A Comparison of Some Thresholding Selection Methods for Wavelet Regression

Alsaidi M. Altaher, Mohd T. Ismail

***Abstract*—**In wavelet regression, choosing threshold value is a crucial issue. A too large value cuts too many coefficients resulting in over smoothing. Conversely, a too small threshold value allows many coefficients to be included in reconstruction, giving a wiggly estimate which result in under smoothing. However, the proper choice of threshold can be considered as a careful balance of these principles. This paper gives a very brief introduction to some thresholding selection methods. These methods include: Universal, Sure, Ebays, Two fold cross validation and level dependent cross validation. A simulation study on a variety of sample sizes, test functions, signal-to-noise ratios is conducted to compare their numerical performances using three different noise structures. For Gaussian noise, EBayes outperforms in all cases for all used functions while Two fold cross validation provides the best results in the case of long tail noise. For large values of signal-to-noise ratios, level dependent cross validation works well under correlated noises case. As expected, increasing both sample size and level of signal to noise ratio, increases estimation efficiency.

Keywords— wavelet regression, simulation, Threshold.

I. INTRODUCTION

ESTIMATING a regression function using wavelet methods is an issue that has received great attention over the last two decades. Many theorems and methods are introduced with emphasis on problems of the choice of the smoothing parameter. The basic idea behind wavelet estimation is to get a relatively small number of wavelet coefficients to represent the underling regression function. A value called (Threshold) is used to kill or keep the wavelet coefficient. Hence, Estimation quality depends strongly on how efficient threshold value would be chosen.

Many different schemes have been proposed for choosing a threshold value begging by Donoho and Johnstone [6, 7], Nason[14], Barber and Nason [3], Johnstone and Silverman [10, 11], Oh, Kim and Lee [12], Silverman [19], Kim and Lee[13]. However, Abramovich et al. [2] and Vidakovic [22] give a review of some of these. In this paper, it would be given a brief description of some of the standard thresholding rules.

As usual, whenever a method is created or developed, a series of investigations comes along with the new development. The main goal of this paper is to investigate and compare the latest methods with previous ones. The remainder of this paper is organized as follows:

Section 2 gives a brief review to wavelet regression and thresholding rules including soft and hard threshold. Then, a brief introduction to: universal, Sure, Ebays, Two fold cross validation and level dependent cross validation. Section 4 examines simulations to investigate the practical performance of the methods mentioned above. Section 5 provides the conclusion of this work.

II. WAVELET REGRESSION

Suppose a data set $y_1, y_2, ..., y_n$ observed from the model:

$$y_i = f(x_i) + \varepsilon_i \quad , i = 1,2,...,n = 2^j \tag{1}$$

Where $\{\varepsilon_i\}$ are $iid\, N(0,\sigma^2)$, $x_i = \frac{i}{n}$, and f is the function to be estimated.

Wavelet estimation of model (1) can be performed in three steps: first, take the discreet wavelet transform of y_i. Next, a "soft" or a "hard" thresholding rule is used to threshold the coefficients. Finally the coefficients are inversely transformed back to the signal space to obtain the estimated $\hat{f}$.

Given a wavelet coefficient w and threshold value λ, the hard threshold value of the coefficient can be written as:

$$\eta_{hard}(w,\lambda) = w\, I(\ |w| > \lambda)$$

while the soft threshold value is

$$\eta_{soft}(w,\lambda) = \operatorname{sgn}(w)\ (\ |w| - \lambda)\ I(\ |w| > \lambda)$$

where I is the usual indicator function. In other words, "hard" means "keep or kill" while "soft" means "shrink or kill".

III. THRESHOLDING SELECTION METHODS

This section is devoted to introduce theoretical descriptions of some thresholding selection methods.

A. *Universal Thresholding Methods*

The universal threshold method was introduced by Donoho and Johnstone [6]. It is given by

Alsaidi M. Altaher is with the Mathematics School, University Sins Malaysia, Penang, 11800, phone: 0060174087402;
(E-mail: assaedi76@yahoo.com).
Mohd T. Ismail is with the Mathematics School, University Sins Malaysia, Penang, 11800, phone: 0060164143464;
(E-mail: mtahir@cs.usm.my)

$$\lambda_{universal} = \sigma\sqrt{2\log(n)}$$

where n is the total number of data points (equivalently the number of wavelet coefficients) and σ is the standard deviation of noise level which almost is unknown and it is usually replaced by a robust estimate, $\hat{\sigma}$ such as the median absolute deviation of the wavelet coefficients at the finest level ($j = \log(n) - 1$)

$$\hat{\sigma} = median(\left|w_{j-1,k} - median(w_{j-1,k})\right| \qquad (2)$$

Using the universal threshold yields the largest thresholds and thus an estimate of regression function with a relatively high degree of smoothing.

B. *Sure Thresholding Method*

This method was introduced by Donoho and Johnstone [7] which was achieved by the principle of minimizing the Stein' Unbiased Risk Estimation (SURE) for each wavelet level j .

Let $\mathrm{X} \sim N_p(\mu, \mathrm{I})$ be multivariate Gaussian observations with mean vector μ and diagonal covariance matrix I .
Stein [19] showed that if:

$$\hat{\mu}(\mathrm{X}) = \mathrm{X} + g(\mathrm{X})$$

Where $\hat{\mu}(\mathrm{X})$ is a particular fixed estimator of μ and $g = (g_i)_{i=1}^{p}$ is a function from R^p into R^p which is assumed to be weakly differentiable, then

$$E_\mu \left\|\hat{\mu}(\mathrm{X} - \mu)\right\| = p + E_\mu\{\left\|\mathrm{g}(\mathrm{X})\right\|^2 + 2\nabla.g(\mathrm{X})\}$$

Where

$$\nabla.g = \sum_{i=1}^{p} \frac{\partial}{\partial x_i} g_i$$

The insight of Donoho and Johnstone [7] was to apply Stein's result in [20] using a soft threshold.

In this case:

$$g(x_i) = \begin{cases} -x_i & \text{if} \quad |x_i| \le \lambda \\ -\lambda & \text{if} \quad x_i > \lambda \\ \lambda & \text{if} \quad x_i < \lambda \end{cases}$$

Then:

$$SURE(\lambda_J, w_{jk}) = p - 2.\#\{i : |x_i| \le \lambda\} + \sum_{i=1}^{p} (|x_i| \wedge \lambda)^2$$

is an unbiased estimate of the risk. This means:

$$E_\mu \left\|\hat{\mu}^{(\lambda)}(\mathrm{X}) - \mu\right\| = E_\mu SURE(\lambda, x)$$

So the SURE threshold can be written as:

$$\lambda_{j,sure} = \arg_{0 \le \lambda \le \sqrt{2\log n}} \min(\lambda_J, w_{jk})$$

C. *Two fold Cross Validation*

Cross validation is a popularly used method in a wild range of statistical procedure; see for example Stone [21], Silverman [18], Green and Silverman [9].

Since the fast wavelet transform methods require input data vectors that are of length $n = 2^j$, leaving out a data point makes the data length is no longer a power of two and thus classic cross validation cannot be done to estimate regression function. For this, Nason [14] proposed dropping half of the data points which whose size is still a power of two. Here is a description of Nason two fold crosses validation.

Let $y_1^o, y_2^o, \ldots, y_{n/2}^o$ represent the odd data points and $y_1^E, y_2^E, \ldots, y_{n/2}^E$ represent the even data points.
Let $\hat{f}^o$, $\hat{f}^E$ denote the wavelet estimators based on the odd index points and even index points respectively. Using the removed odd indexed data, an interpolated version of the odd noise data is formed:

$$\tilde{y}_i^o = \begin{cases} \frac{1}{2}(y_{2i-1} + y_{2i+1}) & , i = 1, 2, \ldots \frac{n}{2}\text{-}1 \\ \frac{1}{2}(y_{n-1} + y_1) & , i = \frac{n}{2} \end{cases}$$

For the even data noise, let:

$$\tilde{y}_i^E = \begin{cases} \frac{1}{2}(y_{2i-2} + y_{2i}) & , i = 2, \ldots \frac{n}{2} \\ \frac{1}{2}(y_n + y_2) & , i = 1 \end{cases}$$

The full cross validation estimate for the risk $M(\lambda)$ is:

$$M(\lambda) = \sum\{(\hat{f}_{\lambda,j}^{E}(\tfrac{2i}{n}) - \tilde{y}_i^o)^2 + (\hat{f}_{\lambda,j}^{o}(\tfrac{2i-1}{n}) - \tilde{y}_i^E)^2\}$$

If $\lambda_{\frac{n}{2}}$ minimizes $M(\lambda)$ then the final threshold is given by

$$\lambda_n = (1-\frac{\log 2}{\log n})^{-1/2}\lambda_{\frac{n}{2}}$$

which can be computed numerically .

D. Level Dependent Cross Validation

Various existing methods for level dependent cross validation have been developed by Donoho and Johnstone [7], Johnstone and Silverman [10]. Oh, Kim and Lee [13] proposed a new method for level dependent cross validation in which a data point is imputed rather than expelling data. A fast imputation method is used to obtain the CV wavelet estimation $\hat{f}_\lambda^{-i}(x_i)$ when the i^{th} observation is deleted. Cross validation estimator is given by

$$CV(\lambda) = \frac{1}{n}\sum_{i=1}^{n}[y_i - f_\lambda^{-i}(x_i)]^2 \qquad (3)$$

Two algorithms can be used to obtain $\hat{f}_\lambda^{-i}(x_i)$: "Leave one out CV" or "k-fold CV". Suppose dataset $\{x_i, y_i\}_{i=1}^{n}$. To apply" Leave one out algorithm" first, remove the data point (x_i, y_i) to be considered as a test dataset. This will leave $y_{obs} = \{y_1, y_2, ..., y_{i-1}, y_{i+1}, ..., y_n\}$ as a training dataset. For y_{obs} , Mallat's fast algorithm cannot be used since the new data length is no longer a power of 2. To overcome this limitation, Oh, Kim and Lee [13] suggested inputting a data point $\tilde{y}_i$ using an iterative imputation procedure to get a new training dataset $y_{new,i} = \{y_{obs}, \tilde{y}_i\}$ with length a power of 2.

Now, for a given threshold value λ , wavelet estimate $\hat{f}_\lambda(x_i)$ at every design point x_i must be found and then, $\hat{f}_\lambda^{-i}(x_i)$ is evaluated. Finally, $CV(\lambda)$ is computed over a certain range for λ , thus, level dependent cross validation threshold $\hat{\lambda}$ that minimize (3).

Traditional "K-fold cross validation "can also be used to obtain $\hat{f}_\lambda^{-i}(x_i)$ .Suppose data sets with size n . Divide the dataset into M blocks, where each block has n/m size. Then, "leave m blocks out CV)" can be performed by dropping m blocks as test data. In general, leave m blocks-out with block size b is as if we perform a k-fold CV with block size b , where $k = \frac{n}{m \times b}$.

Oh, Kim and Lee [13] showed that level dependent thresholds according to levels are obtained by minimizing the level dependent cross validation:

$$(\lambda_1, \lambda_2, ..., \lambda_J) = \arg\min CV(\lambda_1, \lambda_2, ..., \lambda_J)$$

Where

$$CV(\lambda_1, \lambda_2, ..., \lambda_J) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_{\lambda_1,\lambda_2,...,\lambda_j}^{-k(i)})^2$$

Where $f_{\lambda_1,\lambda_2,...,\lambda_j}^{-k(i)}$ represents the wavelet estimate based on the thresholds $\lambda_1, \lambda_2, ..., \lambda_J$ after removing the k^{th} part of the data. Here λ_J represents threshold value at resolution level j .

E. Ebayes Thresholding Methods

Recently various Bayesian approaches are introduced to choose a threshold value. See for example Abramovich et al. [2], Clyde and George [4, 5]. In this paper, we are interested in describing the approach that introduced by Johnstone and Silverman [11].

In this approach, given a single observation X_i subject to noise ε_i we can write

$$X_i = \mu_i + \varepsilon_i$$

Where X_i is drawn independently from a normal distribution with mean μ_i and variance σ^2 and $\varepsilon_i \sim N(0,1)$. An achieving a threshold value using this approach involves three main aspects.

1) First, a Bayesian model is used for the parameters μ_i. In this model μ_i is assumed to be zero with probability $(1-\tau)$ and with probability τ to be drawn from a symmetric heavy-tailed density γ such Laplace or Cauchy density. Define the prior distribution of μ_i as

$$f_{prior}(\mu_i) = (1-\tau)\delta_0(\mu_i) + \tau\gamma(\mu_i) \qquad (4)$$

Here δ_0 is the Dirac function.

2) Given a sequence observations, the weight τ is automatically chosen from the data using a marginal maximum likelihood approach. The marginal maximum likelihood estimator $\hat{\tau}$ of τ maximizes the marginal log likelihood and it can be written as

$$l(\tau) = \sum_{i=1}^{n}\log\{(1-\tau)\varnothing(X_i) + \tau g(X_i)\}$$

Where g denotes the convolution of the density γ with the standard normal $\varnothing$.

3) After obtaining $\hat{\tau}$, an estimate for μ_i is found by substituting $\hat{\tau}$ back into the prior (4) and taking the posterior median of μ given $X_i = x_i$. In this case, let :

$$\hat{\mu}_i = \hat{\mu}(x_i, \hat{\tau})$$

Where $\hat{\mu}(x_i, \hat{\tau})$ is the median of the posterior distribution of μ given $X_i = x_i$. However, we can use also the mean of the posterior distribution as an estimate for μ_i . For any fixed $\tau < 1$, the posterior median will be a thresholding rule, in that there exists a threshold $t(\tau) > 0$ such that $\hat{\mu}(x_i, \tau) = 0$ with the constrain $|x_i| \le t(\tau)$. Hence, the estimated value $\hat{\tau}$ gives an estimated threshold $t(\hat{\tau}) = \hat{t}$.

IV. SIMULATION

In this section, a simulation study was conducted to compare the five methods:

1) Universal: Donoho and Johnstone procedure [6].
2) Sure: Donoho and Johnstone procedure [7].
3) Two fold CV: two fold cross validation of Nason [14].
4) Ebayes: the empirical EBayes procedure of Johnstone and Silverman [11].
5) Level dep. CV : level dependent cross validation introduced by Oh, Kim and Lim [13].

Four test functions were used. Heavsine, Doppler which introduced by Donoho Johnstone [6]. Fg1 of Fan and Gijbels [8], Piecewise polynomial of Nason and Silverman [15].

Three different kinds of noise were used:

1) Independently distributed normal noise,
2) Independently distributed Student's t noise with three degrees of freedom.
3) Correlated normally distributed deviates from AR (1) of lag 1 with parameter 1/2 as in Nason [14]. All errors have zero mean and constant variances.

Five levels of signal to noise ratio (snr) were used: $snr = 2,5,7,9$ and 10. Also two different sample sizes were chosen: $n = 512,1024$. For every combination of test function, noise structure, level to noise and sample size, 1000 samples were generated. For each generated data sets, the five methods were applied to get an estimate for the test functions and then the mean squared error was computed.

For addition information regarding this simulation: mother wavelet $N = 6$ was used in every wavelet transform, Soft thresholding was used for every method, the formula (2) was used to find the variances for Universal and Ebayes, Laplace density with the median of posterior were used for Ebayes, all simulation results were carried out using the waveThresh package of Nason [16] in R.

Table I, Table II and Table III report the average of the mean squared error under 1000 replications. Having closed look at these tables, the following major interpretation can be made: For Gaussian noises, the Ebayes method provides the best results in all cases without any exception. It is remarkable to notice that for $(snr = 2)$, Universal and Sure have the same mean squared errors though the equality disappears when snr increases. Two fold cross validation does the best in the case of Student's t noise (long tail noises t(3)).

For Correlated noises from AR (1) of lag 1 with parameter 1/2, we noticed that: Two fold cross validation and Ebayes perform badly because of the correlation structure within the data set. In most cases large values of snr eg. $(snr = 9,10)$ seems to make level dependent cross validation do better while it does poorly when snr becomes less and less.

For level dependent cross validation 32 blocks were left out with block size 4. Coming back to Oh, Kim and Lim' results [13], they left 64 blocks out with block size 2. The result was not similar to theirs but the same conclusion has been found.

In this work, a different block size is used to show how strongly level dependent cross validation depends on the block sizes. However, finding a good block size for level dependent cross validation is left for further research. Finally, it was expected to notice that increasing sample size and level of signal to noise ratio, improves estimation efficiency of all used methods and this suits theoretical considerations.

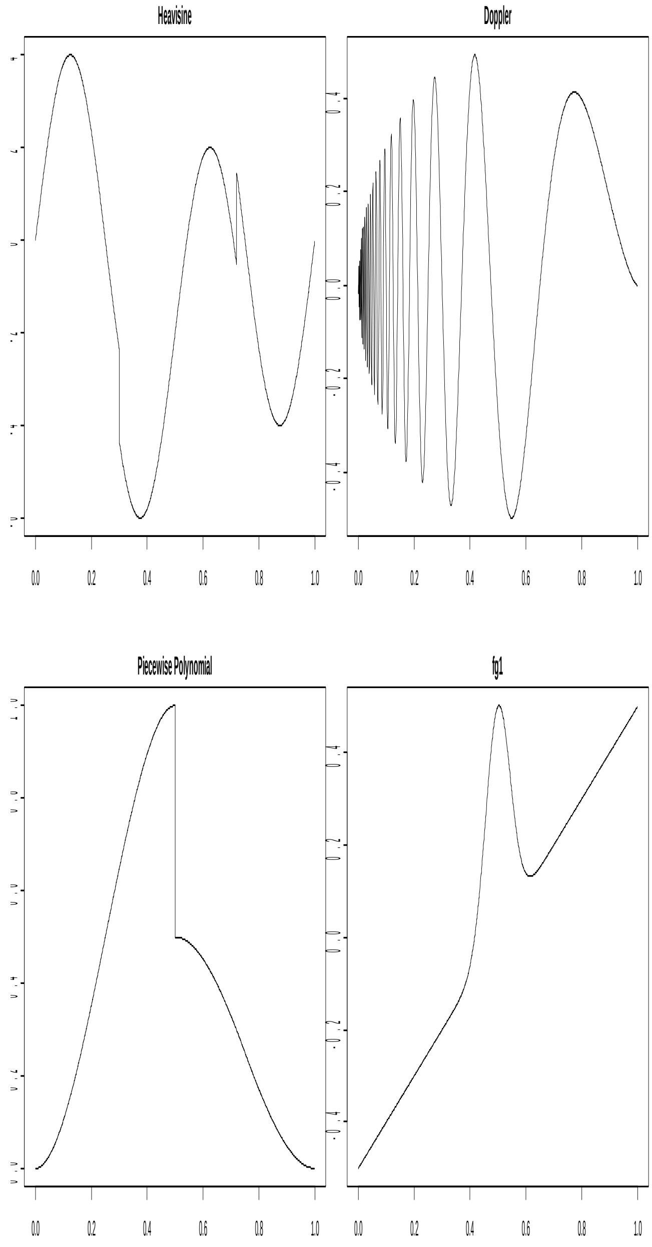

Fig. 1 The four functions used in simulation.

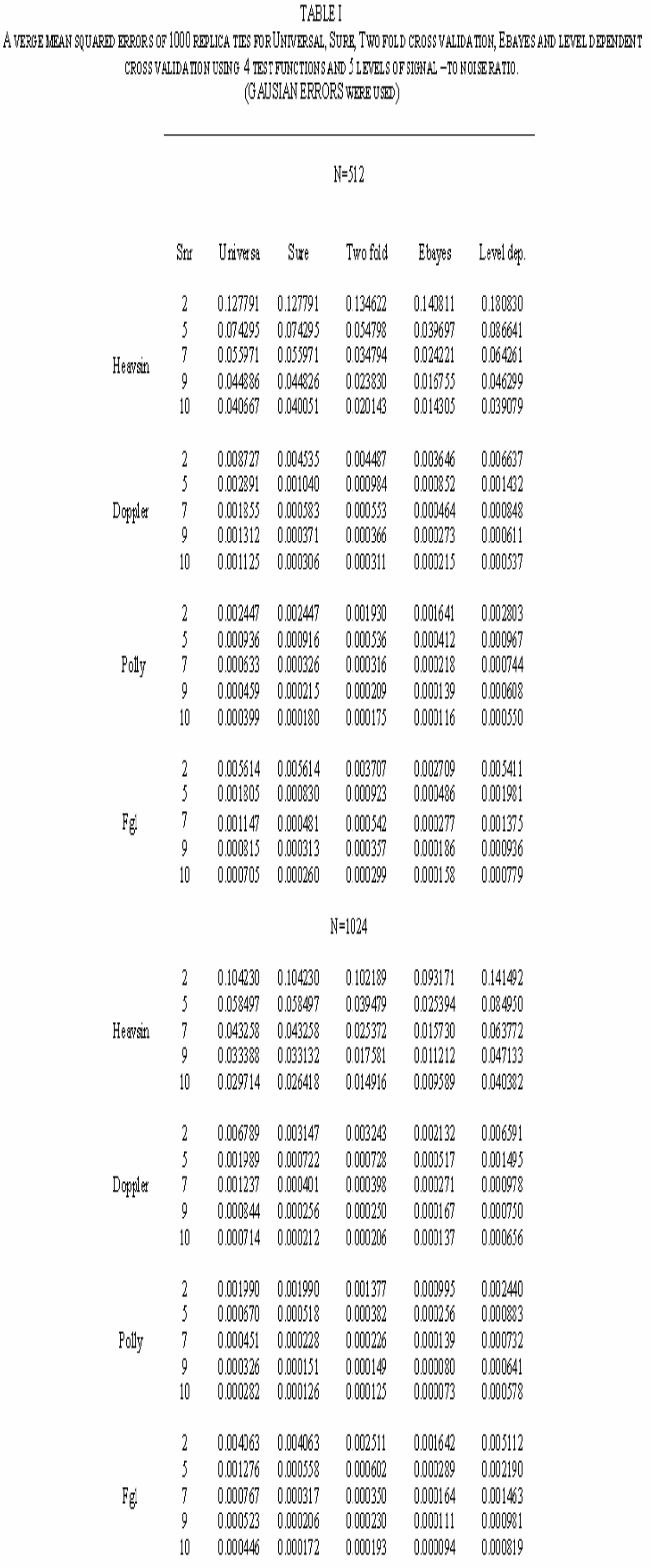

TABLE I
AVERGE MEAN SQUARED ERRORS OF 1000 REPLICATIES FOR UNIVERSAL, SURE, TWO FOLD CROSS VALIDATION, EBAYES AND LEVEL DEPENDENT CROSS VALIDATION USING 4 TEST FUNCTIONS AND 5 LEVELS OF SIGNAL –TO NOISE RATIO.
(GAUSIAN ERRORS WERE USED)

	Snr	Universa	Sure	Two fold	Ebayes	Level dep.
				N=512		
Heavsin	2	0.127791	0.127791	0.134622	0.140811	0.180830
	5	0.074295	0.074295	0.054798	0.039697	0.086641
	7	0.055971	0.055971	0.034794	0.024221	0.064261
	9	0.044886	0.044826	0.023830	0.016755	0.046299
	10	0.040667	0.040051	0.020143	0.014305	0.039079
Doppler	2	0.008727	0.004535	0.004487	0.003646	0.006637
	5	0.002891	0.001040	0.000984	0.000852	0.001432
	7	0.001855	0.000583	0.000553	0.000464	0.000848
	9	0.001312	0.000371	0.000366	0.000273	0.000611
	10	0.001125	0.000306	0.000311	0.000215	0.000537
Polly	2	0.002447	0.002447	0.001930	0.001641	0.002803
	5	0.000936	0.000916	0.000536	0.000412	0.000967
	7	0.000633	0.000326	0.000316	0.000218	0.000744
	9	0.000459	0.000215	0.000209	0.000139	0.000608
	10	0.000399	0.000180	0.000175	0.000116	0.000550
Fgl	2	0.005614	0.005614	0.003707	0.002709	0.005411
	5	0.001805	0.000830	0.000923	0.000486	0.001981
	7	0.001147	0.000481	0.000542	0.000277	0.001375
	9	0.000815	0.000313	0.000357	0.000186	0.000936
	10	0.000705	0.000260	0.000299	0.000158	0.000779
				N=1024		
Heavsin	2	0.104230	0.104230	0.102189	0.093171	0.141492
	5	0.058497	0.058497	0.039479	0.025394	0.084950
	7	0.043258	0.043258	0.025372	0.015730	0.063772
	9	0.033388	0.033132	0.017581	0.011212	0.047133
	10	0.029714	0.026418	0.014916	0.009589	0.040382
Doppler	2	0.006789	0.003147	0.003243	0.002132	0.006591
	5	0.001989	0.000722	0.000728	0.000517	0.001495
	7	0.001237	0.000401	0.000398	0.000271	0.000978
	9	0.000844	0.000256	0.000250	0.000167	0.000750
	10	0.000714	0.000212	0.000206	0.000137	0.000656
Polly	2	0.001990	0.001990	0.001377	0.000995	0.002440
	5	0.000670	0.000518	0.000382	0.000256	0.000883
	7	0.000451	0.000228	0.000226	0.000139	0.000732
	9	0.000326	0.000151	0.000149	0.000080	0.000641
	10	0.000282	0.000126	0.000125	0.000073	0.000578
Fgl	2	0.004063	0.004063	0.002511	0.001642	0.005112
	5	0.001276	0.000558	0.000602	0.000289	0.002190
	7	0.000767	0.000317	0.000350	0.000164	0.001463
	9	0.000523	0.000206	0.000230	0.000111	0.000981
	10	0.000446	0.000172	0.000193	0.000094	0.000819

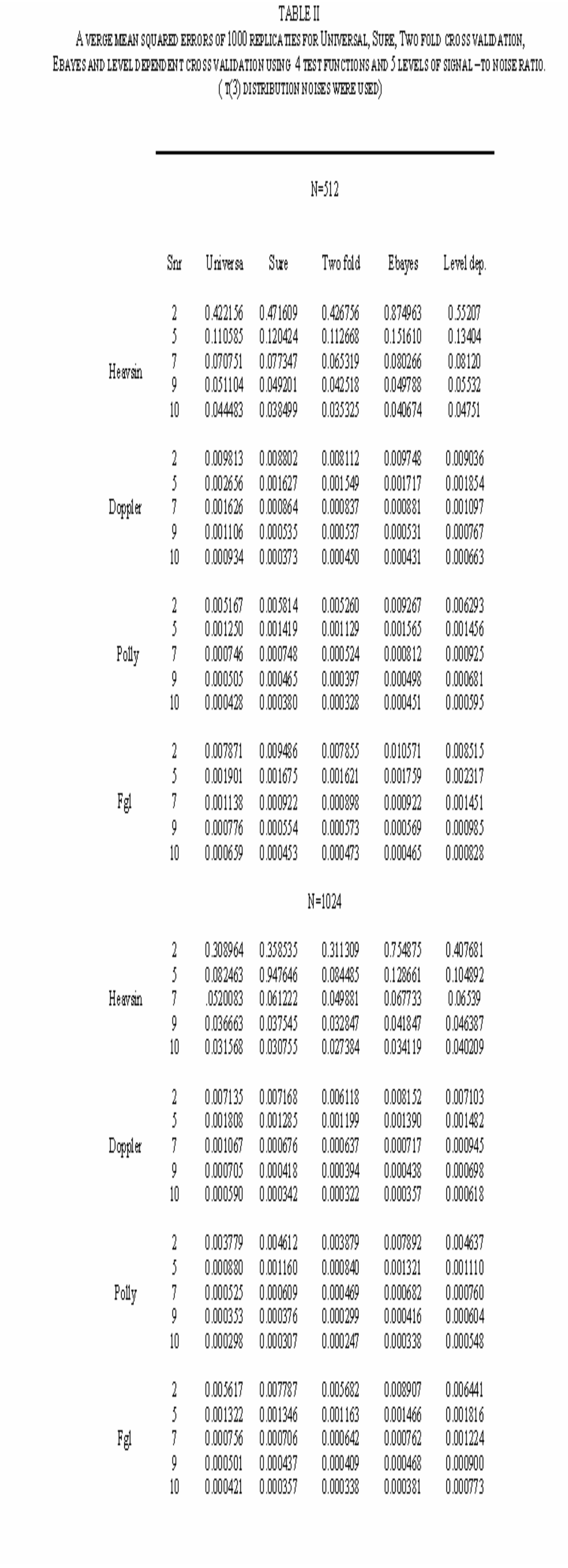

TABLE II
A VERGE MEAN SQUARED ERRORS OF 1000 REPLICA TIES FOR UNIVERSAL, SURE, TWO FOLD CROSS VALIDATION, EBAYES AND LEVEL DEPENDENT CROSS VALIDATION USING 4 TEST FUNCTIONS AND 5 LEVELS OF SIGNAL –TO NOISE RATIO.
(T(3) DISTRIBUTION NOISES WERE USED)

N=512

	Snr	Universa	Sure	Two fold	Ebayes	Level dep.
Heavsin	2	0.422156	0.471609	0.426756	0.874963	0.55207
	5	0.110585	0.120424	0.112668	0.151610	0.13404
	7	0.070751	0.077347	0.065319	0.080266	0.08120
	9	0.051104	0.049201	0.042518	0.049788	0.05532
	10	0.044483	0.038499	0.035325	0.040674	0.04751
Doppler	2	0.009813	0.008802	0.008112	0.009748	0.009036
	5	0.002656	0.001627	0.001549	0.001717	0.001854
	7	0.001626	0.000864	0.000837	0.000881	0.001097
	9	0.001106	0.000535	0.000537	0.000531	0.000767
	10	0.000934	0.000373	0.000450	0.000431	0.000663
Polly	2	0.005167	0.005814	0.005260	0.009267	0.006293
	5	0.001250	0.001419	0.001129	0.001565	0.001456
	7	0.000746	0.000748	0.000524	0.000812	0.000925
	9	0.000505	0.000465	0.000397	0.000498	0.000681
	10	0.000428	0.000380	0.000328	0.000451	0.000595
Fgl	2	0.007871	0.009486	0.007855	0.010571	0.008515
	5	0.001901	0.001675	0.001621	0.001759	0.002317
	7	0.001138	0.000922	0.000898	0.000922	0.001451
	9	0.000776	0.000554	0.000573	0.000569	0.000985
	10	0.000659	0.000453	0.000473	0.000465	0.000828

N=1024

	Snr	Universa	Sure	Two fold	Ebayes	Level dep.
Heavsin	2	0.308964	0.358535	0.311309	0.754875	0.407681
	5	0.082463	0.947646	0.084485	0.128661	0.104892
	7	.0520083	0.061222	0.049881	0.067733	0.06539
	9	0.036663	0.037545	0.032847	0.041847	0.046387
	10	0.031568	0.030755	0.027384	0.034119	0.040209
Doppler	2	0.007135	0.007168	0.006118	0.008152	0.007103
	5	0.001808	0.001285	0.001199	0.001390	0.001482
	7	0.001067	0.000676	0.000637	0.000717	0.000945
	9	0.000705	0.000418	0.000394	0.000438	0.000698
	10	0.000590	0.000342	0.000322	0.000357	0.000618
Polly	2	0.003779	0.004612	0.003879	0.007892	0.004637
	5	0.000880	0.001160	0.000840	0.001321	0.001110
	7	0.000525	0.000609	0.000469	0.000682	0.000760
	9	0.000353	0.000376	0.000299	0.000416	0.000604
	10	0.000298	0.000307	0.000247	0.000338	0.000548
Fgl	2	0.005617	0.007787	0.005682	0.008907	0.006441
	5	0.001322	0.001346	0.001163	0.001466	0.001816
	7	0.000756	0.000706	0.000642	0.000762	0.001224
	9	0.000501	0.000437	0.000409	0.000468	0.000900
	10	0.000421	0.000357	0.000338	0.000381	0.000773

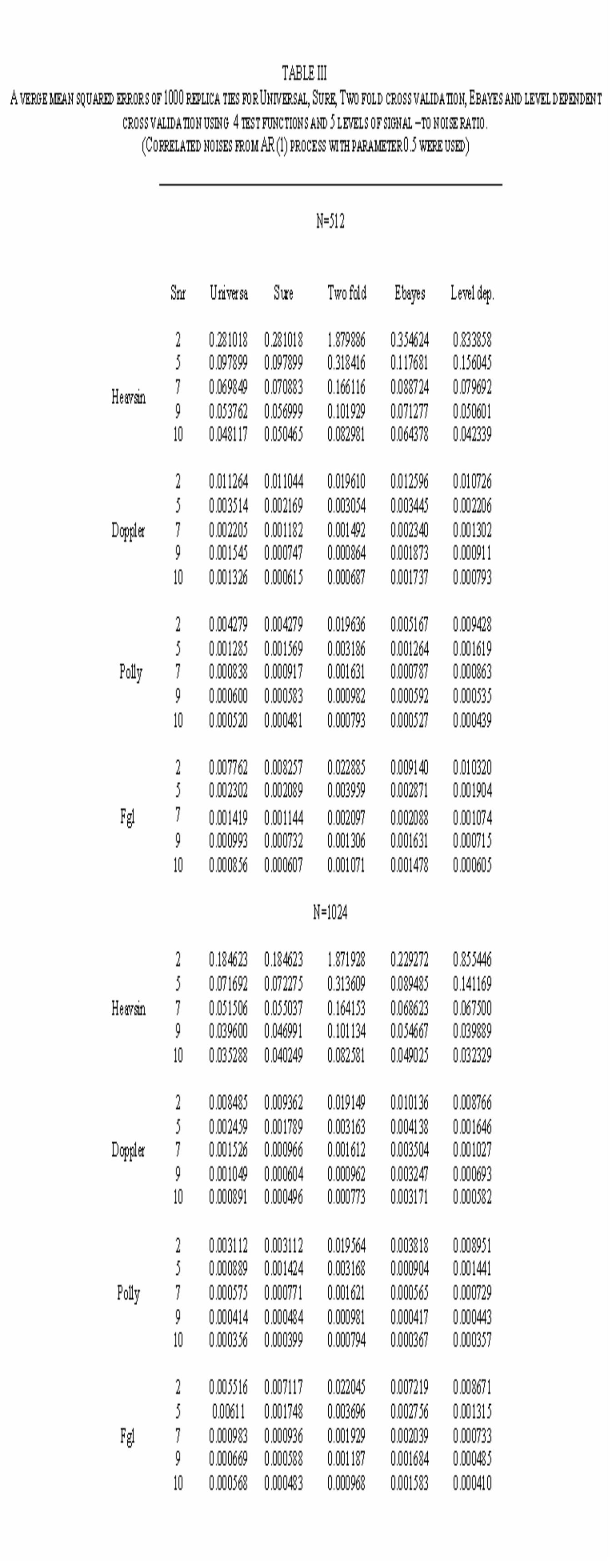

TABLE III
A VERGE MEAN SQUARED ERRORS OF 1000 REPLICA TIES FOR UNIVERSAL, SURE, TWO FOLD CROSS VALIDATION, EBAYES AND LEVEL DEPENDENT CROSS VALIDATION USING 4 TEST FUNCTIONS AND 5 LEVELS OF SIGNAL –TO NOISE RATIO.
(CORRELATED NOISES FROM AR (1) PROCESS WITH PARAMETER 0.5 WERE USED)

N=512

	Snr	Universa	Sure	Two fold	Ebayes	Level dep.
Heavsin	2	0.281018	0.281018	1.879886	0.354624	0.833858
	5	0.097899	0.097899	0.318416	0.117681	0.156045
	7	0.069849	0.070883	0.166116	0.088724	0.079692
	9	0.053762	0.056999	0.101929	0.071277	0.050601
	10	0.048117	0.050465	0.082981	0.064378	0.042339
Doppler	2	0.011264	0.011044	0.019610	0.012596	0.010726
	5	0.003514	0.002169	0.003054	0.003445	0.002206
	7	0.002205	0.001182	0.001492	0.002340	0.001302
	9	0.001545	0.000747	0.000864	0.001873	0.000911
	10	0.001326	0.000615	0.000687	0.001737	0.000793
Polly	2	0.004279	0.004279	0.019636	0.005167	0.009428
	5	0.001285	0.001569	0.003186	0.001264	0.001619
	7	0.000838	0.000917	0.001631	0.000787	0.000863
	9	0.000600	0.000583	0.000982	0.000592	0.000535
	10	0.000520	0.000481	0.000793	0.000527	0.000439
Fgl	2	0.007762	0.008257	0.022885	0.009140	0.010320
	5	0.002302	0.002089	0.003959	0.002871	0.001904
	7	0.001419	0.001144	0.002097	0.002088	0.001074
	9	0.000993	0.000732	0.001306	0.001631	0.000715
	10	0.000856	0.000607	0.001071	0.001478	0.000605

N=1024

	Snr	Universa	Sure	Two fold	Ebayes	Level dep.
Heavsin	2	0.184623	0.184623	1.871928	0.229272	0.855446
	5	0.071692	0.072275	0.313609	0.089485	0.141169
	7	0.051506	0.055037	0.164153	0.068623	0.067500
	9	0.039600	0.046991	0.101134	0.054667	0.039889
	10	0.035288	0.040249	0.082581	0.049025	0.032329
Doppler	2	0.008485	0.009362	0.019149	0.010136	0.008766
	5	0.002459	0.001789	0.003163	0.004138	0.001646
	7	0.001526	0.000966	0.001612	0.003504	0.001027
	9	0.001049	0.000604	0.000962	0.003247	0.000693
	10	0.000891	0.000496	0.000773	0.003171	0.000582
Polly	2	0.003112	0.003112	0.019564	0.003818	0.008951
	5	0.000889	0.001424	0.003168	0.000904	0.001441
	7	0.000575	0.000771	0.001621	0.000565	0.000729
	9	0.000414	0.000484	0.000981	0.000417	0.000443
	10	0.000356	0.000399	0.000794	0.000367	0.000357
Fgl	2	0.005516	0.007117	0.022045	0.007219	0.008671
	5	0.00611	0.001748	0.003696	0.002756	0.001315
	7	0.000983	0.000936	0.001929	0.002039	0.000733
	9	0.000669	0.000588	0.001187	0.001684	0.000485
	10	0.000568	0.000483	0.000968	0.001583	0.000410

ACKNOWLEDGMENT

The authors would like to thank USM for financial support.

REFERENCES

[1] F. Abramovich, F. Sapatinas, and B. W. Silverman ,"Wavelet thresholding via a Bayesian approach," J. R. Stat. Soc., B 60: 725–749, (1998).

[2] F. Abramovich , T. C. Bailey, T. Sapatinas, "Wavelet analysis and its statistical application," The Statistician 49:1-29, 2000).

[3] S. Barber, G. P. Nason, " Real Nonparametric regression using complex wavelets," J. R. Stat. Soc., B 66:927-939, (2004).

[4] M. Clyde, E. I. George, "Empircal Bayes estimation in wavelet nonparametric regression. In Wavelet-Based Models (lecture Notes in Statistcs, Vol. 141)," edit by P. Mller, B. Vidakovice. Oppenheim,309-322. Springer-Verlage, New York, (1999).

[5] M. Clyde, E. I. George , "Flexible empircal Bayes estimation for wavelets," J. R. Stat. Soc., B 62: 681-698, (2000).

[6] D. L. Donoho, I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," Biometrika 81:425–455, (1994).

[7] D. L. Donoho, I. M. , "Johnstone Adapting to unknown smoothing via wavelet shrinkage," J. Am. Stat. Assoc. 90:1200–1224, (1995).

[8] J. Fan, I. Gijbels, "Data-Driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaption," J. R. Stat. Soc., B 7:371-394, (1995).

[9] P. J. Green, B. W. Silverman, "Nonparametric Regression and Generalized Linear Models, A roughness Penelty Approach," Chapman and Hall, London. (1994).

[10] I. M. Johnstone, B. W. Silverman, "Ebayesthresh: R programs for Empirical Bayes thresholding," J. Stat. Softw. 12: 1-38, (2005a).

[11] I. M. Johnstone, B. W. Silverman, "Empirical Bayes selection of wavelet thresholds," Ann. Stat. 33:1700–1752, (2005b).

[12] D. Kim, H. S. Oh , "CVThresh: R Package for Level Dependent Cross –Validation Thresholding", J. Stat. Softw. (2006).

[13] H. S. Oh, D. Kim, Y. Lee, "Cross-validated wavelet shrinkage. Springer," New York 24:497-512, (2008).

[14] G. P. Nason, "Wavelet shrinkage by cross-validation," J. R. Stat. Soc., B 58:463–479, (1996).

[15] G. P. Nason, B. W. Silverman, "The discrete wavelet transform in S," J. Comput. Graph. Stat. 3:163–191, (1994).

[16] G. P. Nason, "(WaveThresh3 Software. Department of Mathematics, University of Bristol, UK. URL http://www.stats.bris.ac.uk/~wavethresh/.1998).

[17] G. P. Nason ,"Wavelet Methods in Statistics with R," Springer, New York. (2006).

[18] B. W. Silverman , "Density estimation for Statistics and Data Analysis," Chapman and Hall, London, (1986).

[19] B. W. Silverman "Empirical Bayes thresholding: adapting to sparsity when it advantageous to do so," J. Korean Stat. Soc. 36:1–29, (2007).

[20] C. Stein , "Estimation of the mean of a multivariate normal distribution," Ann. Stat. 9:1135-1151, (1981).

[21] M. Stone, "Cross –validatory choice and assessment of statistical predictions," J. R. Stat. Soc., B 36:111-147, (1974).

[22] B. Vidakovic, "Statistical modeling by wavelets," Wiley, New York, (1999a).