinkage methods.

Acknowledgements

The authors are grateful to the reviewer for his/her comments which lead to a more succinct version of the paper. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada and the National Science Foundation under Grant No. 0203901.

References

Donoho, D.L., Johnstone, I.M., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika 81, 425–455.

Donoho, D.L., Johnstone, I.M., 1995. Adapting to unknown smoothness via wavelet shrinkage. J. Amer. Statist. Assoc. 90, 1200–1224.

Fan, J., 1992. Design-adaptive nonparametric regression. J. Amer. Statist. Assoc. 87, 998–1004.

Fan, J., Gijbels, I., 1995. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. J. Roy. Statist. Soc. Ser. B 57, 371–394.

Hastie, T., Loader, C., 1993. Local regression: automatic kernel carpentry (with discussion). Statist. Sci. 8, 120–143.

Hastie, T., Tibshirani, R., 1990. Generalized Additive Models. Chapman & Hall, London.

Hwang, J.-N., Lay, S.-R., Maechler, M., Martin, R.D., Schimert, J., 1994. Regression modeling in back-propagation and projection pursuit learning. IEEE Trans. Neural Networks 5, 342–353.

Johnstone, I.M., Silverman, B.W., 2003. Empirical Bayes selection of wavelet thresholds, unpublished manuscript.

Kovac, A., Silverman, B.W., 2000. Extending the scope of wavelet regression methods by coefficient-dependent thresholding. J. Amer. Statist. Assoc. 95, 172–183.

Lee, T.C.M., Oh, H.-S., 2003. Automatic polynomial wavelet regression, unpublished manuscript.

Naveau, P., Oh, H.-S., 2003. Polynomial wavelet regression for images with irregular boundaries. IEEE Trans. Image Process., to appear in June 2004.

Oh, H.-S., Naveau, P., Lee, G., 2001. Polynomial boundary treatment for wavelet regression. Biometrika 88, 291–298.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statist. 6, 461–464.