A NEW THRESHOLDING TYPE TECHNIQUE FOR THE DETECTION OF SEISMIC EVENTS

*Athanasios E. Lois, Emmanouil Z. Psarakis and Erion-Vasilis Pikoulis**

Department of Geology - Department of Computer Engineering and Informatics
University of Patras 26500 Rion-Patras Greece
email:lois@upatras.gr-{psarakis, pikoulis}@ceid.upatras.gr

ABSTRACT

The problem of seismic events detection constitutes one of the most important and vital tasks for the automatic identification of the seismic phase arrivals. In this work, we propose a new thresholding type technique, tailored to fit real world situations where our knowledge on the statistical characteristics of the background noise process are unknown and a strict hypothesis testing framework can not be followed. In such cases the replacement of the unknown probability density function under the null hypothesis by its empirical counterpart, constitutes a possibility. In this work, a two stage procedure is proposed. The first one concerns the estimation of the empirical functions of the noise process itself as well as its whitened counterpart. In the second stage, using the above empirical functions, a thresholding scheme is proposed in order to solve the problem of the detection of seismic events in a non strict hypothesis testing framework. The performance of the proposed technique is confirmed by its application in a series of experiments both in synthetic and real seismic data sets.

Index Terms— Seismic events detection, Hypothesis testing, Estimation of empirical *pdf*, Thresholding.

1. INTRODUCTION

Methods that deal with the seismic event detection and phase identification are a topic of ongoing research. A robust signal detection algorithm plays a crucial role in events location, focal mechanism determination as well as in other applications such as passive seismic tomography investigations, early warning to name a few. Most of the techniques that have been proposed and applied up to now for seismic event detection, are mainly based on energy criteria and referred in the literature as Short Term Average/Long Term Average (STA/LTA) detectors. In its simplest form [1], the STA/LTA algorithm processes filtered seismic signals in two moving windows and compares their absolute values ratio to a user threshold value in order to declare a seismic event.

This idea was adopted by Allen [2, 3], Baer and Kradolfer [4] and Earle and Shearer [5], who propose different characteristic functions, based on the trace amplitude, which must be compared to an empirically set threshold. Apart from the well known STA/LTA detectors, seismic event detection algorithms that utilize polarization properties of the seismic signals, (by Magotra et al. [6], and Ruud and Husebay [7]) in order to enhance the detection rate of the seismic events have also been proposed. A comparison of existing time and frequency domain detection techniques can be found in the work of Withers et al. [8]. There are also numerous techniques that have been developed, based on the idea of discriminating the seismic signals from the background noise, but are mainly used for the automatic identification of the seismic phases (automatic picking) rather than event detection. Such well known methods are based on higher order statistics [9, 10] and wavelet analysis [11]. Moreover, hybrid methods which incorporate the wavelet transform with the Akaike Information Criterion or energy criteria with the polarization features of the seismic signals are presented in [12] and [13] respectively.

In this work we propose a new method, based on the estimation of an empirical distribution of the noise process as well as a residual based empirical distribution, in order to use a non strict hypothesis testing scenario for discriminating the recorded noise from the seismic signals.

2. PROBLEM FORMULATION

Let us denote with x_n, $n = 0, 1, \cdots, T-1$, the record from a given station and let us also assume that during the recording interval occurred an unknown number K, of seismic events. If we denote with s_n^k, $n = 0, 1, \ldots, T_k - 1$, the signal produced by the k-th event and with n_k the corresponding wave arrival time, then x_n can be expressed as:

$$x_n = \sum_{k=1}^{K} s_{n-n_k}^k + w_n, \tag{1}$$

where w_n is a realization of the noise process $\mathcal{W}_n$. The problem at hand is then the detection of the presence of the events in the record and the segmentation of the given record into signal and noise intervals. In the proposed methodology we assume that the record x_n is sparse. In addition, let us consider the set $\mathcal{T} = \{1, \cdots, T\}$, where T is the duration of the record, and its subset $\mathcal{N}$, containing all the time points n related with the noise intervals of the record, with $|\mathcal{N}| = N$ being its cardinality. Then, the set $\mathcal{E} = \mathcal{T} - \mathcal{N}$, will contain all the time points n related to the seismic events contained in the given record, with $|\mathcal{E}| = T - N$.

2.1. Events Detection as a Hypothesis Testing Problem

Let $\mathbf{x}_i$ be a block of length L of consecutive i.i.d. noise samples, drawn from a known probability density function $f_{\mathcal{W}}(w)$. Then the event detection problem can be formulated

*This work was financed by the University of Patras, "Karatheodori" research programs, entitled "*Developement of Automatic Seismic Data Processing Techniques for Application to Seismic Tomography,*" and "*The Relocation Problem of Seismic Event Hypocenter Parameters*".

as a simple, block based hypothesis testing one with the null and alternative hypotheses defined as follows:

$$\begin{aligned} \mathcal{H}_0 : \hat{f}_{\mathcal{W}}(w;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}}) &\sim f_{\mathcal{W}}(w) \\ \mathcal{H}_1 : \hat{f}_{\mathcal{W}}(w;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}}) &\not\sim f_{\mathcal{W}}(w) \end{aligned} \quad (2)$$

where $\hat{f}_{\mathcal{W}}(w;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}})$ constitutes an estimation of the probability density function of the RV $\mathcal{W}$, based on the histogram $\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}} = [h_1 \quad h_2 \quad \ldots \quad h_{M_{\mathcal{W}}}]$ of the data block $\mathbf{x}_i$, and $\mathcal{B}_{\mathcal{W}}$ is a $M_{\mathcal{W}}$-bins partitioning of the support of the *pdf* $f_{\mathcal{W}}(w)$. By considering, that $\mathcal{J}_L(\hat{f}_{\mathcal{W}}(.;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}})||f_{\mathcal{W}}(.))$ is a statistic whose statistical distribution under the null hypothesis is known (even asymptotically, i.e. when $L \to \infty$), it can be used to measure the discrepancy existing between the theoretical probability density function and an empirical counterpart. Indeed, we can set the desired level of significance α to decide between the hypotheses defined in Equ. (2) thus solving the hypothesis testing problem.

However, in real cases the above mentioned *pdf* is unknown and safe assumptions for the noise process except maybe wide sense stationarity and second order ergodicity, can not be made. In such cases, a possibility is the replacement of the unknown probability density function $f_{\mathcal{W}}(w)$ by its empirical counterpart $\hat{f}_{\mathcal{W}}(w;\bar{\mathbf{h}}_{\mathcal{B}_{\mathcal{W}}})$, where $\bar{\mathbf{h}}_{\mathcal{B}_{\mathcal{W}}}$ is the average of the histograms $\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}},\ i = 1, 2, \ldots$ resulting from data blocks drawn from the noise process. Note though that the estimation of the above mentioned empirical function from the given data record is not an easy problem. In addition, the statistical distribution of the modified sample statistic $\mathcal{J}_L(\hat{f}_{\mathcal{W}}(.;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}})||\hat{f}_{\mathcal{W}}(.;\bar{\mathbf{h}}_{\mathcal{B}_{\mathcal{W}}}))$ becomes unknown, and a strict hypothesis testing scenario can not be followed. In the next section we are going to propose a solution for the above mentioned problems.

3. THE PROPOSED SOLUTION

In this section we propose the use of a two stage procedure for the solution of the above mentioned problems. The first stage of the proposed procedure, accepts as input a record of seismic data and the noise based as well as the residual based empirical distributions of the background noise process are produced in its output. In the second stage, having obtained the aforementioned functions, a simple thresholding type test based on the statistic $\mathcal{J}_L(\hat{f}_{\mathcal{W}}(.;\mathbf{h}^i_{\mathcal{B}_{\mathcal{W}}})||\hat{f}_{\mathcal{W}}(.;\bar{\mathbf{h}}_{\mathcal{B}_{\mathcal{W}}}))$ is applied, for achieving the desired discrimination between “noise” and “signal” samples.

3.1. Estimation of Empirical Distribution

To begin with, let us concentrate ourselves on the first stage of the proposed procedure which is responsible for the estimation of the empirical *pdfs* of the noise itself as well as its whitened counterpart, from the given data record. According to Equ. (1) a seismic record constitutes the superposition of noise and a number of seismic signals that has occurred. Thus, initially a “sufficient” subset of the noise samples of the given record has to be isolated and then the desired estimation can be obtained by means of a detailed histogram. This specific stage consists of four steps, including sampling and modeling of a number of blocks from the given data record, clustering using a special algorithm based on the use of the correlation coefficient as a similarity measure, in order to isolate the “noise” part of the record, and finally the estimation of the desired empirical probability density functions.

3.1.1. Sampling

Let us consider the following set of I blocks of samples that have been drawn independently and with replacement from the given record $x_n, n \in \mathcal{T}$:

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_I\} \quad (3)$$

where the i-th element of set $\mathbf{X}$ is a block of length L of consecutive samples, that is $\mathbf{x}_i = [x_{n_i}, x_{n_i+1}, \ldots, x_{n_i+L-1}]^t$, with its start time index n_i being a uniformly distributed discrete random variable that ensures equiprobable selection of any L-length block from the $T - L + 1$ contained in the given record.

We must stress at this point that the choice of the number of blocks I is crucial in the sense that a“sufficiently” large I is needed to ensure that the proposed method will result to a good estimation of the noise distribution, while at the same time I must not be much larger than $T - L + 1$ to avoid the sacrification of the independence among them [14].

Having formed the set $\mathbf{X}$ of Equ. (3) our goal now is to identify the blocks that have been drawn from the noise process. This is exactly the goal of the second stage of the pipeline, where the use of data modeling for the reduction of the dimensionality of the problem at hand as well as for the whitening of the noise process is proposed.

3.1.2. Modeling

We are going to solve the modeling problem by using the maximum likelihood method. Thus, we denote with $f_{\mathcal{X}_i}(\mathbf{x}_i|\boldsymbol{\theta}_i)$ the conditional joint *pdf* of the random vector $\mathcal{X}_i$ represented by the observed sample vector $\mathbf{x}_i$ and $\boldsymbol{\theta}_i$ is a parameter vector whose structure and length are strongly related to the specific choice of model we select to use for the given data. In this work we propose the use of a P-th order autoregressive model denoted by $AR(\boldsymbol{\theta},\ P)$.

Given the length L block of data $\mathbf{x}_i$ we must estimate the parameters of the model. To this end we need to define an appropriate measure of the fit between the model and the data. This implies that (unless we have some prior information) the estimation procedure demands the use of a cost function for selecting the model order. Such a function which is based on the likelihood and the Kullback-Leibler Divergence measure, is the corrected Akaike Information Criterion (AICc) [15] which, for finite sample sizes and for a model of order P_i, is defined as follows:

$$AICc(P_i) = 2P_i - 2\ln(l(\boldsymbol{\theta}^*_i)) + \frac{2P_i(P_i+1)}{L - P_i - 1}, \quad (4)$$

where $\boldsymbol{\theta}^*_i$ is the maximum likelihood parameter estimator of length P_i, resulting from the solution of the following optimization problem:

$$\boldsymbol{\theta}^*_i = \arg\max_{\boldsymbol{\theta}_i} l(\boldsymbol{\theta}_i) = \arg\max_{\boldsymbol{\theta}_i} f_{\mathcal{X}_i}(\mathbf{x}_i|\boldsymbol{\theta}_i), \quad (5)$$

$l(\boldsymbol{\theta}^\star_i)$ is the achieved maximum of the likelihood function and L the number of the available data points.

We can now find out the optimal model for the given data block by solving the following optimization problem:

$$P_i^\star = \underset{P_i \in \mathbb{P}}{\arg\min}\{AICc(P_i)\} \tag{6}$$

where $\mathbb{P}$ is a subset of the set of natural numbers with its cardinality defining the maximum order of the candidate AR models.

Solving the above defined problem for each member of set $\mathbf{X}$ defined in Equ. (3) and by selecting $P^\star = \max_{i \in \{1,2,\cdots,I\}}\{P_i^\star\}$, we can form a $P^\star \times I$ matrix $\mathbf{\Theta}$ by placing the optimal parameters $\boldsymbol{\theta}_i^\star$, resulting from the modeling of the data block $\mathbf{x}_i$, in its i-th column, and compute its column standardized counterpart $\bar{\mathbf{\Theta}}^o$.

Our goal now is the grouping or clustering of the "similar" columns of the above defined matrix. That is, we have to solve a clustering problem. This is exactly the task implemented in the third step of the proposed procedure.

3.1.3. Clustering

It is clear that each column of parameters matrix $\bar{\mathbf{\Theta}}^o$ can be represented by a point in the Euclidean parameter space $\mathbb{R}^{P^\star}$. Moreover, because of their standardization all these points are located on the surface of the hypersphere $\mathcal{S}^{(P^\star-1)}$. Thus, in order to identify a representative cluster of the noise model, that is a dense subset of noise models located on a small area of the surface of the aforementioned hypersphere, let us denote by $\bar{\boldsymbol{\theta}}^o_{\text{med}}$ the standardized counterpart of the median vector, which can be evaluated from the medians of the $P^\star$ rows of matrix $\bar{\mathbf{\Theta}}^o$. Then, we can compute the following inner products:

$$\rho_i = < \bar{\boldsymbol{\theta}}^o_{\text{med}}, \, \bar{\boldsymbol{\theta}}^o_i >, \; i = 1, 2, \cdots, I, \tag{7}$$

and use them, to form the subset $\mathbf{\Theta}^\star$ of columns of matrix $\mathbf{\Theta}$ by using the following set of indeces:

$$\mathcal{Q} = \{i = 1, 2, \cdots, I \; : \rho_i \geq \cos(\phi)\} \tag{8}$$

where ϕ is a small angle[1] that we use in order to control the tightness of the noise cluster. Since ρ_i quantifies the correlation existing between the standardized median and the $i-$th parameter vector, we are expecting that the more its value tends to unity the more alike the parameter vector to the standardized median is, and thus its proximity to the unity can be used as a safe criterion to form the desired noise models cluster. Finally, because of the sparsity (i.e. $N >> T/2$) of the seismic record, the cardinality Q of set $\mathcal{Q}$ defined in Equ. (8) will be large enough, thus ensuring the critical size of samples needed for the estimation of the desired Empirical pdfs. As it is expected, the proposed procedure leads to a robust clustering technique that identifies succesfully the kernel of the desired noise cluster.

3.1.4. Empirical pdfs

Let us assume that from the I initial blocks $\mathbf{x}_i$, Q have been isolated by the above mentioned clustering procedure. Then, these Q blocks have been drawn from the noise process and an estimation of the desired empirical probability density function $\hat{f}_{\mathcal{W}}(.;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{W}})$ of the noise process, can be directly obtained by evaluating the mean of their respective histograms, that is:

$$\hat{f}_{\mathcal{W}}(w;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{W}}) = \sum_{m_w=1}^{M_w} \bar{h}_{m_w} \mathbb{1}_{\mathcal{B}_{m_w}}(w), \tag{9}$$

where $\bar{h}_{m_w}$ are the average relative frequencies with respect to each bin $\mathcal{B}_{m_w}$ of the partition of the support of $\hat{f}_{\mathcal{W}}(w;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{W}})$, for all Q blocks $\mathbf{x}_q$, $q = 1, 2, \cdots, Q$ that belongs to the noise process, and $\mathbb{1}_{\mathcal{B}_{m_w}}(w)$ is the indicator function of set $\mathcal{B}_{m_w}$. In addition, we can exploit the average model $\bar{\boldsymbol{\theta}}^\star$ resulting from the clustering step, in order to obtain a distribution based on the residuals after the whitening of the noise process blocks. Specifically, by filtering each noise process block $\mathbf{x}_q$, by the average model $\bar{\boldsymbol{\theta}}^\star$, whose $p-$th element is defined as follows:

$$\bar{\theta}^\star_p = \frac{1}{Q}\sum_{q=1}^{Q} \mathbf{\Theta}^\star_{pq}, \quad p = 1, 2, \ldots, P^\star, \tag{10}$$

we obtain Q blocks containing the residuals $\mathbf{r}_q$, that correspond to the noise process blocks. Following the same procedure as above, we can obtain an estimation of a residual based empirical pdf, that is:

$$\hat{f}_{\mathcal{R}}(r;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{R}}) = \sum_{m_r=1}^{M_r} \bar{h}_{m_r} \mathbb{1}_{\mathcal{B}_{m_r}}(r), \tag{11}$$

where $\bar{h}_{m_r}$ are the average relative frequencies with respect to each bin $\mathcal{B}_{m_r}$ of the partition of the support of $\hat{f}_{\mathcal{R}}(r;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{R}})$ for all Q blocks of residuals $\mathbf{r}_q$, $q = 1, 2, \cdots, Q$ and $\mathbb{1}_{\mathcal{B}_{m_r}}(r)$ is the indicator function of set $\mathcal{B}_{m_r}$.

What remains in order to complete the presentation of our technique, is a procedure for detecting the seismic events. This is exactly the goal of the next subsection.

3.2. A Thresholding type Statistical Test and Events Detection

As it has already been mentioned, the fact that the real distribution of the statistic $\mathcal{J}_L(\hat{f}_{\mathcal{W}}(.;\mathbf{h}^i_{\mathcal{B}_\mathcal{W}})||\hat{f}_{\mathcal{W}}(.;\bar{\mathbf{h}}_{\mathcal{B}_\mathcal{W}}))$ is unknown does not allow us to accept or reject the null hypothesis by setting a level of significance a. In order to overcome this obstacle, we define the following statistics:

$$\mathcal{S}_j(i) = \mathcal{J}_L(\hat{f}_{\mathcal{Z}}(.;\mathbf{h}^i_{{}_j\mathcal{B}_\mathcal{Z}})||\hat{f}_{\mathcal{Z}}(.;\bar{\mathbf{h}}_{{}_j\mathcal{B}_\mathcal{Z}})), \quad j = 1, 2, \tag{12}$$

where superscript i refers to the data block $\mathbf{x}_i$ and use them for solving the problem at hand. In order to achieve our goal, let us make now the following choices for the parameters of the above defined statistics:

$${}_1\mathcal{B}_\mathcal{Z} = \bigcup_{m_\mathcal{Z}=1}^{M_\mathcal{Z}} \mathcal{B}_{m_\mathcal{Z}}, \tag{13}$$

and ${}_2\mathcal{B}_\mathcal{Z}$ such that the following inequality be held:

$$\sum_{m_\mathcal{Z}=m_0}^{M_\mathcal{Z}-m_0} \bar{h}_{m_\mathcal{Z}} \int_{{}_2\mathcal{B}_{m_\mathcal{Z}}} \mathbb{1}_{{}_2\mathcal{B}_{m_\mathcal{Z}}}(x)dx \geq 1 - \alpha. \tag{14}$$

[1] In all experiments we have conducted ϕ was set to $\pi/180$, i.e. one degree.

Note that although ${}_1\mathcal{B}_\mathcal{Z}$ and ${}_2\mathcal{B}_\mathcal{Z}$ correspond to the same partitioning, it is clear that the later is a subset of the former. Specifically, given the level of significance α the value of the lower limit m_0 of the summation is selected such that the integral under the "truncated" pdf to be greater than or equal to the right hand side of Inequality (14). Finally, note that we have the possibility to use the noise or the residual based empirical functions in the evaluation of the above defined statistics by assigning to the RV $\mathcal{Z}$ the RV $\mathcal{W}$ or RV $\mathcal{R}$ respectively.

Let us now consider that with the aim of a sliding window of length L, we are evaluating the above defined statistics for each sample of the given data set, thus forming two sequencies. The values of the aforementioned statistics, are expected to be close to each other, when the data block $\mathbf{x}_i$ contains only samples drawn from the seismic noise. On the other hand, when data block $\mathbf{x}_i$ enters a seismic signal, the values attained by the statistics will be in general different and strongly related to the used level of significance α in Equ. (14). This relation is exactly what we propose to exploit in order to solve the detection problem at hand. In particular, the following simple thresholding scheme:

$$\mathcal{S}_1(i) > \max_{i \,\in\, [1\cdots T]} \mathcal{S}_2(i), i = 1,\ 2, \cdots,\ T \qquad (15)$$

is proposed for the detection of the seismic events, and this concludes the presentation of the proposed technique. In the next section we are going to evaluate its performance.

4. EXPERIMENTAL RESULTS

In this section we evaluate the performance of the proposed method by applying it in both synthetic and real seismic data sets. The synthetic seismic signals, described by Equ. (1), were modeled as low-pass filtered Gaussian noise, multiplied by a half-Gaussian window for the effect of amplitude shaping, and a constant gain, controlling the signal-to-noise ratio (SNR). In order to construct a data record, first the noise process w_n was created and then the synthetic signals s_n^k with the desired SNR value in the range $[0\ 10]$ were formed. Moreover, the onset times were obtained by random selection in the interval $[1\ T]$ and the resulting signal x_n was calculated, by using Equ. (1). We must stress at this point that, since the proposed test defined in Equ. (15) does not depend on the specific form of statistic $\mathcal{J}_L(\,.\,||\,.\,)$ used for its definition, any of the well known statistics that measures the discrepancy between two density functions can be used. In our experiments Pearson's Chi-Squared goodness-of-fit statistic [16] is used. Finally, we must say that in all experiments we are going to present, the level of significance α was set to 10^{-2} and the number of blocks, of length $L = 500$ samples each, that have been drawn independently and with replacement from each seismic record was $I = 500$.

4.1. Experiment I-Synthetic Data

In this first example we are going to evaluate the performance of the ingredients of the proposed technique by applying it to synthetic seismic data. To this end we constructed 500 records (10 min of duration each), of synthetic seismic data containing a total number of 2500 events, by using the above mentioned process generator with an ARMA(2, 2) process used as the noise process w_n in Equ. (1).

Aiming to test the modelling capabilities of the proposed technique, in the first part of this experiment, the input of the above mentioned ARMA(2,2) model was feeded by white Gaussian noise of unit power. Part of a synthetic record containing three events with SNR 5 is shown in Fig. 1.(a). The values of the parameters of the mean model resulting from the application of the modeling procedure described in Subsection 3.1.2 as well as their confidence intervals before(red-dotted line) and after the clustering step (green-dotted line), are shown in Fig. 1.(b). Moreover, a detailed histogram, that could be considered as the empirical pdf of the noise process, resulting by the blocks of the noise process isolated by the clustering technique proposed in Subsection 3.1.3, is presented in Fig. 1.(c). Note the departure of the empirical distribution from the normal one. This departure mainly happens because the samples of the noise process are correlated. Finally, in Fig. 1.(d) we can see a detailed histogram resulting from the residuals obtained after the whitening of the blocks of noise process. As we can see from this figure, the estimated histogram is perfectly matching the theoretical Gaussian function, thus revealing its effectiveness.

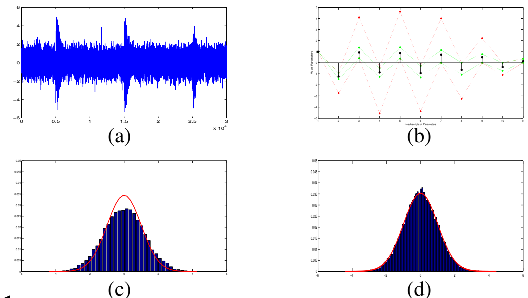


Fig. 1. A sample of a synthetic seismic signal (a), the parameters of the mean model (black) and their confidence values before (red-dotted) and after the clustering (green-dotted) step (b). Empirical pdfs of the noise process resulting from noisy blocks after the clustering step (c) and their whitened residuals (d) obtained from the application of the proposed technique. For comparison purposes, in both Empirical pdfs the Gaussian kernel (red) is superimposed.

Furthermore, we tested our technique by using the same synthetic data set but with the white Gaussian noise source replaced by a mixture of a gaussian and a uniform distribution, and the percentage of the successfuly detected events as well as the percentage of the false alarms were calculated. Only the cases where the estimated onset time was in a reasonable neighborhood [2] of the real one, were considered as successful detections.

Table I: Detection Rate (DR) & False Alarms (FA)

	Residuals		Noise	
SNR (dB)	DR (%)	FA (%)	DR (%)	FA (%)
0	91.6	8.8	92.4	10.7
2	97.2	1.5	98.2	3.8
5	100	0	100	2.0
10	100	0	100	0.5

The results we obtained for four different SNR values are contained in Table I. From the contents of the table, it is evident that in low and medium SNR values the noise based empirical pdf marginaly exhibits a better detection rate performance

[2] The neighborhood of a real onset time is considered as an interval of length $L/2$ samples ($L/4$ samples before and after that time) where L is the length of the used sliding window.

while in high SNR values both approaches seem to have an excellent performance. Finally, as it is clear, the noise based empirical pdf suffers from a higher number of false alarms for all SNR values.

4.2. Experiment II-Real Data

In this experiment we apply the proposed technique in real seismic data and its performance is compared against the STA/LTA technique, which is one of the most frequently used techniques by the seismologists. The real data set was comprised by 50 pre-cut records of continuously recorded seismic data, during a period of high seismicity. The "true" number of events, counted by an expert analyst, contained in the above mentioned 10 hours duration records were 500 with different amplitudes and durations. By using a window of length L = 200 samples (2 sec), the proposed detector and its rival were applied to the above data set and the results we obtained are contained in Table II. As we can see from the contents of Table II, the proposed method succeeded in identifying 489 and 486 seismic events with the use of residuals and the noise process respectively, and with the corresponding number of false alarms being approximately 22 and 30 respectively. On the other hand STA/LTA, even the special way we have treated it [3] , succeeded in identifying 417 seismic events and with the number of False Alarms increaced to 83.

These results reinforce the findings of Experiment I and confirm the appropriateness of the proposed technique for the problem at hand. This is also evident in Fig. 2 where the solutions to the detection problem obtained by the two rivals for a specific record of the real data set are shown. More specificaly, in Fig. 2.(a) curves of the statistics $\mathcal{S}_1(.)$ (black) and $\mathcal{S}_2(.)$ (green) in logarithmic scale, as well as the threshold value (dashed line) corresponding to the maximum value of the sequence $\mathcal{S}_2(.)$ defined in Equ. (12) are depicted. Finally, in Fig. 2.(b), the performance of STA/LTA algorithm on the same record with the empirically set threshold (dashed line) are illustrated. Note that in the specific example, although the detection performance of both rivals can be considered comparable, STA/LTA seems to suffer from the problem of false alarms.

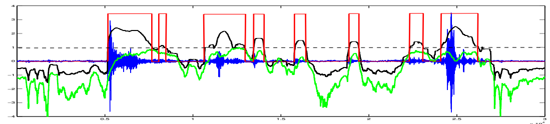
Table II: Detection Rate (DR) & False Alarms (FA)

Proposed (%)				STA/LTA (%)	
Residuals		Noise			
DR	FA	DR	FA	DR	FA
97.8	4.4	97.2	6.0	83.4	17.6

5. CONCLUSION

In this paper a new non strict hypothesis testing based method for the solution of the seismic events detection problem was proposed. The effectiveness of the proposed technique in identifying seismic events was confirmed from its application in a series of experiments both in synthetic and real seismic data sets. Although the advantages of the proposed technique against the widely used STA/LTA technique were demonstrated, the developement of a mechanism for controling the way the level of significance of the test affects the detection

[3]For each record the value of the needed threshold was selected by an expert analyst, such that the maximum possible number of detected events to be achieved while at the same time the number of False Alarms to be minimum.


(a)

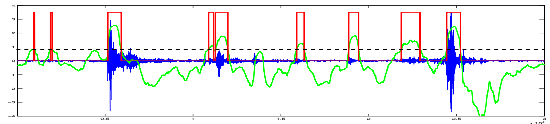

(b)

Fig. 2. Test and Detection curves in logarithmic scale achieved from the application of the proposed method (a) and STA/LTA (b) in a real data set. The threshold values for both algoritmhs are indicated by dashed lines (see the text for more details).

and the false alarms should be considered. This is an issue that is currently under investigation.

6. REFERENCES

[1] A. Trnkoczy, "Understanding and parameter setting of sta/lta trigger algorithm," *IASPEI, New Manual of Seismological Observatory Practice*, 2002.

[2] R. V. Allen, "Automatic earthquake recognition and timing from signal traces," *Bull. Seismol. Soc. Amer.*, vol. 68, pp. 1521–1532, 1978.

[3] R. V. Allen, "Automatic phase pickers: their present use and future prospects," *Bull. Seism. Soc. Am.*, vol. 72, pp. 225 – 242, 1982.

[4] M. Baer and U. Kradolfer, "An automatic phase picker for local and teleseismic events.," *Bull. Seism. Soc. Am.*, vol. 77, pp. 1437–1445, 1987.

[5] E. S. Earle and E. M. Shearer, "Characterization of global seismograms using an automatic-picking algorithm," *Bull. Seism. Soc. Am.*, vol. 84, pp. 366–376, 1994.

[6] N. Magotra N. Ahmed and E. Chael, "Seismic event detection and source location," *Bull. Seism. Soc. Am.*, vol. 77, pp. 958–971, 1987.

[7] B. O. Ruud and E. S. Husebye, "A new three-component detector and automatic single station bulletin production," *Bull. Seism. Soc. Am.*, vol. 82, pp. 221–237, 1992.

[8] M. Withers R. Aster C. Young J. Beiriger M. Harris S. Moore and J. Trujillo, "A comparison of select trigger algorithms for automated global seismic phase and event location," *Bull. Seism. Soc. Am.*, vol. 88, pp. 95–106, 1998.

[9] C. D. Saragiotis L. J. Hadjileontiadis and S. Panas, "Pai-s/k: A robust automatic seismic p phase arrival identification scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, pp. 1395–1404,, 2002.

[10] J. Galiana-Merino and S. Parolai, "Seismic p phase picking using a kurtosis-based criterion in the stationary wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, pp. 3815–3825, 2008.

[11] K. S. Anant and F. U. Dowla, "Wavelet transform methods for phase identification in three-component seismograms," *Bull. Seism. Soc. Am.*, vol. 87, pp. 1598–1612, 1997.

[12] H. Zhang C. H. Thurber and C. A. Rowe, "Automatic p-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings," *Bull. Seism. Soc. Am.*, vol. 93, pp. 1904–1912, 2003.

[13] T. Diehl N. Deichmann and S. Husen, "Automatic s-wave picker for local earthquake tomography," *Bull. Seism. Soc. Am.*, vol. 99, pp. 1906–1920, 2009.

[14] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.

[15] J. Cavanaugh, "Unifying the derivations of the akaike and corrected akaike information criteria," *Statistics and Probability Letters*, vol. 31, pp. 201–208, 1997.

[16] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philosophical Magazine*, vol. 5, pp. 157–175,, 1900.