FUZZY THRESHOLD BASED WAVELET PROCESSING OF RANDOM TIME SERIES AND HIGHLY NOISY ACCELEROGRAMS

M. Beena Mol, S. Prabavathy and J. Mohanalin

1 Department of Civil Engineering, Noorul Islam University, Kumaracoil, Thuckalay -629 180, India
2 Department of Civil Engineering, Mepco Schlenk Engineering College, Anna University, Sivakasi – 626 005, India
3 Department of Electrical Engineering, College of Engineering, Pathanapuram (Under CAPE, Government of Kerala) – 691 005, India

Received: 26 March 2016; Accepted: 27 May 2016

ABSTRACT

Accelerograms are the random time series representation of the strong ground motions called earthquake. Processing of accelerograms is done before using them for any seismic and engineering applications. This article focuses the impact of threshold and noise in wavelet processing using experiments on random generated time series signals. Experiments conducted on the global thresholds currently in use lead us to propose a novel fuzzy based threshold to remove the random noise. Therefore a fuzzy based threshold determination in the domain of wavelet coefficients has been proposed and effectively used in denoising the real observed accelerograms. Also, the experimental results on synthetic signals has outperformed the existing universal and sure thresholds in terms of SNR of 20.3 and RMSE of 0.0019, respectively.

Keywords: Accelerograms; earthquake; fuzzy; de-noising; wavelet transforms.

1. INTRODUCTION

Processing the accelerograms [1-3] is done to analyze the strong ground motion (SGM) observations, for a meaningful seismic analysis of structures [4-6] and other engineering and seismic applications [7, 8]. An accelerogram contains a lot of coherent and random noises of environmental disturbances [1, 2, 8]. To remove these unwanted noisy disturbances, the
wavelet based processing tools [2, 3, 6] can be used to bring out the actual information of earthquake. Wavelet Transform (WT), originally introduced to improve the seismic signal analysis is the non-parametric representation of a real or complex valued squared integral function by a certain orthonormal series generated by a mother wavelet called the wavelet series [9]. It aids to detect and analyse abrupt changes in signals [10, 11] and is suitable for highly non-stationary signals characterised with sudden picks and discontinuities. High and low frequency scales, allow the analysis of rapidly and slowly changing details respectively. Low scales correspond to compressed wavelets, which are capable of capturing rapidly changing features of the signal linked to high frequencies. The high scales correspond to dilated wavelets that are able to capture the slowly changing features of the signal linked to low frequencies [12].

2. DISCRETE WAVELET TRANSFORM (DWT)

DWT [3] has been defined as the integration of the signal multiplied by the scaling \( a_0 \) and shifting \( b_0 \) parameters over the signal length. If \( X_m(t) \) is the discrete time series representation of the signal, then the forward discrete wavelet transform (FDWT) on the discrete time series \( X_m(t) \) using a mother wavelet function \( \psi \) and a scaling function \( \phi \) produces a set of wavelet coefficients \( C(j, k) \). It is a combination of approximations \( C_A(j, k) \) and details \( C_D(j, k) \) as expressed in Equation (1). The inverse discrete wavelet transform (IDWT) reconstruct the discrete time series representation of the signal \( X_m(t) \) as expressed in Equation (2).

\[
\text{WT}_{\psi, \phi}[X_m(t)] = \frac{1}{\sqrt{M}} \sum_{m} X_m(t) \phi_{j_0, k}(m) + \frac{1}{\sqrt{M}} \sum_{m} X_m(t) \psi_{j, k}(m) = C_A(j, k) + C_D(j, k) = C(j, k)
\]

\[
\text{IDWT}_{\psi}[C(j, k)] = \frac{1}{\sqrt{M}} \sum_{k} C_{\phi}(j_0, k) \phi_{j_0, k}(m) + \sum_{j=0}^{\infty} \sum_{k} C_{\phi}(j, k) \psi_{j, k}(m) = X_m(t)
\]

Where \( m = 0, 1, 2, 3, \ldots, (M-1) \); \( M \) is a power of 2 \( (M = 2^J) \); \( j \) and \( k \) are the translation and dilation parameters; \( j > j_0 \); \( j = 0, 1, 2, \ldots, (J-1) \) and \( k = 0, 1, 2, \ldots, 2^{J-1} \).

2.1 Statement of the problem

Let \( X = \{X_m\}_{m=1}^{M} \) be the contaminated accelerogram of size \( M \). It is considered to be a non-parametric regression problem at \( 2^J \) regularly spaced time interval and is analytically expressed as in Equation (3).

\[
X(m) = f(m) + W(m), \quad m = 1, 2, 3, \ldots, M
\]

and \( M = 2^J \), \( J \in \mathbb{Z} \), the real numbers,

Where \( X(m) = \) observed accelerogram
FUZZY THRESHOLD BASED WAVELET PROCESSING OF RANDOM TIME ...

\( f(m) = \{ f_m \}_{m=1}^M \) is the unknown desired accelerogram to be recovered, and \( W(m) = \{ W_m \}_{m=1}^M \) is the non-stationary noise in the accelerogram.

The aim is to process and denoise the accelerogram \( X(m) \) and to obtain the estimate \( f'(m) \) of the desired accelerogram \( f(m) \) that minimises the root mean square error (RMSE) with greater SNR. The MSE and SNR are expressed in Equations (4) and (5) respectively.

\[
\begin{align*}
\text{RMSE} &= \frac{1}{M} \sum_{m=1}^M (f(m) - f'(m))^2 \\
\text{SNR} &= \frac{\sum [f(m)]^2}{\sum [f(m) - f'(m)]^2}
\end{align*}
\]

2.2 Existing thresholds in wavelet domain

The universal threshold [13-15], empirical Bayes estimate [15] and Stein's unbiased risk estimate (sure) threshold [2] are the thresholds found in literature used for processing the accelerograms. The universal threshold estimates asymptotically the optimal threshold by minimising the maximum error over the entire signal. Bayes threshold is estimated by assuming that the wavelet coefficients have a priori distribution and the sure threshold is determined by minimising the unbiased risk estimate.

Universal threshold is the optimal threshold in asymptotic sense and is expressed as in Equation (6).

\[
\lambda_{\text{uni}} = \sigma \sqrt{2 \log M}
\]

where, \( \sigma = \frac{\text{MAD}}{0.6745} = \text{noise variance} = \frac{\text{median of the absolute value of wavelet coefficients}}{0.6745} \)

In statistics, sure is an unbiased estimator of the mean-squared error of a given estimator providing an indication of the accuracy of a given estimator. Sure threshold is based upon the method for estimating the loss in an unbiased fashion. Let \( \lambda \) be an unknown threshold parameter and \( \lambda \in \mathbb{R}^d \), where, \( \mathbb{R}^d \) is the vector of the universal wavelet coefficients resulting from the FDWT. Let \( C(j,k) \in \mathbb{R}^d \) be a measurement vector of wavelet coefficients whose components are independent and identically distributed normally with mean \( \mu \) and variance \( \sigma^2 \). If \( h(C(j,k)) \) is an estimator of \( \lambda \) from \( C(j,k) \in \mathbb{R}^d \), and can be written as in Equation (7).

\[
h(C(j,k)) = C(j,k) + g(C(j,k))
\]

where, \( g \) is a weak differentiable derivative. Then, the sure threshold is expressed as in Equation (8).

\[
\text{Sure} \left( h(C(j,k)) \right) = d \sigma^2 + ||g(C(j,k))||^2 + 2 \sigma^2 \sum_{i=1}^d \frac{\partial}{\partial C(j,k)_i} g_i C(j,k)
\]
where, \( g_i(C(j,k)) \) is the \( i \)th component of the function \( g(C(j,k)) \), and \( \| \cdot \| \) is the Euclidean norm. The importance of sure is that it is an unbiased estimate of the mean-squared error or squared error risk of \( h(C(j,k)) \) as expressed in Equation (9) and (10).

\[
E_{\mu} \left\{ \text{Sure} \left( h(C(j,k)) \right) \right\} = \text{RMSE} \left( h(C(j,k)) \right) \tag{9}
\]

where,

\[
\text{RMSE} \left( h(C(j,k)) \right) = E_{\mu} \| h(C(j,k)) - \mu \|^2 \tag{10}
\]

The Bayes threshold technique utilizes the Bayesian framework for signals to calculate the threshold. It is assumed that all the \( M \) observations of SGM signal are independent of each other. Let \( d_{jk} \) and \( w_{jk} \) be the wavelet coefficients of desired and observed signal. In the Bayesian approach, a prior distribution is placed on each coefficient. A generalized Gaussian distribution (GGD) prior is used on the wavelet coefficients and the threshold is derived in the Bayesian framework. It is also assumed that the signal and the noise are independent and identically distributed Gaussian random variables having normal with zero mean as expressed in equation (11). \( X(m) \sim N(0, \tau^2) \) and is independent of the signal \( f(m) \). The variance is given in Equation (12).

\[
i.e., f(m) \sim N(0, \sigma^2) \text{ and } W(m) \sim N(0, \rho^2) \tag{11}
\]

Also \( \tau^2 = \sigma^2 + \rho^2 \tag{12} \)

The mixture prior distribution of the wavelet coefficients of the desired signal is assumed as in Equation (13).

\[
p_j(d_{jk} \mid \pi_j, \sigma_j^2) = (1 - \pi_j)\delta(d_{jk}) + \pi_j N(0, \sigma_j^2) \tag{13}
\]

where \( \pi_j \) is the probability that \( d_{jk} \) is non-zero. The standard deviation of the noise \( \rho \) is estimated using the median estimator from the first decomposition level wavelet coefficients sub band as in Equation (14).

\[
\rho = \frac{\text{med}(|w_{jk}|)}{0.6745} \tag{14}
\]

The variance of the observed signal \( \tau^2 \) is calculated as in Eq. (15).

\[
\tau^2 = \frac{1}{M} \sum_{m=1}^{M} w_{jk}^2 \tag{15}
\]

The standard deviation of the signal \( f(m) \) is calculated as in Eq. (16).
FUZZY THRESHOLD BASED WAVELET PROCESSING OF RANDOM TIME

\[ \sigma = \sqrt{\text{max}(\tau^2 - \rho^2)}, \tau^2 > \rho^2 \]  
(16)

Then the Bayes threshold value \( \lambda_{BS} \) is given by Eq. (17).

\[ \lambda_{BS} = \begin{cases} \frac{\rho^2}{\sigma}, & \text{if } \tau^2 > \rho^2 \\ \text{max}\{|w_{ij}|\}, & \text{otherwise.} \end{cases} \]  
(17)

2.3 Thresholding rules
After determining the appropriate global threshold, the hard, soft, semi-soft or Garrotte thresholding rule is applied to obtain the modified wavelet coefficients. In hard thresholding [15], the coefficients less than the threshold are set to zero and the coefficients greater than the threshold are retained. In soft thresholding [16], the wavelet coefficients below the threshold are set to zero and the wavelet coefficients above the thresholds are reduced by the threshold. The semi-soft thresholding is the combination of both hard and soft thresholding based on two positive thresholds and Garrotte thresholding is also a shrunk procedure. The hard thresholding rule (HTR), soft thresholding rule (STR), semi-soft thresholding rule (SSTR) and the Garrotte thresholding rule (GTR) are expressed in Equations (18), (19), (20) & (21) respectively.

HTR, \( C'(j, k)(F(t)) = \begin{cases} C(j, k) & |C(j, k)| > \lambda \\ 0 & |C(j, k)| < \lambda \end{cases} \)  
(18)

STR, \( C'(j, k)(F(t)) = \begin{cases} \text{sgn}(C(j, k))(|C(j, k)| - \lambda), & |C(j, k)| > \lambda \\ 0, & |C(j, k)| \leq \lambda \end{cases} \)  
(19)

SSTR, \( C'(j, k)(F(t)) = \begin{cases} \frac{\lambda_2 |C(j, k)| - \lambda_1}{\lambda_2 - \lambda_1}, & \lambda_1 < |C(j, k)| \leq \lambda_2 \\ C(j, k), & |C(j, k)| > \lambda_2 \end{cases} \)  
(20)

GTH, \( C'(j, k)(F(t)) = \begin{cases} C(j, k) - \frac{\lambda^2}{C(j, k)}, & |C(j, k)| > \lambda \\ 0, & |C(j, k)| < \lambda \end{cases} \)  
(21)

3. ROLE OF THRESHOLD IN PROCESSING
The role of threshold in processing the desired signal is investigated by varying the threshold manually. The Gaussian non-stationary signals are generated using Matlab. The signal has been corrupted by adding the noise variance of 0.0075. The noisy signal has then been subjected to wavelet decomposition and is processed by varying the thresholds manually and by applying the HTR. Denoising results throughout this thesis contain x-axis as data points and y-axis as the corresponding amplitudes. The investigation has been carried out by varying the thresholds from 0.1 to 4.0. The processed results for the impact of threshold from 0.1 to 0.4 and from 0.5 to 3.0 are shown in Figs. 1 & 2 respectively. It is
observed that at very low thresholds of 0.1 & 0.15 the added noises exists in the signal. The existence of noise suggests that the threshold has to be increased.

Figure 1. Impact of threshold (0.1 to 0.4) in wavelet processing

Processed signals with thresholds 0.2 and greater have shown the removal of noise variance from the signal. The signals denoised with thresholds of 0.2 & 0.25 have shown close similarity with the original signal. Therefore, it is inferred that the thresholds of 0.2 & 0.25 have effectively removed the noise without disturbing the features of original signal.

The denoised signals processed with thresholds greater than 0.25 gradually have influenced in modifying the shape of the signal and hence have failed to preserve the shape of the original signal. Signals processed using thresholds greater than 1.0 have almost completely distorted the signal making it unusable to any future analysis and applications. The quality of the processed signals is verified using the RMSE and the SNR and are
described in Table 1. It is observed that the threshold 0.2 has produced the optimum denoised signal with an increased SNR of 17.7547 and a low RMSE of 0.0552. The thresholds greater than and equal to 0.5 have produced poor signal with very low SNR and hence degrades the quality of the signal under consideration. The graphical investigations of the RMSE and SNR for the various manually varied thresholds are shown in Figs. 3 and 4 respectively.

Figure 2. Impact of threshold (0.5 to 3.0) in wavelet processing
Table 1: RMSE and SNR of processed Gaussian signal

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
<th>RMSE</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0755</td>
<td>15.0352</td>
<td>0.7</td>
<td>0.1158</td>
<td>11.3231</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td>0.0635</td>
<td>16.5399</td>
<td>0.8</td>
<td>0.1348</td>
<td>9.9994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.0552</td>
<td>17.7547</td>
<td>0.9</td>
<td>0.1499</td>
<td>9.0791</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.0598</td>
<td>17.0564</td>
<td>1.0</td>
<td>0.1661</td>
<td>8.1862</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>0.0689</td>
<td>15.8333</td>
<td>1.5</td>
<td>0.2164</td>
<td>5.8891</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.35</td>
<td>0.0822</td>
<td>14.3027</td>
<td>2.0</td>
<td>0.2750</td>
<td>3.8098</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.45</td>
<td>0.1099</td>
<td>12.6053</td>
<td>2.5</td>
<td>0.2934</td>
<td>3.2451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>0.1068</td>
<td>12.0222</td>
<td>3.0</td>
<td>0.3072</td>
<td>2.8472</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. RMSE for processed synthetic signal

Figure 4. SNR for processed synthetic signal

4. ROLE OF THRESHOLD VS NOISE VARIANCES

The results obtained from the above experiment are extended to the Gaussian signals with varying variance. The signals are added with the noise of variance 0.001, 0.005 and 0.0075. The thresholds selected are 0.1, 0.2, 0.3, 0.4, 0.5, 1.0 and 2.0. RMSE and the SNR performance measures of the thresholds are described in Table 2. It is observed that as the level of noise variance increases, the threshold needed to process the signal is also increased. It is also understood that in all cases very high thresholds result in degrading the quality of the signal under investigation. Also, it is clear that choosing any of the random thresholds will not result in the optimum denoising.
The quantitative performance of the synthetic signals with the manually selected thresholds for the added noise variance 0.001, 0.005 and 0.0075 in terms of RMSE is shown in Fig. 5. Irrespective of the noise variance, the lower thresholds have produced the reduced RMSE and thereby have shown the better performance. Also, for the signal with very low noise variance of 0.001, there is an increase in RMSE for the selected thresholds from 0.1 to 2.0. But for the signals with increased variance 0.005 and 0.0075, the RMSE is slightly higher for the threshold value of 0.1. But, by increasing the threshold to 0.2 has produced the optimum minimum RMSE. On further increasing the thresholds, the RMSE also is increased indicating the gradual degradation in the quality of the signal with a gradual increase of threshold beyond the optimum threshold. The quantitative performance of the Gaussian synthetic signals with the manually selected thresholds for the added noise variance 0.001, 0.005 and 0.0075 in terms of SNR are shown in Fig. 6.

Table 2: RMSE and SNR of the processed Gaussian signal (threshold Vs noise variances)

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Variance = 0.001</th>
<th>Variance = 0.005</th>
<th>Variance = 0.0075</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>SNR</td>
<td>RMSE</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0219</td>
<td>24.9747</td>
<td>0.0596</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0404</td>
<td>18.9236</td>
<td>0.0508</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0583</td>
<td>15.7346</td>
<td>0.0666</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0697</td>
<td>14.1874</td>
<td>0.0791</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0912</td>
<td>11.8523</td>
<td>0.1017</td>
</tr>
<tr>
<td>1.0</td>
<td>0.1515</td>
<td>7.4440</td>
<td>0.1698</td>
</tr>
<tr>
<td>2.0</td>
<td>0.2327</td>
<td>3.7152</td>
<td>0.3093</td>
</tr>
</tbody>
</table>

Figure 5. RMSE for the processed Gaussian signals (threshold Vs noise variances)

Figure 6. SNR for the processed Gaussian signals (threshold Vs noise variances)
It is observed that, irrespective of the noise variance, lower thresholds have produced increased SNR. Also, for the signal with very low noise variance of 0.001, there is a decrease in SNR for the selected thresholds from 0.1 to 2.0. But, for the signals with increased variance 0.005 and 0.0075, the SNR is slightly lower for the threshold value of 0.1. By increasing the threshold to 0.2 has produced the optimum maximum SNR. On further increasing the thresholds, the SNR has decreased, indicating the gradual degradation in the quality of the signal. It has also been observed that irrespective of the noise variances, the lower thresholds have produced the least RMSE. The reason is that, most of the wavelet coefficients representing the noise variances takes the lower value in wavelet domain.

From the above investigations on the random Gaussian synthetic signals, it can be concluded that as the noise variance increases, the threshold to be chosen for producing the optimum reconstruction is also increased. Choosing the optimum random threshold is time consuming and has to be done in a trial and error basis. Therefore, there is a need for the automatic identification of the optimum threshold that will end with the optimum reconstruction of the decomposed signal without removing and disturbing the essential and important features of engineering concern.

5. ROLE OF GLOBAL THRESHOLD VS NOISE VARIANCES

Experiments have been conducted on the random Gaussian signals to investigate the role of the existing global thresholds in processing. The performance of the universal, sure and Bayes thresholds is quantitatively measured using the RMSE and SNR. The quantitative measures are also compared with the performance of the manual thresholds. The experiments are carried out on the random synthetic signals with added noise variance of 0.01, 0.0025, 0.005 and 0.0075. The processed synthetic signals with an added noise variance of 0.001, 0.005 and 0.0075 are shown in Figs. 7, 8 and 9 respectively. Through visual inspection, it has been observed that for the signal with a noise variance of 0.001, the signals processed using the universal threshold and the manual threshold 0.1 have reconstructed the original signal without disturbing the features to the maximum. For the signal with a noise variance of 0.005, the signal processed using the universal threshold and the manual threshold of 0.2 has reconstructed the original signal without disturbing the features to the maximum. But for the signal with a noise variance of 0.0075, the signals processed using the universal, Bayes and the manual threshold of 0.2 have reconstructed the original signal without disturbing the features to the maximum.

For all the synthetic signals, irrespective of the level of noise variance, the wavelet processed signals using the sure threshold have shown poor performance.
The sure threshold has overly smoothened all the noisy signals. Also, it is observed from the Figs. 5, 6 and 7 that the sure threshold has removed the significant features of the signal together with the added noise variance. Moreover, for all the cases of noise variances, the quality of the processed signal is better for the manually selected threshold than the universal, sure and the Bayes threshold.

The comparative quantitative measures of RMSE and SNR of the processed signals using the universal, sure and the Bayes thresholds for added noise variances of 0.001 and 0.0025 is tabulated in Table 3. Similarly, the RMSE and SNR of the processed signals using the universal, sure and the Bayes thresholds for added noise variance of 0.0050 and 0.0075 is tabulated in Table 4. It is observed that the processing based on the universal threshold has produced more efficient wavelet processing than the sure and Bayes threshold.
Also, for signals with increased noise variances, the Bayes threshold and the universal threshold have equally performed better when compared with the sure threshold. The graphical comparison of noise variance for universal, sure and Bayes thresholds is shown in Fig. 10. From the figure, it is observed that the universal threshold increases gradually and slowly with respect to the increase in the noise variance. For very low noise variance, the Bayes threshold has shown an increased threshold and as the noise variance increases the bayes threshold is slightly decreased. But, the sure threshold always shows a higher value which indicates that the sure threshold removes the essential features of the signal at all levels of noise variances.
Table 4: Threshold, RMSE and SNR of the Gaussian signal (global thresholds Vs noise variances = 0.005 & 0.0075)

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Variance = 0.005</th>
<th>Variance = 0.0075</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Threshold</td>
<td>RMSE</td>
</tr>
<tr>
<td>Universal</td>
<td>0.3195</td>
<td>0.0694</td>
</tr>
<tr>
<td>Sure</td>
<td>1.2055</td>
<td>0.2227</td>
</tr>
<tr>
<td>Bayes</td>
<td>0.5411</td>
<td>0.1041</td>
</tr>
</tbody>
</table>

The graphical comparison of the RMSE of the signals processed using universal, sure and Bayes thresholds with respect to the noise variance 0.001, 0.0025, 0.005 and 0.0075 are shown in Fig. 11. From the figure, it is observed that for all cases of noise variances, the universal threshold has shown a reduced RMSE and therefore universal threshold can be positively preferred for processing the non-stationary accelerograms.

![Figure 10. Global thresholds Vs noise variances for Gaussian synthetic signals](image1)

For very low noise variance, the Bayes threshold has shown an increased RMSE and as the level of noise variance increases, the RMSE is gradually decreased. Also, at higher noise variances, there is a chance that Bayes threshold may outperform the universal threshold in obtaining the low RMSE with respect to the desired signal. But, the processed signal with sure threshold for all noise variances always shows a higher RMSE value which indicates that the sure threshold removes the essential features of the signal at all level of noise variances. The graphical comparison of the SNR of the signals processed using universal,
sure and the Bayes thresholds with respect to the noise variance 0.001, 0.0025, 0.005 and 0.0075 are shown in Fig. 12. From the figure, it is observed that for all cases of noise variances, the universal threshold has shown an increased SNR and therefore universal threshold can be positively preferred for processing the non-stationary accelerograms. For very low noise variance, the bayes threshold has shown a decreased SNR and as the level of noise variance increases the SNR is also gradually increased. Also at higher noise variances, there is a chance that Bayes threshold may outperform the universal threshold in obtaining the higher SNR with respect to the desired signal. But the processed signal with sure threshold for all noise variances always shows a lower SNR value which indicates that the sure threshold removes the essential features of the signal at all level of noise variances.

In the process of wavelet processing, the thresholds play a major role in denoising the noisy signals. The experiments conducted on the Gaussian synthetic signals with different noise variances show that the existing universal, sure and Bayes thresholds denoise the signals to an extent. But, the signal denoised using the existing global thresholds are not optimum as inferred from the experiments by manually varied thresholds. It has also been observed that the sure threshold did not show a uniform trend of increasing or decreasing the SNR and RMSE with increased variance. Hence, sure threshold cannot be used effectively for noise removal of accelerograms because; accelerograms are mixed with different levels of noise variances. Therefore, in the field of earthquake and seismological engineering, there is a need for determining the new optimum threshold that will further improve the SNR with reduced RMSE.

6. FUZZY BASED ACCELEROMETER PROCESSING

To handle complex situations where the difference between classes of wavelet coefficients is ill defined, type I fuzzy is required [17]. Fuzzy sets are the classes of data whose elements can be expressed in degrees of membership. This means, both the data which belong and do not belong in a class, will enjoy certain degree of relation. Hence the boundary is not sharp, but blunt. This relation is mathematically written in form of membership function ($\mu(x)$). It assigns grade in the interval $[0, 1]$ to each of the values. If $X$ is the collection of points $x_1, x_2, x_3, \ldots, x_n$, then fuzzy set $F$ in mathematical form is
FUZZY THRESHOLD BASED WAVELET PROCESSING OF RANDOM TIME … 1145

expressed as in Equation (22).

\[ F = (x_i, \mu_F(x_i)) \] (22)

Where, \( x_i \in X \) and \( i \) varies from 1 till \( n \). The points in \( F \) map the input function with appropriate values ranging between 0 and 1.0. There are so many membership functions available that can do the mapping. However, this part of experiment utilizes \( pi \) membership function derived from \( S \) membership function as in Equation (23) owing to its smoothness and popularity.

\[
S(C;x,y,z) = \begin{cases} 
0 & C \leq x \\
\frac{1}{2} \left( \frac{(z-x)^2}{(z-x)} \right) & x \leq C \leq y \\
1 - \frac{1}{2} \left( \frac{(C-z)^2}{(z-x)} \right) & x \leq C \leq y \\
1 & C \geq z 
\end{cases} 
\] (23)

Where as, a \( pi \) membership function can be written as in Equation (24) as

\[
\Pi(C;b,t) = \begin{cases} 
S(C;\frac{t-b}{2}, C) & C \leq t \\
1 - S(C;\frac{t+b}{2},t+b) & otherwise 
\end{cases} 
\] (24)

The bandwidth \( b \) has to be calculated using \( b = \{max(t - \mu), (maxN - t)\} \). Here, \( \mu \) represents the average value of coefficients and \( maxN \) to the maximum value of the coefficients calculated. Hence, blunting of spikes is carefully avoided, where important earthquake details lie. But, the high frequency and low frequency pulses have been discarded. \( t \) has to be found using the Minkowski distance between two fuzzy sets. If \( F&R \) are any two fuzzy sets, then fuzzy similarity measure is a distance measure \( d(F,R) \) which can be mathematically written as in Equation (25).

\[
d(F,R) = \sum_{i=1}^{n} |\mu_F(x_i) - \mu_R(x_i)| 
\] (25)

The distance measure reflects the degree of ambiguity that exists between \( F \) and \( R \). Fig. 13 demonstrates the main idea of the proposal.
Typically, a low index denotes a low ambiguity; indicating the elements present in \( F \) and \( R \) tend to be similar. A high index value indicates the dissimilarity between the functions. Optimal threshold is the value which provides one the maximum value of \( d(F, R) \) and can be written as in Equation (26).

\[
t = \text{argmax}\{d(F, R)\}
\]

(26)

By investigating the strength and weakness of wavelet and fuzzy, a novel idea which fuses the advantages of both has arisen. Strength of wavelet lies in its ability to work in various resolutions; yet it does not rely on quality thresholds as noted in literature survey. At the same time, fuzzy index scheme is ideal in locating the threshold for complicated signals like accelerograms. Hence, the idea of clubbing wavelet with fuzzy for accelerograms processing has arisen.

Before denoising, the level of ambiguity presents in wavelet coefficients is higher. In theory, accelerogram is a non-linear mix of noise and signal in time domain, which may be inseparable. i.e., if the noisy part of earthquake signal is termed as class F and region of interest as class R, then problem of identifying the position which separates class R from class F is vague and ambiguous. To minimize the ambiguity, \( pi \) membership function is used for each element corresponding to the frequency coefficients. It is followed by identification of the maximum fuzzy index which will give one the optimal threshold to separate the information from noise. It is followed by hard thresholding technique to suppress the noise from region of interest. Then the reconstructed wavelet coefficients would provide the de-noised accelerogram.

6.1 Implementation and result analysis
Table 5 describes the details of accelerogram components chosen for implementing type I fuzzy based processing. Like in the previous cases, first set of experiments are executed with synthetic Gaussian signals with sharp edges and drifts. This is followed by the experiments with real accelerograms taken from the strong motion centre database.
Table 5: Details of accelerograms used in type I fuzzy based processing

<table>
<thead>
<tr>
<th>Details</th>
<th>Accelerogram components</th>
<th>N-S</th>
<th>E-W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station Code</td>
<td></td>
<td>6503</td>
<td>6503</td>
</tr>
<tr>
<td>Station Name</td>
<td>Van Muradiye meteorology</td>
<td>Van Muradiye meteorology</td>
<td></td>
</tr>
<tr>
<td>Epicentre (km)</td>
<td></td>
<td>46.6</td>
<td>46.6</td>
</tr>
<tr>
<td>Earthquake date</td>
<td></td>
<td>23/10/2011</td>
<td>23/10/2011</td>
</tr>
<tr>
<td>Magnitude</td>
<td></td>
<td>7.2</td>
<td>7.2</td>
</tr>
<tr>
<td>Sampling interval</td>
<td></td>
<td>0.001(sec)</td>
<td>0.001(sec)</td>
</tr>
<tr>
<td>Sampling points</td>
<td></td>
<td>10714</td>
<td>10714</td>
</tr>
<tr>
<td>PGA</td>
<td></td>
<td>178.5 cm/s²</td>
<td>169.5 cm/s²</td>
</tr>
<tr>
<td>Earthquake</td>
<td></td>
<td>Eastern Turkey earthquake</td>
<td></td>
</tr>
</tbody>
</table>

The important steps in the type I fuzzy based accelerogram processing are as follows

- Choose db4 as mother wavelet.
- Derive the wavelet coefficients of the accelerogram under investigation.
- Select the shape of membership function.
- Initiate the position of the membership function.
- Calculate the S membership function.
- Calculate the pi membership function.
- Calculate the fuzzy index in each possible location.
- Calculate the optimal index and fix this as the optimal threshold.
- Perform the hard thresholding rule.
- Perform IDWT to reconstruct the processed accelerogram.

The SNR and RMSE comparison of type I fuzzy based threshold with universal and sure thresholds in accelerogram processing is described in Table 6. The RMSE comparison of the type I fuzzy based threshold with universal and sure thresholds based processing is shown in Fig. 14.

Table 6: RMSE and SNR for universal, sure and type I fuzzy based processing

<table>
<thead>
<tr>
<th>Added noise variance</th>
<th>Universal</th>
<th>Sure</th>
<th>Type I Fuzzy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>SNR</td>
<td>RMSE</td>
</tr>
<tr>
<td>0.27</td>
<td>0.0023</td>
<td>19.20</td>
<td>0.0028</td>
</tr>
<tr>
<td>0.7</td>
<td>0.027</td>
<td>17.19</td>
<td>0.031</td>
</tr>
<tr>
<td>2.7</td>
<td>0.0307</td>
<td>15.28</td>
<td>0.038</td>
</tr>
</tbody>
</table>
Figure 14. RMSE comparison of universal, sure and type I fuzzy based thresholds

It is obvious from the Fig. 14 that the proposal is superior in their denoising performance. Type I fuzzy based threshold supports the claims in the form of an improved SNR and decreased RMSE respectively. The clearance level between the plots drawn for type I fuzzy based processing and other ones is significant. It strongly emphasises the fact for accelerograms, better thresholds are indeed needed. Moreover, Fig. 15 characterizes the impact of the type I fuzzy, universal and sure thresholds in processing the real accelerograms. Type I fuzzy based processing has preserved the information like shape of sharp and peak edges efficiently than its counterparts.

Figure 15. Raw and processed accelerograms using the universal, sure and Type I fuzzy based thresholds
The proposed type I fuzzy based threshold preserves the information of engineering interest by reconstructing the sharp edges more effectively than the universal and sure threshold. Also, from the experiments performed on the synthetic and real accelerograms, it was observed that the proposed type I fuzzy based processing is technically superior, as fuzzy way of defining wavelet coefficients of seismic signal is more appropriate. Moreover, the S membership function selected is a smoother function and hence edges are well preserved.

7. CONCLUSION

It has been found that for representing the wavelet coefficients of time series accelerograms, fuzzy concept would be ideal. This has paved the way for a new fuzzy based denoising algorithm. Results are encouraging, when compared to the popularly existing thresholding schemes.

REFERENCES